2015

Assessing the Accuracy of Fish-Based Indicators of Biological Condition in Coastal Wetlands of the Great Lakes

Jeffrey D. Buckley
University of Windsor

Assessing the Accuracy of Fish-Based Indicators of Biological Condition in Coastal Wetlands of the Great Lakes

By

Jeffrey Buckley

A Thesis

Submitted to the Faculty of Graduate Studies
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

2015

Assessing the Accuracy of Fish-Based Indicators of Biological Condition in Coastal Wetlands of the Great Lakes

by

Jeffrey Buckley

APPROVED BY:

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

____
C.A.D. Semeniuk
Great Lakes Institute for Environmental Research

____
D.M. Higgs
Department of Biological Sciences

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

3 February 2015

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Assessing the quality of biological communities is important in the management of Great Lakes Coastal Wetlands. Biological indicator models can be used to quantify the condition of biotic communities. A number of biological indicators have been developed for use with fish communities in Great Lakes Coastal wetlands. The overall goal of this thesis was to assess the performance of various biological indicators in their ability to identify degradation in wetland fish communities. Biological indicators were assessed with respect to the disturbance gradient against which they was originally derived. Subsequently, the models' utility as diagnostic tools was assessed for use in identifying sources of anthropogenic stress. Overall, the Cooper-IBIs and Wetland Fish Indices demonstrated the highest classification accuracy, although factors such as their relative sensitivity and specificity, and the purposes for which they were originally designed should be taken into account when applying each indicator.

To my mother, Michel

& to my aunt, Donna

ACKNOWLEDGEMENTS

I would like to acknowledge and thank my supervisor Dr. Jan Ciborowski. He plucked me from obscurity at Environment Canada and has guided me through my thesis. I will be forever thankful for his generosity, his wisdom, and for always playing a support order when a move was justified.

I would also like to thank my committee members, Dr. Lucinda Johnson, Dr. Christina Semeniuk, and Dr. Dennis Higgs as well as members of the Great Lakes Environmental Indicators and Coastal Wetland Monitoring projects: Dr. Katya Kovalenko, Dr. Joe Gathman, Dr. Valerie Brady, Dr. Terry Brown, and Josh Dumke for offering their valuable input throughout this process.

I would like to thank the countless members of the Ciborowski Lab including fellow grad students Kaitlin, Sheeva, Kate, Jasmine, Chantal, and Kellie, as well as everyone else for all their help: Li, Michelle, Lyndon, Hannah, Julia, Angela, Jordan and Manasi.

My great field crew, Joe, Jasmine, Curtis, Justin, Celine, Jess, and Carla deserve endless thanks for making the long, hot, tiring, and smelly months spent in wetlands, vans, boats, and motels, fun and enjoyable.

Dr. Sue Watson gave me a chance and brought me back to the science I love. Without her guidance I wouldn’t be here.

To friends new and old: Jesse, Vince, Tom, Ryan, and Kyle, thanks for the good times.

I would like to thank my family: Mom and Laura, Aunt Donna, and my Nanny and Poppa. It is from their encouragement and love throughout my life that I was able to get here. And through their constant support that I could finish.

Finally, most importantly, Lauren. Your love humbles me and your determination inspires me. I could not ask for a more supportive, caring, and patient partner. Thank you for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS/SYMBOLS

AOC – Area of Concern

AUC - Area under the [ROC] curve

BRC - Biotic Response Curve

CWM - Coastal Wetland Monitoring

FCI - Fish Condition Index

GLEI - Great Lakes Environmental Indicators

GLWQA - Great Lakes Water Quality Agreement

IBI - Index of Biological Integrity, Index of Biotic Integrity

IEC - Index of Ecological Condition

MMI - Multimetric Index

pAUC - Partial area under the [ROC] curve

ROC - Receiver-Operating Characteristic

SAV - Submerged Aquatic Vegetation

WFI - Wetland Fish Index

WFI-AB - Wetland Fish Index - Abundance

WFI-PA - Wetland Fish Index - Presence/Absence

WQI - Water Quality Index

CHAPTER 1

GENERAL INTRODUCTION

The overall objective of this thesis is to evaluate indicator models of biological condition to determine which composite indices are able to accurately classify wetlands as being degraded. These models quantify the condition of a site, based on the biological community found there. Indicators of biological condition can be an important tool in the monitoring, assessment and restoration of biological communities and wetland sites in the Great Lakes. Few, if any, models have seen thorough testing and validations as accurate indicators of site condition.

Great Lakes Coastal Wetlands

Wetlands are an important part of the Great Lakes ecosystem. Coastal wetlands provide many valuable services including disturbance regulation, nutrient retention and cycling, commercial fishing, and recreation (Costanza *et al.* 1989, Costanza *et al.* 1997, Sierszen *et al.* 2012). Overall, wetlands in the Great Lakes provide approximately $2 billion in ecosystem services each year (Costanza *et al.* 1997, Sierszen *et al.* 2012). Additionally, coastal wetlands are ecologically important to multiple taxa including birds, mammals, amphibians, and invertebrates as habitat and breeding grounds (Sierszen *et al.* 2012). Coastal wetlands are particularly important to fish communities as seventy-five to ninety percent of Great Lakes fish species spend at least part of their life cycle in a coastal wetland (Brazner & Beals 1997, Sierszen *et al.* 2012). Fish can use coastal wetlands as a primary habitat, spawning and nursing habitat, and as a source of protection from predation (Jude & Pappas 1992).

Despite their economic and ecological importance, many Great Lakes coastal wetlands have experienced extreme degradation from anthropogenic activities. Since

European colonization of the Great Lakes, approximately 75% of wetlands in developed areas have been lost due to land drainage, commercial and industrial land use, dyking, and dredging (Whillans 1982). Wetland communities currently face risk of degradation from agricultural run-off, point-source pollution, shoreline development and other anthropogenic stressors (Danz *et al.* 2007). Due to the importance of the Great Lakes ecosystem and the threats posed to it, the governments of Canada and the United States developed the Great Lakes Water Quality Agreement (GLWQA) to address these issues (Canada and the United States 1972). The most recent amendment to the GLWQA calls for an increased emphasis on monitoring nearshore Great Lakes waters and for the development of indicators for use in monitoring (Canada and the United States 2012).

Indicators are tools used by scientists and local managers to assess to assess the condition of a site and have been applied to biological communities in coastal wetlands. A biotic community is commonly defined as being in good condition if it has experienced minimal or no impact from anthropogenic activity, that is, the condition of a community is defined by its 'naturalness' (Cains *et al.* 1993, Karr 1999, Davies & Jackson 2006, Stoddard *et al.* 2006, Hawkins *et al.* 2010). A community in poor condition is thus one that has been subject to significant impact from anthropogenic stress. Biological condition is therefore intrinsically tied to the condition of the environment and the effects of anthropogenic activity. Environmental Indicators of human disturbance of wetlands include altered water chemistry (for example, concentrations of nutrients such as phosphorus and nitrogen, or surrogate measures such as chlorophyll-a (algal biomass), chloride (ions) or dissolved oxygen (heterotrophic activity) (Uzarski *et al.* 2005, Seilheimer & Chow-Fraser 2006). These indicators have been used to assess the environmental expression of anthropogenic activity. However, reliance on chemical based indicators alone does not necessarily reflect the full effect of human activity on biological communities and the environment (Cairns *et al.* 1993). Biological indicators integrate the net effect of anthropogenic stressors and reflect the level of impact experienced directly by the biological community (Karr 1991, Karr 1999, Niemi *et al.* 2007, Cvetkovic & Chow-Fraser 2011). Finally, by establishing the connection between biological indicators and measures of the direct risk of anthropogenic impact (e.g. human land use [Danz *et al.* 2007]) management decisions can can address the specific causes of impairment affecting biological communities.

Biological Indicators

A biological indicator can be defined as a measurable aspect of the biotic community that changes predictably with changes in anthropogenic disturbance (Caro & O'Doherty 1999, McGeogh 1998, Heink & Kowerick 2010). A good biological indicator will: reliably and accurately change as a response to disturbance, distinguish between natural and anthropogenic disturbance, diagnose specific environmental stressors, be quick and easy to employ, and is easily interpretable by managers and end-users of the indicator (Kurtz *et al.* 2001).

A simple biological indicator may consist of a single metric; for example, the presence or abundance of individuals of an indicator species (McGeogh 1998, Heink & Kowarik 2010). However, Karr (1981) argued that a single metric may reflect the effects of only a particular environmental stress. According to Karr, the effects of multiple and synergistic stressors can be inferred by summing the scores of many individual metrics to create a multimetric index (MMI), which he called the Index of Biotic Integrity (IBI). In developing the original IBI, Karr (1981) used fish community data to generate a single composite score for a site that would indicate the site's condition from the perspective of the fish assemblage. These scores were based on an *a priori* decision of the community attributes (individual metrics) that distinguish a community in good condition from one in poor condition. The MMI approach has been greatly refined since the original Index of Biotic Integrity was proposed (Karr & Chu 1997, Schoolmaster *et al.* 2012). While Karr (1981) originally suggested that the IBI should be a composite of twelve fish community-based metrics, the US Environmental Protection Agency now lists over 60 possible fish-based metrics (Barbour *et al.* 1999). The MMI approach is now used throughout the world (Dos Santos *et al.* 2011, Shah & Shah 2012, Wu *et al.* 2012) as a model for developing biological indicators, and is the dominant approach in the United States (Reynoldson *et al.* 1997). Other approaches to deriving composite indices of biological condition have been developed in Australia (Smith *et al.* 1999), Canada (Reynoldson *et al.* 1995) and the United Kingdom (Wright *et al.* 1997) that use multivariate analyses to model community composition and integrate it into a single measure of biological condition. Collectively, MMIs, multivariate models, and other integrated measures of community composition can be referred to as biological indicator models.

The specific methods used to develop indicators can vary widely among different models. However, some steps are common in the development of all varieties of indicator. Firstly, variability in community composition due to natural environmental variation within the region of interest (e.g., the Great Lakes) must be taken into account. This is accomplished by identifying the primary drivers of natural variation in community composition across a region either by using best professional judgment (Karr & Chu 1997) or by evaluating the composition of communities at reference sites (Stoddard *et al.* 2006) and assigning homogeneous communities to groups across the range of natural variation. Secondly, the community composition is typically integrated into a single, composite value representing the biological condition of the community, either through combining individual metrics of community composition, or through multivariate analysis of community composition. Finally, the model as a whole is evaluated to determine whether it truly indicates the biological condition of a community. This testing requires applying the model to data from many sites and determining if the indicator score generated by the model accurately matches an independent measure of site condition.

Fish-Based Biological Indicator Models in the Great Lakes

Biological indicator models have been developed for Great Lakes wetlands using a variety of taxa including birds, invertebrates, diatoms and fishes (see Niemi *et al.* 2007, or Cvetkovic & Chow-Fraser 2001). Among these taxa, fishes are particularly useful as biological indicators of wetland habitats. Fishes can be sampled through electro-fishing, seine netting, fyke-netting or other commonly-used methods, and can be non-destructively identified in the field, permitting the community to be quickly evaluated. Most fish species in the Great Lakes spend a portion of their life in coastal wetlands and occupy many different niches and trophic levels (Jude & Pappas 1992). Therefore, the community composition will integrate effects of stressors that occur across different trophic levels

A number of fish-based biological indicator models have seen initial development for use within the Great Lakes. While some of these models are only applicable to littoral habitats (Minns *et al.* 1994), four models have been designed as indicators of Great Lakes coastal wetland quality: Indices of Biological Integrity developed by Cooper *et al.* (in review, based on Uzarski *et al.* 2005), the Fish Condition Index (Bhagat *et al.*, in prep.), the Wetland Fish Index (Seilheimer & Chow-Fraser 2006, 2007) and the Index of Ecological Condition (Howe *et al.* 2007a, 2007b).

Cooper Indices of Biological Integrity (Cooper-IBIs)

The Cooper-IBIs are multimetric indices of wetland condition. According to Wilcox & Meeker (1992), Uzarski *et al.* (2005), and Cvetkovic *et al*. (2010), emergent wetland plants are the primary factor driving natural variation in fish communities across the Great Lakes. Therefore, fish IBIs were developed by Cooper *et al.* under the framework of dominant plant zones. That is, individual IBIs were developed for areas sampled within a wetland that contained a single dominant plant type. Currently, there are published IBIs for bulrush (*Schoenoplectus* spp.), cattail (*Typha spp.*), lily (Nymphaceae) and submerged aquatic vegetation (SAV) – dominant zones (Cooper *et al*. in review.), hereafter referred to as the Bulrush-IBI, Cattail-IBI, Lily-IBI and SAV-IBI respectively and the Cooper IBIs collectively.

To develop these models, a list of 154 individual, candidate metrics were initially investigated for inclusion in each of the final Cooper IBIs. These metrics were based on fish community attributes including: total number of individuals caught, species richness, relative omnivore abundance (number of omnivorous individuals as a proportion of the total number of individuals caught) and Centrarchidae abundance, among others. This initial list of metrics was chosen based on best professional judgment of fish community characteristics expected in a community that is in good condition (i.e., equivalent to reference). Metrics were calculated in each zone of a wetland, based on the pooled catch of three fyke nets set overnight.

To test each individual metric for its relationship to human impact on fish communities, a disturbance scale ('gradient') was created as a surrogate for overall anthropogenic stress and was used as an independent variable against which to ordinate candidate individual metrics and the final IBIs. This scale consisted of a combination of land use (e.g., proportion of watershed area consisting of developed land, forest, etc.) in surrounding areas extending up to both 1 km and 20 km from the sampling location, as well as water quality (e.g. specific conductance, pH, turbidity, dissolved oxygen concentration, etc.). Each sampled site was then ordered according to its combined level of disturbance and assigned a rank-order score. All data were stratified by wetland plant zone. Therefore, each fish community metric was tested against the disturbance scale separately for each zone type and the rank-order disturbance scale was recalculated for each plant zones.

Individual metrics that correlated highly with the anthropogenic disturbance scale were retained for use within the final IBIs. Ranges of values for each metric were assigned metric scores such that a higher metric score implied greater biotic integrity. For example, for the metric 'total number of fishes caught' : < 10 = 0; 10-30 = 1; > 30 = 2. A list of individual metrics and their scoring criteria were then used as the final IBI, with a separate list of metrics used for each plant zone type. The summed scores for all metrics in an IBI therefore represent the biologically inferred condition of the fish community. All four Cooper-IBIs generate a score from 0 to 100 with a lower score indicating a poorer condition (see Appendix 3 for full IBIs).

Fish Condition Index (FCI)*:*

The Fish Condition Index is a multivariate measure of fish community composition that indicates community and wetland quality and is based on the Benthic Assessment of Sediment (BEAST) model of indicator development (Reynoldson *et al.* 1995). As a first step in its development, the FCI stratified groups of sites based on cluster analysis of fish assemblages found at minimally disturbed sites. Five distinct fish assemblages were identified in minimally disturbed reference sites across the Great Lakes (Bhagat *et al.* 2005). Discriminant function analysis (DFA) was used to identify the environmental variables that best characterize the sites for each of the five groups. Multiple regression was then used to develop a predictive classification model based on the DFA. This predictive model was used to classify additional (i.e., nonreference) sites and determine their membership in a specific group (i.e., expected reference assemblage) based on the naturally occurring environmental features of the site.

An anthropogenic disturbance scale was developed based on watershed land use (Host *et al*. 2005; Danz *et al*. 2007). Bray-Curtis two-endpoint ordination analysis was used to determine which species best characterized wetlands at the reference end of the disturbance gradient, and which species characterized wetlands at the degraded end. Finally, multiple regression was used to generate an equation predicting the Bray-Curtis ordination score for each environmental class of wetlands based on the relative abundance of key species designated by the Bray-Curtis ordination.

To assess the biological condition at a site, a wetland is first assigned to one of the five environmental classes based on environmental variables in the DFA model. The fish assemblage is sampled, and the relative abundances of the species caught are used in the predictive equation to generate a single value that represents inferred wetland condition. Assemblages that are more similar to the expected reference communities of a given group will produce a higher score, while communities that are more similar to the degraded communities will produce a lower score.

Wetland Fish Index (WFI)

The Wetland Fish Index is a fish-based measure of wetland condition that is meant to be applicable across all coastal wetlands within the Great Lakes. This index was developed by ordinating fish species relative abundances against a disturbance scale (Chow Fraser 2006). This scale, the Water Quality Index (WQI) was itself derived from an ordination of ten water quality variables (total phosphorus, nitrate, ammonia, etc.) that were considered indicative of anthropogenic disturbance. Each species was given a tolerance score between 1 and 5 based on its position on the disturbance axis. A score of 1 indicates a high tolerance for degraded water quality conditions while a score of 5 indicates a very low tolerance for degraded conditions. Additionally, each species was given a niche breadth score between 1 and 3, again based on its distribution across the disturbance axis.

To apply the WFI, tolerance, niche breadth and abundance values for each fish species found in a wetland are used to calculate the WFI composite score according to a formula detailed in Seilheimer and Chow-Fraser (2006; 2007). The final score generated by the WFI for a wetland is between 1 and 5, with a lower value implying more degraded conditions. The WFI has been developed for use based on either the abundance of individual fish species (WFI-AB) or the presence/absence of species (WFI-PA).

Index of Ecological Integrity (IEC)

The IEC, as described in Howe *et al*. (2007 a, 2007b), is a general, multivariate method for computing the expected biological condition of a community. Development of a specific IEC model consists of using biological community data, collected across a large number of sites, based on either the abundance or the presence/absence of species and ordinating the presence or abundance of each taxon along an anthropogenic disturbance scale. The disturbance scale is chosen *a priori* to the development and can be based on professional judgment, or a quantitative measure of human impact. A maximum-likelihood based approach is used to determine the expected 'biotic response' function for each species. The biotic response function relates the probability of observing a species to the site's condition (anthropogenic stress score) characterizes a species' tolerance to stress and the likelihood of observing the species at a given point on the disturbance scale.

The biotic response functions calculated during the initial development of the IEC can then be applied to a novel site to infer the site's biological condition. A second maximum-likelihood based approach is used to generate a score for the novel site based on the observed community composition and the previously derived biotic response functions. The score is a value between 1 and 10 with a higher number indicating a better biological condition.

IEC models were initially developed for bird communities, but can be applied to any taxon, or combination of taxa, so long as the sampling of sites is standardized between those that are used to develop the initial biotic response functions and those that are novel sites in which the user wishes to infer condition.

Overall Goals and Study Objectives

Although several fish-based biological indicator models have been published and are in common use, none have been validated through prospective sampling using a new and independent dataset. In this thesis, I test these models by validating their accuracy, determine which are suitable indicators of biological condition, and finally assess the ability of each model to diagnose specific types of disturbance.

The thesis consists of four chapters. In chapter 2, I comparatively assess each biological indicator model with respect to the data and disturbance gradient with which it was originally derived. While these models are intended to indicate overall anthropogenic stress, the measure of stress used in the development of each model differs. The Fish Condition Index and the Index of Ecological Integrity were calibrated with a purely land use based measure of stress, the Wetland Fish Index was initially calibrated with reference to a water quality-based scale, whereas the Cooper IBIs were calibrated with a scale consisting of a combination of land use and water quality to characterize stress. In testing against each model's original disturbance gradient I how reliably and accurately each model indicated wetland condition under its own terms.

In chapter 3 I tested each fish-based biological indicator model's ability to identify degraded wetlands based on a direct measure on human impact - land use. As well, I tested these indicator models as measures of various types of anthropogenic disturbance. A useful characteristic of an index is the ability to correctly diagnose specific types of anthropogenic stress. In testing indices against different types of stress I determined which models are best as diagnostic tools. Additionally, I used receiver-operating characteristic curve analysis to determine how models can be optimized for use as indicators of specific stressors.

Chapter four integrates the findings of the previous two chapters, recommends which indicator models should be implemented within the Great Lakes, and discusses how the validated models can be used in a monitoring and restoration capacity, as well as how they can be used to address other questions related to wetland ecology and conservation.

GENERAL METHODS

Data were collected through two multi-investigator, multi-year projects – the Coastal Wetland Monitoring Project and the Great Lakes Environmental Indicators (GLEI) Project. Each project extensively sampled wetlands throughout the Great Lakes. However, the primary purpose of the Coastal Wetland Monitoring Project is to monitor the environmental health of Great Lakes coastal wetlands, while the purpose of the GLEI Project is to develop and test indicators or environmental and biological condition, with a focus on relating biological condition to land use-based sources of anthropogenic impact. Similar sampling protocols for fish and water quality among these projects meant that the two sources of data could be used interchangeably.

Site Selection

A total of 254 coastal wetland sites were sampled between 2011 and 2013 (Appendix 1). All sites were sampled between June and September. Sites were selected according to a stratified random sampling design following the protocols of the Great Lakes Coastal Wetland Monitoring Project (Uzarski and Otieno 2008). All wetlands within the Great Lakes Basin has been identified using GIS coverages created by Albert and Simpson (2004) and Ingram and Potter (2004). Sites were stratified according to wetland type (riverine, lacustrine or barrier-protected), ecoregion (Olmernik 1987), and Great Lake. Each year, sites were randomly chosen from each stratum such that 20% of all sites in a stratum were sampled each year. The intention of this stratification was to ensure that sampling occurred across the range of geomorphic variation in the Great Lakes. Note that sites were not stratified by risk of anthropogenic impact (see GLEI methods below). All sites were at least 4 ha in area and had a close connection via surface water to a Great Lake. Sites were not sampled if they were deemed unsafe to reach from shore or by motorboat, on private property, or no longer existed due to lake level fluctuations.

Site Sampling

Upon arrival at a site, the distribution of plant morphotypes was assessed across the entire site. Morphotypes generally included plants classified by genus (e.g. *Typha* spp., *Schoenoplectus* spp., *Pontedaria* spp.), but also included more general categories such as submerged aquatic vegetation (SAV) and water lily (Nymphaeaceae spp.) (see Uzarski *et al.* 2005). Across the entire site, up to three dominant vegetation zones were identified. A zones was defined as an area within the wetland containing a single, monodominant morphotype. A single plant type had to have a coverage of at least 75 % of the zone to be considered monodominant. A vegetation zone could be represented by a single, large expanse, or by a collection of disconnected patches. However, one vegetative type had to cover a total area of least 400 m^2 to be considered large enough to warrant fish sampling. Physical characteristics of each wetland, including near shore land cover, shoreline structures (docks, gabions, groynes, etc.), and hydrologic connection to the main lake were recorded for each site.

Fish Sampling

Fish were sampled using large (3/16" mesh, 25' x 3' lead, 6' x 3' wings, 4' x 3' front opening) and small (3/16" mesh, 25' 1.5 ' lead, 6' x 1.5' wings, 3' x 18" front opening) fyke nets. Three nets were set overnight (12 - 24 h) in each vegetation zone. A zone was considered to have been successfully fished if at least two nets remained upright overnight and a total of at least ten fish were caught among the three nets set within a vegetation zone. If a zone was not successfully fished it was re-sampled the following night. Most fish were identified to species according to Holm *et al*. (2010) and released alive in the field. Fish that could not be identified in the field were preserved in ethanol and identified at a later date with the aid of a dissecting microscope. Some species, particularly young-of-year individuals, were identified to a higher taxonomic level. These included *Ameiurus* spp., *Notropis* spp., *Lepomis* spp., and *Lepisosteus* spp. Fish representing each species found at a site were also preserved in ethanol and kept as a voucher specimen according to the protocols of Portt *et al.* (2008). For larger species, photographs were taken of key identifying characteristics as vouchers. A random subset of up to 25 individuals per species at each net was measured for total length. Young-of-year status as well as any lesions or deformities were noted when caught. Zone fish counts were pooled counts among all nets in a zone (3). Catch per unit effort (CPUE) was defined as total catch per net per night.

Water Sampling

In situ water temperature, dissolved oxygen concentration, pH and specific conductance were recorded at each net using a multi-probe meter (YSI 556 MPS). Probes were calibrated before visiting each site. A water sample was collected into a 1-L acid-washed Nalgene container each net at at least 1 cm in depth. Samples were pooled together for each zone in an acid-washed 18-L plastic container. Samples were collected before nets were set and, where possible, from a boat to minimize disturbance of the substrate. Turbidity was measured approximately 3 – 5h after samples were collected using a Hach 2100Q portable turbidity meter. At that time subsamples were filtered and/or stabilized according to standard methods for the analysis of total phosphorus, total nitrogen, alkalinity, turbidity, chlorophyll-a, chloride, colour, soluble reactive phosphorus, nitrate, and nitrite (Uzarski *et al.* 2008) and stored refrigerated for later analyses. These analyses were performed within a week of collection by the Canadian National Laboratory of Environmental Testing (Burlington, Ontario).

Great Lakes Environmental Indicators Project (GLEI)

The GLEI Project was a collaborative initiative to assess biological condition across a human disturbance gradient (Niemi *et al.* 2007). The initial GLEI Project (GLEI 1, 2001 - 2003) collected data for the training and development of new biological indicator models, while the current GLEI Project (GLEI 2, 2011 - 2013) focused on assessing indicators and calibrating their relationship to human a disturbance scale (Niemi *et al.* 2007). Sampling for the GLEI project followed the same methods and protocols as the Coastal Wetland Monitoring Project, except for the following:

Site Selection

A total of 82 and 60 coastal wetland sites were sampled for GLEI 1 and GLEI 2, respectively (Appendix 2). Site selection was based on a stratified random sampling design as described in Danz *et al.* (2005). In short, wetlands within the Great Lakes were stratified by wetland type (riverine, lacustrine, barrier-protected, or high energy), ecoprovince (Keys *et al*. 1995) and by environment, representing potential sources of anthropogenic-based stress. Importantly, this design ensured sampling across the entire gradient of anthropogenic –based disturbance.

Site Sampling

Zonation of sites was based on the dominant near shore land use. Dominant land use was defined as a single land use type encompassing at least 10 % of the site's shore line. Land use types included residential, industrial, agricultural, and forested land. Up to three zones were identified at each site. Fish and water samples were collected along transects randomly placed in each zone extending from the shoreline. Fish were sampled with paired fyke net arrays at approximately 50 cm and 100 cm depth placed parallel to shore, for a total of 4 nets per zone. Fish were sampled along only a single transect and in up to 2 zones. Site-wide fish abundance estimates consisted of the counts of individuals pooled from all nets (2 – 8) at a site. Catch per unit effort was calculated as the total number of fish caught at a site, per net per night.

Water samples were collected along both transects at 50 cm and 100 cm depths in all zones. Samples at each depth in each zone were pooled for later analysis, resulting in 2 samples per zone.

Indicator Model Calculation

Biological indicator models developed using four independent methods were assessed in this study, the Fish Condition Index (FCI), the Wetland Fish Index (WFI), the Cooper-IBIs, and the Index of Ecological Condition (IEC). For a full description of indicator models, see the General Introduction.

Fish Condition Index

All FCI scores were calculated based on Bhagat *et al.* (in prep.). The FCI takes as input fish abundance data. Abundances were calculated as the total catch per net per night from all nets at a site and were converted to relative abundances, and $\log_2+1$ (i.e. octave) transformed. Site scores were calculated using the equation:

$$\text{FCI} = 0.616 + \sum_i^n A_i C_i$$

Where A is the relative abundance (octaves) and C is the coefficient for species i. C values are derived from a multivariate analysis of fish communities from both reference and degraded wetlands across the Great Lakes. In the initial development of the FCI, 5 unique fish community reference assemblages were found in the Great Lakes. Unique sets of species coefficients were derived for each of the assemblages. To determine which set of coefficients should be applied to a novel site, the expected assemblage can be determined based on the physical characteristics of the site. Currently, species coefficients are only available for the assemblage that corresponds to the Southern Superior Uplands (SSU) and Northern Great Lakes (NGL) ecoregions in Omernik (1987). In this study, data were not available for this ecoregion classification. However, the SSU and NGL correspond approximately to the Northern ecoprovince (Keys *et al*. 1995) classification used in the study. Therefore, the FCI was only applied to sites in the Northern ecoprovince in the GLEI 2 database.

Wetland Fish Index

All WFI scores were calculated using the methods found in Seilheimer & Chow-Fraser (2006, 2007). WFI scores are based on a weighted average equation:

$$WFI = \frac{\sum_{i=1}^{n} Y_i T_i U_i}{\sum_{i=1}^{n} Y_i T_i}$$

Where Y is either the log+1 transformed total abundance (WFI-AB) or the presence/absence (1/0; WFI-PA) of species i. U and T are constants determined for common species in the Great Lakes (see Appendix 4) and have been calculated based on a multivariate analysis of each species' response to a water quality based disturbance gradient. T represents a species' likelihood to be found across a range of disturbance (i.e. niche breadth), with a 3 signifying a very specific range and 1 signifying a broad range. U represents a species' tolerance to degradation, with a 1 signifying high tolerance and a 5 signifying low tolerance.

Cooper IBIs

Cooper-IBI scores were calculated from the methods described in Cooper *et al.* (in review), which are based on the methods of Uzarski *et al.* (2005). All IBI scores take as input fish abundance data. Abundance values were calculated as the pooled catch per net per night within a given vegetation zone (see General Introduction). Abundance values were used to calculate individual metrics within each zone. Metric values were then scored as either 0, 1, or 2. The final IBI score for a zone was the sum of all metric scores. The specific metrics used and scoring criteria for each metric are unique to each vegetation zone (see Appendix 3 for all metrics and scoring criteria). IBI score were calculated for Lily, *Typha*, Bulrush, and SAV zones.

Index of Ecological Integrity

IEC scores were calculated based on Howe *et al.* (2007a, 2007b). The IEC generates site scores using fish species abundance data and previously established Biotic Response Curves (BRC). BRCs represent a species' predicted occurrence across a disturbance gradient and are derived through a maximum likelihood-based analysis. Each set of BRCs is therefore meant to be applied to a specific population and as a measure of a specific disturbance gradient. No BRCs are currently available for Great Lakes wetland fish communities, so the GLEI 1 database was used as a training set to develop an IEC model for disturbance based on cumulative anthropogenic land use (Danz *et al.* 2007) in the Great Lakes.

All abundance data were calculated as the pooled catch per net per night for all nets at a site and were log +1 transformed. The 'est_brc()' function in the R package "IEC" was used to train BRCs. A measure of cumulative agricultural and urbanization-based land use (see chapter 2) was used as the disturbance gradient. The disturbance gradient was rescaled to a scale of 0 – 10 where the wetland in poorest condition in the training dataset was given a score of 0 and the wetland in best condition in the training dataset was given a score of 10. Separate BRCs were generated for the Northern and Southern ecoprovinces (Keys *et al.* 1995) of the Great Lakes. The 'est_iec()' function in the "IEC" package was used to generate final IEC based site scores. Site scores for both GLE 1 and GLEI 2 sites were generated using the BRCs derived from the GLEI 1 (i.e. training) database.

References

Albert, D. A., J. Ingram, T. Thompson, and D. Wilcox. 2003. Great Lakes coastal wetlands classification. Pages 239–246 Great Lakes Coastal Wetland Monitoring Plan.

Bennett, N. D., B. F. W. Croke, G. Guariso, J. H. A. Guillaume, S. H. Hamilton, A. J. Jakeman, S. Marsili-Libelli, L. T. H. Newham, J. P. Norton, C. Perrin, S. A. Pierce, B. Robson, R. Seppelt, A. A. Voinov, B. D. Fath, and V. Andreassian. 2013. Characterising performance of environmental models. Environmental Modelling & Software 40:1–20.

Bhagat, Y. 2005. Fish indicators of anthropogenic stress at Great Lakes coastal margins: multimetric and multivariate approaches. M.Sc. Thesis, University of Windsor.

Bhagat, Y., and J. J. H. Ciborowski. The Fish Condition Index (FCI). in prep.

Booth, D. B., and C. R. Jackson. 1997. Urbanization of aquatic systems: degradation thresholds, stormwater detection, and the limits of mitigation. Journal of the American Water Resources Association 33:1077–1090.

Booth, D. B., and L. E. Reinelt. 1993. Consequences of urbanization on aquatic systems - measured effects, degradation thresholds, and corrective strategies. Pages 545–550 Proc. of Watershed ’93. U.S. Environmental Protection Agency.

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.

Cairns, J., P. V. McCormick, and B. R. Niederlehner. 1993. A proposed framework for developing indicators of ecosystem health. Hydrobiologia 263:1–44.

Caro, T. M., and G. O’Doherty. 1999. On the use of surrogate species in conservation biology. Conservation Biology 13:805–814.

Chatfield, C. 1995. Model uncertainty, data mining and statistical inference. Journal of the Royal Statistical Society. Series A. 158:419.

Cooper, M. J., G. A. Lamberti, A. H. Moerke, C. R. Ruetz III, D. A. Wilcox, V. J. Brady, T. N. Brown, J. J. H. Ciborowski, J. P. Gathman, G. P. Grabas, L. B. Johnson, and D. G. Uzarski. An expanded fish-based index of biotic integrity for Great Lakes coastal wetlands. in review

Costanza, R., R. D’Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O’Neill, J. Paruelo, R. G. Raskin, P. Sutton, and M. van den Belt.

1997. The value of the world's ecosystem services and natural capital. Nature 387:253–260.

Costanza, R., S. C. Farber, and J. Maxwell. 1989. Valuation and management of wetland ecosystems. Ecological Economics 1:335–361.

Croft, M. V., and P. Chow-Fraser. 2007. Use and development of the Wetland Macrophyte Index to detect water quality impairment in fish habitat of Great Lakes coastal marshes. Journal of Great Lakes Research 33:172–197.

Dale, V. H., and S. C. Beyeler. 2001. Challenges in the development and use of ecological indicators. Ecological Indicators 1:3–10.

Davies, S. P., and S. K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16:1251–1266.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38–49.

Hawkins, C. P., J. R. Olson, and R. A. Hill. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society 29:312–343.

Heink, U., and I. Kowarik. 2010. What are indicators? On the definition of indicators in ecology and environmental planning. Ecological Indicators 10:584–593.

Howe, R. W., R. R. Regal, J. Hanowski, G. J. Niemi, N. P. Danz, and C. R. Smith. 2007a. An index of ecological condition based on bird assemblages in Great Lakes coastal wetlands. Journal of Great Lakes Research 33(sp3):93–105.

Howe, R. W., R. R. Regal, G. J. Niemi, N. P. Danz, and J. M. Hanowski. 2007b. A probability-based indicator of ecological condition. Ecological Indicators 7:793–806.

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. 1991. Biological integrity: A long-neglected aspect of water resource management. Ecological Applications 1:66.

Karr, J. R. 1999. Defining and measuring river health. Freshwater Biology 41:221–234.

Keys, J.E., J., C. A. Carpenter, S. L. Hooks, F. G. Koeneg, W. H. McNab, W. E. Russell, and M. . Smith. 1995. Ecological units of the eastern united states: First approximation. Atlanta, GAA: U.S. Department of Agriculture, Forest Service.

Kurtz, J. C., L. E. Jackson, and W. S. Fisher. 2001. Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development. Ecological Indicators 1:49–60.

Lougheed, V. L., and P. Chow-Fraser. 2002. Development and use of a zooplankton index of wetland quality in the Laurentian Great Lakes basin. Ecological Applications 12:474–486.

McGeogh, M. A. 1998. The selection, testing and application of terrestrial insects as bioindicators. Biological Reviews 73:181–201.

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.

Mouton, A. M., B. De Baets, and P. L. M. Goethals. 2010. Ecological relevance of performance criteria for species distribution models. Ecological Modelling 221:1995–2002.

Niemi, G. J., J. R. Kelly, and N. P. Danz. 2007. Environmental indicators for the coastal region of the North American Great Lakes: Introduction and prospectus. Journal of Great Lakes Research 33:1–12.

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

Portt, C.B., G.A. Coker, N.E. Mandrak and D.L. Ming. 2008. Protocol for the detection of fish Species At Risk in Ontario Great Lakes Area (OGLA). Canadian Science Advisory Secretariat Research Document 2008/026.

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. Austral Ecology 20:198–219.

Robson, B., D. Hamilton, I. Webster, and T. Chan. 2008. Ten steps applied to development and evaluation of process-based biogeochemical models of estuaries. Environmental Modelling & Software 23:369–384.

Dos Santos, D. A., C. Molineri, M. C. Reynaga, and C. Basualdo. 2011. Which index is the best to assess stream health? Ecological Indicators 11:582–589.

Sierszen, M. E., J. A. Morrice, A. S. Trebitz, and J. C. Hoffman. 2012. A review of selected ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. Aquatic Ecosystem Health & Management 15:92–106.

Simon, T. P. 1998. Modification of an index of biotic integrity and development of reference condition expectations or dunal, palustrine wetland fish communities along the southern shore of Lake Michigan. Aquatic Ecosystem Health & Management 1:49–62.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16:1267–76.

Uzarski, D. G., T. M. Burton, and J. J. H. Ciborowski. 2008. Chemical/physical and land use/cover measurements. Pages 29–31 *in* J. T. M. Burton, J. C. Brazner and (Eds.) J. H. Ciborowski, G. P. Grabas, J. Hummer, J. Schneider, and D. G. Uzarski, editors. Great Lakes Coastal Wetland Monitoring Plan. Ann Arbor, MI.

Uzarski, D. G., and S. Otieno. 2008. Statistical design. Pages 20–30 *in* J. T. M. Burton, J. C. Brazner and (Eds.) J. H. Ciborowski, G. P. Grabas, J. Hummer, J. Schneider, and D. G. Uzarski, editors. Great Lakes Coastal Wetland Monitoring Plan. Ann Arbor, MI.

Verbyla, D. L., and J. A. Litvaitis. 1989. Resampling methods for evaluating classification accuracy of wildlife habitat models. Environmental Management 13:783–787.

Whillans, T. H. 1982. Changes in marsh area along the Canadian shore of Lake Ontario. Journal of Great Lakes Research 8:570–577.

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.

Zhou, X.-H., N. A. Obuchowski, and D. K. McClish. 2011. Statistical methods in diagnostic medicine. Second edition. John Wiley & Sons.

CHAPTER 2

VALIDATING FISH-BASED INDICATORS OF BIOLOGICAL CONDITION FOR GREAT LAKES COASTAL WETLANDS

Abstract

Assessing the quality of biological communities is important in the management of Great Lakes Coastal Wetlands. Biological indicator models can be used to quantify the condition of biotic communities. An essential step in model development is assessing the performance of the model and validating it with new, independent data. The Wetland Fish Index, Fish Condition Index, Index of Ecological Integrity, and the Cooper IBIs have seen initial development for use in Great Lakes coastal wetlands. However, none have been thoroughly tested. Using data collected by the Great Lakes Environmental Indicators (GLEI) and Great Lakes Coastal Wetland Monitoring (CWM) projects, biological indices were evaluated to determine which indicators were able to accurately classify fish communities as degraded or not degraded. Both regression analysis and receiver operating characteristic (ROC) curve analysis, indicated that the Wetland Fish Index – Abundance, Lily-IBI, and Cattail-IBI had high classification accuracy when evaluated according to the stressor scales against which they were calibrated.

Introduction

Coastal wetlands are ecologically important to many taxa, but are particularly important to fish communities (Jude & Pappas 1992, Brazner & Beals 1997, Sierszen *et al.* 2012). Despite their importance, coastal wetlands have seen large amounts of degradation over the past century (Whillans 1982) and are subject to continued disturbance from anthropogenic activity (Danz *et al.* 2007, Host *et al.* 2001, Hondrop *et al.* 2014).

Assessing the quality of biological communities is important in the management, and restoration of wetlands (Karr 1991, Reynoldson *et al.* 1995, Wilcox *et al.* 2002, Bailey *et al.* 2004). Biological indicator models quantify the condition of a whole community and can be used to infer the impact of anthropogenic stress on that community. Through evaluating the entire assemblage at a site, a model produces a single score (index value) argued to represent the biological condition. These indicator models are useful because they purport to generate easily interpretable and biologically relevant summary assessments suitable for wetland management that integrate information from the entire assemblage (Karr 1981, Cairns *et al.* 1993). In recent years a number of fish-based biological indices have seen initial development for evaluating the condition of fish communities in Great Lakes coastal wetlands, including The Wetland Fish Index (WFI, Seilheimer & Chow-Fraser 2007), Fish Condition Index (FCI, Bhagat *et al.* in prep.), Index of Ecological Integrity (IEC, Howe *et al.* 2007a, 2007b) and the Cooper IBIs (Cooper *et al.* in review).

An essential step in model development is to assess its performance, often called model validation (Hawkins 2006, Mouton *et al.* 2010). This exercise ensures that a model’s outputs accurately and reliably indicate the condition of the biological community. Despite their potential value, to date no fish-based biological indicator models have been thoroughly validated as measures of biological condition in the Great Lakes coastal wetlands. Using novel, contemporary data, I assessed the accuracy of these models and compared these findings to previous assessments of the models performed during their initial development

Assessing the Accuracy of Biological Indicators

An important aspect in the performance of a biological indicator model is the accuracy with which it indicates the biological condition of a community (Hawkins 2006). To assess the accuracy of a model, it must be compared to a standard measure of the biological condition. Biological community condition as first defined by Frey (1977), and later adopted by others (e.g. Karr & Dudley 1981, Hughes *et al.* 1998, Hawkins 2006, Stoddard *et al.* 2006) is "the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region". An accurate model should indicate these qualities when applied to a community. However, there is no standard, quantifiable measure of community quality with which to evaluate the accuracy of biological indicators. Therefore, surrogate measures of community condition are used that reflect the ‘naturalness’ (Stoddard *et al.* 2006) of the habitat (Hawkins 2006, Yates & Bailey 2010).

Common surrogates for community condition are either anthropogenic stressors present at a site or measures of the potential risk of stress at a site, and are expressed as a disturbance scale (or ‘gradient’). Stressors are often measured in terms of the water chemistry-based water quality. Their components include measures of nutrients (e.g. total phosphorus & total nitrogen), chloride, and turbidity (Chow-Fraser 2003, Seilheimer & Chow-Fraser 2006, Seilheimer & Chow-Fraser 2007). Risk of stress in wetlands can be measured in terms of the human activity in the surrounding area. These measures include agricultural land use, urbanization, and point sources of pollution (Danz *et al.* 2007, Host *et al.* 2001, Niemi *et al.* 2011). Within the context of fish-based coastal wetland indicators, all currently used indices have been developed using different surrogates. Seilheimer & Chow-Fraser (2006, 2007) used a purely water quality-based measure when developing the WFI; Bhagat *et al.* (in prep.) and Howe *et al.* (2007a, 2007b) used different land use-based measures when the developed the FCI and IEC, respectively; and Cooper *et al.* (in review) used a combination of both water quality and land use based measures when developing the Cooper IBIs.

Biological indicator models are commonly evaluated through regression analysis, with the Coefficient of Determination (R^2) used as the performance criterion. To date, R^2 is the only criterion that has been used to evaluate fish biological indicators for Great Lakes coastal wetlands (Uzarski *et al.* 2005, Seilheimer & Chow-Fraser 2006, Bhagat *et al.* 2007, Seilheimer & Chow-Fraser 2007). Regression analysis calculates the least-squares relationship between biological indicator scores (the dependent variable) and scores along a disturbance gradient (the independent variable). The coefficient of determination is a measure of the precision of this relationship. However, the end-use of biological indicators by managers is not to determine the precise amount of disturbance at a site, but to determine whether or not it is degraded. For example, the European Water

Framework Directive (Herring *et al.* 2010) calls for monitoring of water resources based on classes of water quality (e.g. high, moderate, poor). Therefore, biological indicators should be assessed in terms of their classification accuracy (Dos Santos *et al.* 2011) rather than their goodness-of-fit to a regression equation.

The purpose of this study was to validate the accuracy of the Wetland Fish Index, the Fish Condition Index, the Index of Ecological Integrity and the Cooper IBIs. In order to give the fairest assessment, this study evaluated each indicator using the environmental disturbance gradient scale with which it was originally developed. Specifically, this study asks which fish-based biological indicator models are able to accurately classify communities as being degraded by assessing model classification accuracy using receiver operating characteristic curve analysis.

Methods

See General Methods (Chapter 1) for all data collection and indicator calculation procedures.

Datasets Used in Model Assessment

A biological indicator model is evaluated throughout its entire cycle of development (Mouton *et al.* 2010). A *training dataset* is used to initially create the model (e.g. to find individual metrics in the case of IBIs (Uzarski *et al.* 2007), or to ordinate communities using multivariate methods (Reynoldson *et al.* 1995)). Once the final model is developed, it is validated against a gold standard using a *test dataset*. The means by which data are partitioned into training and test datasets can affect the measured accuracy of the model when it is evaluated (Fielding & Bell 1997). To date, most published models relating fish biological indices to anthropogenic stress in Great Lakes wetlands have been assessed only by resubstitution (i.e., the same dataset was used for both the training and test data) rather than through true validation, thus overestimating the accuracy (Fielding & Bell 1997). Ideally, a model is evaluated with reference to a dataset that has been collected separately and is independent of the training dataset (Fielding & Bell 1997).

Training Data sets

In this study, *training datasets* refer to data used in the original creation of a particular indicator model (Table 2.1). Training data for the WFI-AB and WFI-PA indices were obtained from Seilheimer & Chow-Fraser (2007) (n = 100). Training data for IEC-North (n = 78), IEC-South (n = 58), and FCI (n = 73) were taken directly from the Great Lakes Environmental Indicator (GLEI 1) database (T. Brown, Natural Resources Research Institute, University of Minnesota Duluth, pers. comm.). For the Cooper-IBIs the training dataset was provided by M. Cooper (Central Michigan University, pers. comm.) *et al.* using the Great Lakes Coastal Wetland Monitoring (CWM) database. One-half of the samples from each vegetation zone had been randomly designated by Cooper *et al.* as training sites (lily: n = 54; bulrush: n = 58; SAV n = 60; cattail n = 37). These samples were used as the training dataset for the current study. The remaining samples were used as the test dataset. Training data for the WFI, FCI and IEC were collected from 2001 – 2005. Training data for the Cooper-IBIs were collected from 2011 – 2013.

Test Datasets

In this study, *test datasets* refer to data used to evaluate indicator models that had not been used during the training of the model (Table 2.1). For the WFI, the entire GLEI 2 database (n = 60) was used as the test dataset. All sites in the GLEI 2 database from the Northern ecoprovince (Keys *et al.* 1995) were used as the test dataset for the FCI and IEC-North. All sites in the GLEI 2 database from the Southern ecoprovince (Keys *et al.* 1995) were used as the test dataset for the IEC-South. For the Cooper-IBIs the test dataset was generated by Cooper *et al.* using the CWM database. One-half of the samples from each vegetation zone were randomly assigned by Cooper *et al.* as test sites (lily: n = 54; bulrush: n = 58; SAV n = 59; Typha n = 35). The same sites assigned by Cooper *et al.* for testing were used as the test dataset for the current study. All test datasets were collected between 2011 and 2014. Test datasets for assessment of the Cooper IBIs were collected contemporaneously with training datasets. Test datasets for the WFI, FCI and IEC were collected approximately 6 – 10 years after training datasets.

Disturbance Scales

Water Quality Index (WQI)

The WQI is a measure of water quality developed by Chow-Fraser (2006) and was designed as a measure of anthropogenic impact on wetlands. This index was derived from water chemistry data sampled from 110 wetlands across each of the five Great Lakes and has been shown to be significantly correlated with forested land cover and negatively correlated with altered (agriculture + urbanization) land cover. Twelve water quality variables were used to generate the index, they were: turbidity, temperature, pH, conductivity, chlorophyll-*a*, total suspended solids, total inorganic suspended solids, total phosphorus, soluble reactive phosphorus, total ammonium nitrogen, total nitrate, and total nitrogen. A Principal Components Analysis (PCA) was used to ordinate wetlands with respect to water quality. A final water quality score was generated for each wetland using the PC site scores for each of the twelves axes found in the original PCA. The score for an individual site was the sum of its PC scores, with each score weighted by the proportion of explained variable for each axes based on the eigenvalue. Finally, a stepwise multiple regression was used to generate predictive equations that can produce WQI scores based on raw water quality data. In total, 9 predictive equations, each using a different combination of water chemistry variables were found to correlate highly with the original PC derived scores. In this study, WQI #8 was used. The formula for WQI 8 is:

$$0.523 - (0.832 * \log \text{total phosphorus}) - (0.313 * \log \text{total nitrogen}) - (0.983 * \log \text{conductivity}) - (0.583 * \log \text{chlorophyll-}a).$$

The WQI generally produces scores between 3 and -3. Higher scores indicate better water quality, and lower scores indicate worse water quality. It is suggested by Seilheimer & Chow-Fraser (2007) that a score of < 0 denotes a degraded wetland.

RankSum

RankSum is a relative measure of overall anthropogenic disturbance developed by Uzarski *et al.* (2005) and Cooper *et al.* (in review) and is based on both land cover and water quality measures of anthropogenic impact. These variables include land use in both a 1-km and 20-km buffer from the site, turbidity, chlorophyll *a*, total phosphorus, soluble reactive phosphorus, total nitrogen, ammonium-N, nitrate-N, dissolved oxygen, pH, specific conductivity and the first principal component score based on a principal components analysis of all disturbance variables. RankSum is not an absolute measure of anthropogenic disturbance, but instead only shows the relative amount of impact found among a defined group of wetlands and therefore must be recalculated for each dataset to which it is applied. To calculate RankSum, each individual disturbance variable is ranked among all sites in a given dataset, then the rank values are summed for each wetland and rescaled to a range of 0 to 100, with higher values indicating less anthropogenic disturbance and lower values indicating higher anthropogenic disturbance.

For all analyses in this study, RankSum values were recalculated for each biological indicator being assessed (e.g. RankSum was only calculated for Lily-IBI test sites when assessing the Lily-IBI and independent RankSum scores were calculated when assessing the Cattail-IBI test sites).

Land Use (Agriculture, Development, Cumulative, SumRel)

Anthropogenic land use is a direct measure of the risk of anthropogenic stress within the watershed surrounding the wetland. These measures were derived by the Great Lakes Environmental Indicators project and are based on previous research by Danz *et al.* (2005) and Host *et al.* (2011). In a Principal Components Analysis of over 200 individual measures of land-based anthropogenic activity (e.g. Road density, row crop count, forested land cover, residential land cover, etc.) Danz *et al.* and Host *et al.* found that the dominant stressors were: percent agricultural land cover, percent urban land cover, road density, and population density. These variables represented the two primary axes of anthropogenic stress to wetlands: agriculture and urbanization/population pressures. In the current study, percent agricultural land cover within the surrounding watershed was used as a measure of agriculture-based stress. Percent developed (i.e. urban) land cover within the surrounding watershed was used as a measure of urbanization/population pressures.

A measure of cumulative anthropogenic land use was generated by combining both agricultural and development based stressors. Development based land use was summarized in each watershed as the maximum normalized value of percent urban land cover, road density, and population density. For each watershed, the Euclidean distance from the origin of the development and percent agriculture scales was used to combine scores according to the two axes into a single measure of overall anthropogenic land use. Finally, a second measure of overall land use, SumRel, was derived from these data. SumRel is the sum of standardized and normalized land use stressors (agricultural land cover, residential land cover, road density, population density, and point source discharge).

All land use variables (percent agriculture, percent development, cumulative land use, and SumRel) were scaled from 0.0 to 1.0, with 1.0 being the maximum observed value across the entire Great Lakes basin. For all land use measures, higher values indicate greater anthropogenic disturbance, while lower values indicate less anthropogenic disturbance.

The Degraded Condition

To determine observed degradation status, each disturbance scale was divided into two portions, corresponding to nondegraded (little disturbance) and degraded (much disturbance). The value of the disturbance gradient that divides degraded sites from non-degraded sites was called the 'disturbance threshold'. For the WFI disturbance threshold values were taken from Seilheimer & Chow-Fraser (2007), while Bhagat *et al.* (in prep.) provided thresholds for the FCI. Explicit disturbance thresholds were not identified by the developers of the IEC or Cooper IBIs. Therefore, an operational definition of a degradation boundary was used. Disturbance thresholds were defined by the 20 % of wetlands exhibiting the greatest amount of disturbance within the training dataset (IEC and Cooper-IBIs). For example, if 100 wetlands were sampled in the training dataset, the disturbance scale value that corresponded to the wetland rank 80^{th} (from least to most disturbed) was used as the threshold for degradation.

Statistical Analysis

All previous evaluations of the performance of models assessed in this study have used linear regression, with R^2 (Coefficient of Determination) as the performance criterion (Seilheimer & Chow-Fraser 2007, Howe *et al.* 2007a, 2007b, Cooper *et al.* in review, Bhagat *et al.* in prep.). To compare current test datasets with previous analyses, linear regression analysis was performed using the lm() function in the base package of the R Statistical Programming Language (R Core Team 2014) on test datasets. All R^2 values were deemed significant at $p < 0.05$.

Receiver Operating Characteristic (ROC) curve Analysis

Receiver Operating Characteristic (ROC) curve analysis was used to assess classification accuracy (Fawcett *et al.* 2006) of biological indicators. While this method is commonly used in medical research (Park *et al.* 2004, Obuchowski *et al.* 2005) and machine learning (Hand & Hill 2001, Fawcett 2006), only recently has it seen use in ecology (Mouton *et al.* 2010, Liu *et al.* 2011) and specifically with biological indicators (Dos Santos *et al.* 2011).

Biological indicator models were evaluated on their ability to accurately classify sites as being in the degraded condition. To do this, the predicted degradation status (biological condition model score) was compared to the observed degradation status (amount of environmental stress) of each wetland site to assess the classification accuracy of models. The observed degradation status was determined by an anthropogenic disturbance scale (see the degraded condition above).

In this classification scheme there are four possible outcomes from an indicator's classification of a wetland (summarized as a confusion matrix, Fig. 2.1). If a wetland is degraded according to its score on the environmental disturbance scale and the biological index score derived from an indicator model also classifies the site as degraded then this is a true positive result; if the wetland is not degraded according to the disturbance scale and the biological indicator model classifies it as degraded then it is a false positive result; if the wetland is not degraded according to the disturbance scale and the indicator model classifies the site as not degraded it is a true negative result and if the site is not degraded but the indicator says it is degraded then it is a false negative result.

When a set of wetlands are assessed, there are a number of criteria that can be used to evaluate the classification accuracy of a model given the number of sites that fall into each of these four categories. Two of the primary measures of accuracy are sensitivity and specificity.

Sensitivity is the ability of a model to correctly identify a degraded site. That is, the proportion of sites that are degraded according to the disturbance scale, that the biological indicator model classifies as degraded (A / [A+C], Fig. 2.1); A highly sensitive indicator is more likely to tell the user that a site is degraded. Specificity is the ability of a model to correctly identify a non-degraded site, calculated as the proportion of sites that are not degraded according to the disturbance scale, that the indicator says are not degraded (D / [B+D], Fig. 2.1); A highly specific model is more likely to not say a site is degraded. Both sensitivity and specificity are important metrics when evaluating the accuracy of an indicator. A useful indicator will have both high sensitivity and specificity, however it is possible for an indicator to be very sensitive but not specific, and vice versa.

The output for most biological indicator models is not an explicit classification of sites as degraded or not degraded, but a continuous score. To evaluate the classification accuracy of an indicator using just a confusion matrix (as in Fig. 2.1), a cut-point would need to be set for the indicator which delineates what the model classifies as degraded vs. non-degraded. ROC analysis evaluates classification accuracy, but does not require a single cut-point to be specified. Instead, in ROC analysis the sensitivity and specificity for all possible cut-points are determined. The sensitivity/specificity of each cut-point of the model is plotted to generate the ROC curve (e.g. Fig. 2.4a.). By convention, the sensitivity is plotted on the y-axis and specificity plotted (and reversed) on the x-axis. The overall accuracy of the model can then be quantified with the area under the curve (AUC). ROC analysis is also useful as it provides a graphic representation on both the sensitivity and specificity of a model; a ROC curve that encompasses more of the left side of the plot is more specific and minimizes the likelihood of committing a Type II error, while a curve that encompasses more of the top of the plot is more sensitive and minimizes likelihood of committing a Type I error.

All ROC curve analysis was performed using the package "pROC" (Robin *et al.* 2011) of the R Statistical Programming Language (R Core Team 2014). The Area Under the ROC Curve (AUC) was used as the performance criterion for classification accuracy. An AUC of 1.00 signifies a model with perfect classification accuracy while an AUC of 0.500 signifies a model that is no better at classifying sites than random chance. Confidence intervals of AUCs are generated using the DeLong method described in DeLong *et al.* (1988) and Robin *et al.* (2011).

Results

Disturbance Thresholds

For the Water Quality Index (WQI, Seilheimer & Chow-Fraser 2006, 2007), a value of 0 was used in the evaluation of the Wetland Fish Index – Abundance (WFI-AB) and the Wetland Fish Index – Presence/Absence (WFI-PA) as the disturbance threshold. Sites with a WQI value lower than this value were designated as degraded (Table 2.1.).

For SumRel (Bhagat *et al.* in prep.), a value of 0.745 was used in the evaluation of the Fish Condition Index (FCI) as the disturbance threshold. Sites with a SumRel value higher than this were designated as degraded (Table 2.1.).

For Cumulative Land Use, a value of 0.385 was used in the evaluation of the IEC-North as the disturbance threshold while a value of 0.755 was used in the evaluation of the IEC-South. In both cases sites with higher Cumulative Land Use values were designated as degraded (Table 2.1.).

For RankSum, values of 16, 24, 41, and 42 were used as the disturbance threshold in the evaluation of the Lily-IBI, Cattail-IBI, Bulrush-IBI, and SAV-IBI, respectively. In all cases sites with lower values than this were designated as degraded (Table 2.1.).

Regression Analysis:

For all models, except the FCI (Fig. 2.2c, $R^2 = 0.05$, $p = 0.944$, $n = 38$), a statistically significant linear relationship was observed between the predicted condition (biological index score) and the observed condition (disturbance scale score) against which each of the models were regressed. The Lily-IBI showed the highest overall precision (Fig. 2.2f, $R^2 = 0.597$, $p = <0.001$, $n = 54$) while the SAV-IBI showed the lowest precision (Fig. 2.2i, $R^2 = 0.133$, $p = <0.01$, $n = 59$). Overall, the order of indices, ranging from most precise to least precisce was Lily-IBI > Cattail-IBI > WFI-AB > Bulrush-IBI > WFI-PA > IEC-South > IEC-North > SAV-IBI > FCI (Fig. 2.2). Most R^2 values for biological indicator models in the test dataset were lower than in their respective training dataset (Table 2.2). Only the Lily-IBI and SAV-IBI showed an increases in R^2.

Classification Accuracy:

With the original training datasets (Fig. 2.3), the IEC-South and Cattail-IBI had the highest classification accuracy with AUCs of 0.961 and 0.955, respectively, while the SAV-IBI showed the lowest overall classification accuracy with an AUC of 0.694. All other models were also highly accurate, with AUCs of 0.800 or greater. The order of classification accuracy for models based on training datasets was IEC-South > Cattail-IBI > Bulrush-IBI > IEC-North> WFI-PA > WFI-AB > FCI > Lily-IBI > SAV-IBI (Table 2.3).

When the test datasets were assessed (Fig. 2.4), the Lily-IBI and the Cattail-IBI exhibited the highest overall classification accuracy, with AUCs of 0.901 and 0.847, respectively. The FCI and IEC-South were the least accurate, with AUCs of 0.640 and 0.612, respectively. Overall, the order of classification accuracy for all models in the test dataset was Lily-IBI > Cattail-IBI > Bulrush-IBI > WFI-AB > IEC-North > WFI-PA > SAV-IBI > FCI > IEC-South (Table 2.3).

Discussion

Although a large body of research exists describing the development of biological indicators (Karr 1981, Whittier *et al.* 2007, Aparicio *et al.* 2011 Grabas *et al.* 2012, Wu *et al.* 2012) including many specifically for Great Lakes communities (Minns *et al.* 1994, Burton *et al.* 199, Uzarski *et al.* 2005, Niemi *et al.* 2007), relatively few models have been validated with independent data (Bhagat *et al.* 2007, Bailey *et al.* 2014, Strachan *et al.* 2014). Validation is an essential step in model development (Fielding & Bell 1997) to ensure that the model is generalizable independently of the data with which it was originally developed.

In this study, the accuracy of nine biological indicator models of the condition of fish communities in Great Lakes coastal wetlands was assessed. Accuracy was defined as the ability of a model to correctly classify sites as being in the degraded condition or not. Classification accuracy of the models of nearly all biological indicators was high when they were assessed with training datasets (Fig. 2.3, Table 2.3). Only one model, the SAV-IBI, had an AUC value <0.700, while the IEC-South had near perfect classification accuracy. However, despite strong classification accuracy during model training, most models less accurately classified sites in the test dataset (Table 2.3).

No previous studies have directly tested the accuracy of the Cooper-IBIs. However these models were derived using the methods originally published by Uzarski *et al.* (2005). That paper reported a strong relationship between bulrush spp.-based and a cattail spp.-based IBI when assessed with a RankSum disturbance gradient. Further tests of these IBIs by Bhagat *et al.* (2007) demonstrated that these IBI models were sensitive to only particular types of stress. Recent application of the Uzarski-derived IBIs by Calabro *et al.* (2013) revealed no correlation between IBI scores and an independent disturbance gradient. They argued that the wetlands used to develop the original IBIs did not cover a wide enough range of the disturbance gradient and is not transferable to wetlands at or near the reference condition. While bulrush and cattail IBIs developed by Uzarski *et al.* were based on a limited number of sample sites (i.e. sites in only two lakes, with most sites clustered around three locations in those lakes), the Cooper-IBIs are based on data from across all five Great Lakes which may explain their better performance when applied to independent data.

The current study is the first application of the IEC method to fish communities in the Great Lakes. To date, the IEC has only been used as a measure of the condition of bird assemblages (Howe *et al.* 2007a, 2007b). In these previous studies, IEC scores based on the condition of bird assemblages were strongly correlated with site condition when applied to a set of reserved test sites. Interestingly, when predicted site condition (IEC score) and observed condition (disturbance scale score) were plotted together, Howe *et al.* (2007a, 2007b) found a relatively tight cluster of points at the non-degraded end of the disturbance scale, but much greater variability at the degraded end. This is similar to the relationship found in the current study for both IEC-South (Fig. 2.2e.) and especially IEC-North (Fig. 2.2d). As IEC models are initially developed with respect to a single type of explicitly defined disturbance, this may suggest that fish assemblages found at non-degraded sites are similar, while the expected assemblage found at degraded sites is dependent on the type of stress impacting a site.

Despite strong classification accuracy and presicion when assessed with the training dataset, the FCI demonstrated the poorest performance when assessed with the test dataset. The FCI is meant to measure the similarity of a fish community to the community expected at a 'reference' site (i.e. a minimally disturbed site). However, Bhagat *et al.* (in prep.) identified five distinct reference assemblages, and proper application of the index requires first determining which reference assemblages a test site should be compared to. However, data were not available to properly classify sites as to their expected assemblage. Therefore, the low accuracy in site classification by the FCI may be the result of its improper application to Great Lakes wetlands.

Overall, the Cooper-IBI were found to have the highest classification accuracy when applied to a test dataset. However, comparisons between the performance of Cooper-IBIs and other biological indicators can only be made tentatively. Training and test datasets for the Cooper-IBIs are based on a partitioning of sites sampled across the same time period while there is a 6-7 year separation in training and test datasets for the IEC, FCI and WFI (Table 2.1). Validation of test sites using a partitioned dataset is cautioned against by Fielding & Bell (1997) as this may overestimate model accuracy. Cooper-IBIs are promising, but further assessment with truly independent data is warranted. Despite the time differences between training and test datasets, the WFI and IEC were still found to accurately classify sites as being degraded suggesting that these indicators are reliable and accurate measures of biological condition interannually.

Tables and Figures

Table 2.1. Databases used to generate training and test datasets, years during which data were collected, and disturbance scale used to construct and test each biological indicator model. The disturbance threshold is the amount of anthropogenic disturbance delineating degraded and non-degraded sites, in units of the disturbance scale. † Seilheimer & Chow-Fraser 2007.

Index	Training Dataset	Training Dataset Years	Test Dataset	Test Dataset Years	Disturbance Scale	Disturbance Threshold
WFI-AB	Original paper†	2001-2005	GLEI 2	2011-2013	WQI	0
WFI-PA	Original paper†	2001-2005	GLEI 2	2011-2013	WQI	0
FCI	GLEI 1	2002-2003	GLEI 2	2011-2013	SumRel	0.745
IEC-North	GLEI 1	2002-2003	GLEI 2	2011-2013	Cum. Land Use	0.385
IEC-South	GLEI 1	2002-2003	GLEI 2	2011-2013	Cum. Land Use	0.755
Lily-IBI	CM 2	2011-2013	CWM 2	2011-2013	RankSum	16
Cattail-IBI	CM 2	2011-2013	CWM 2	2011-2013	RankSum	24
Bulrush-IBI	CM 2	2011-2013	CWM 2	2011-2013	RankSum	41
SAV-IBI	CM 2	2011-2013	CWM 2	2011-2013	RankSum	42

Table 2.2. Coefficients of determination for training and test datasets of biological indicator models based on a linear regression analysis. All significant correlations ($p < 0.05$) are noted with an asterisk (*).

Index	Training R^2	Test R^2
WFI-AB	0.642*	0.342*
WFI-PA	0.669*	0.225*
FCI	0.446*	0.05
IEC-North	0.49*	0.2*
IEC-South	0.816*	0.213
Lily-IBI	0.481*	0.597*
Cattail-IBI	0.67*	0.478*
Bulrush-IBI	0.415*	0.327*
SAV-IBI	0.097*	0.133*

Table 2.3. Overall classification accuracy of biological indicator models measured as the AUC of ROC curves for training and test datasets. Bracketed values are the 95 % Confidence intervals of the AUC.

Index	Training AUC	Test AUC
WFI-AB	0.866 (0.79 - 0.942)	0.804 (0.66 - 0.948)
WFI-PA	0.874 (0.8 - 0.947)	0.767 (0.618 - 0.916)
FCI	0.826 (0.721 - 0.93)	0.640 (0.255 - 1)
IEC-North	0.895 (0.817 - 0.974)	0.779 (0.622 - 0.935)
IEC-South	0.961 (0.917 - 1)	0.612 (0.219 - 1)
Lily-IBI	0.815 (0.704 - 0.926)	0.901 (0.814 - 0.988)
Cattail-IBI	0.955 (0.892 – 1)	0.847 (0.686 – 1)
Bulrush-IBI	0.924 (0.847 - 1)	0.822 (0.652 - 0.993)
SAV-IBI	0.694 (0.507 - 0.882)	0.734 (0.592 - 0.876)

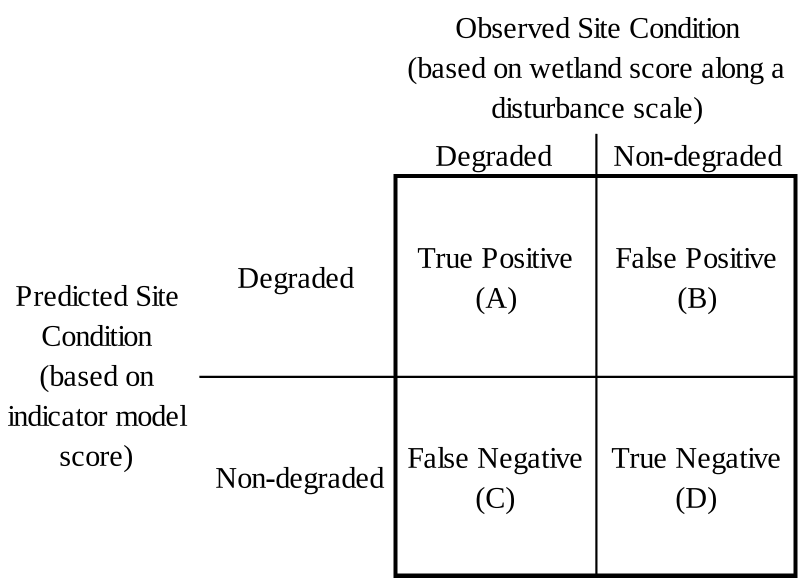

Fig. 2.1. Confusion matrix for assessing classification accuracy of biological indicators.

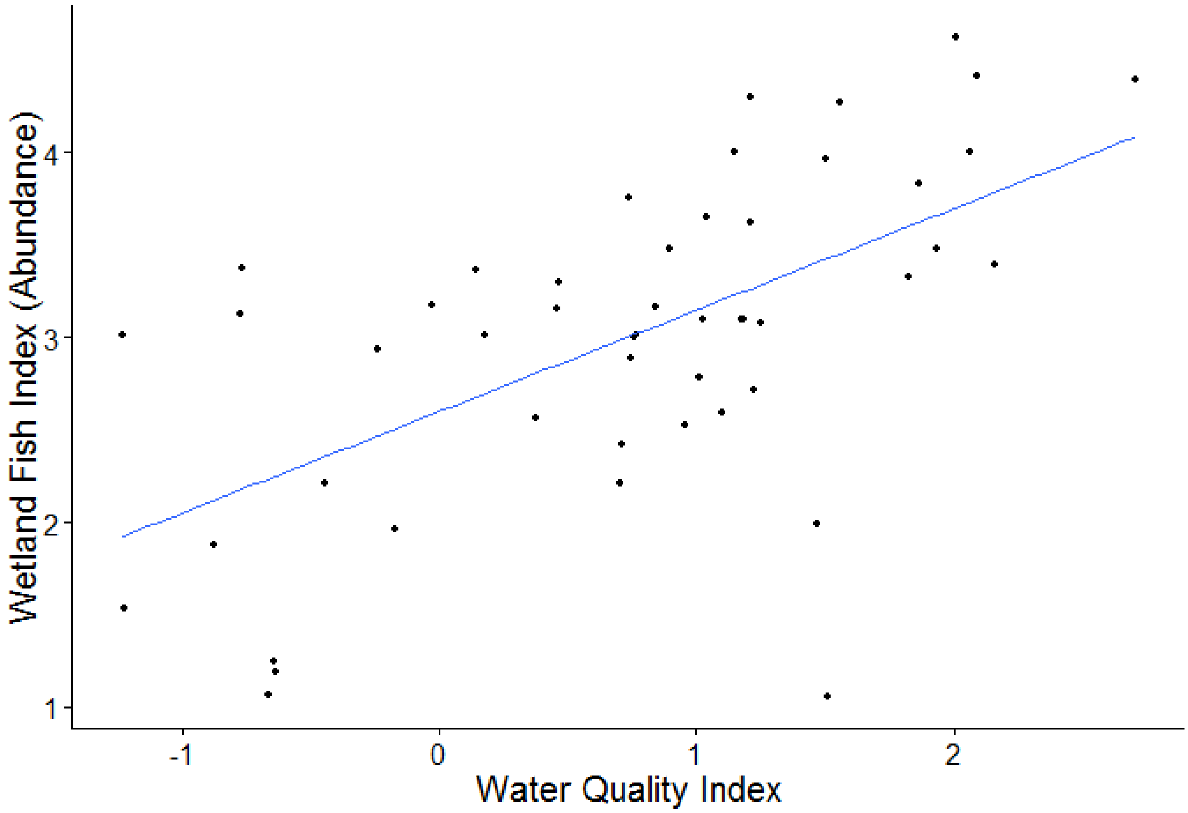

Fig. 2.2a. The relationship between the index score for WFI – AB and WQI score using the GLEI 2 test dataset. Increasing values for both the WFI and WQI indicate greater site degradation. The trendline is the least squares linear regression fitted to the data, R^2 = 0.343, p = <0.001, n = 49.

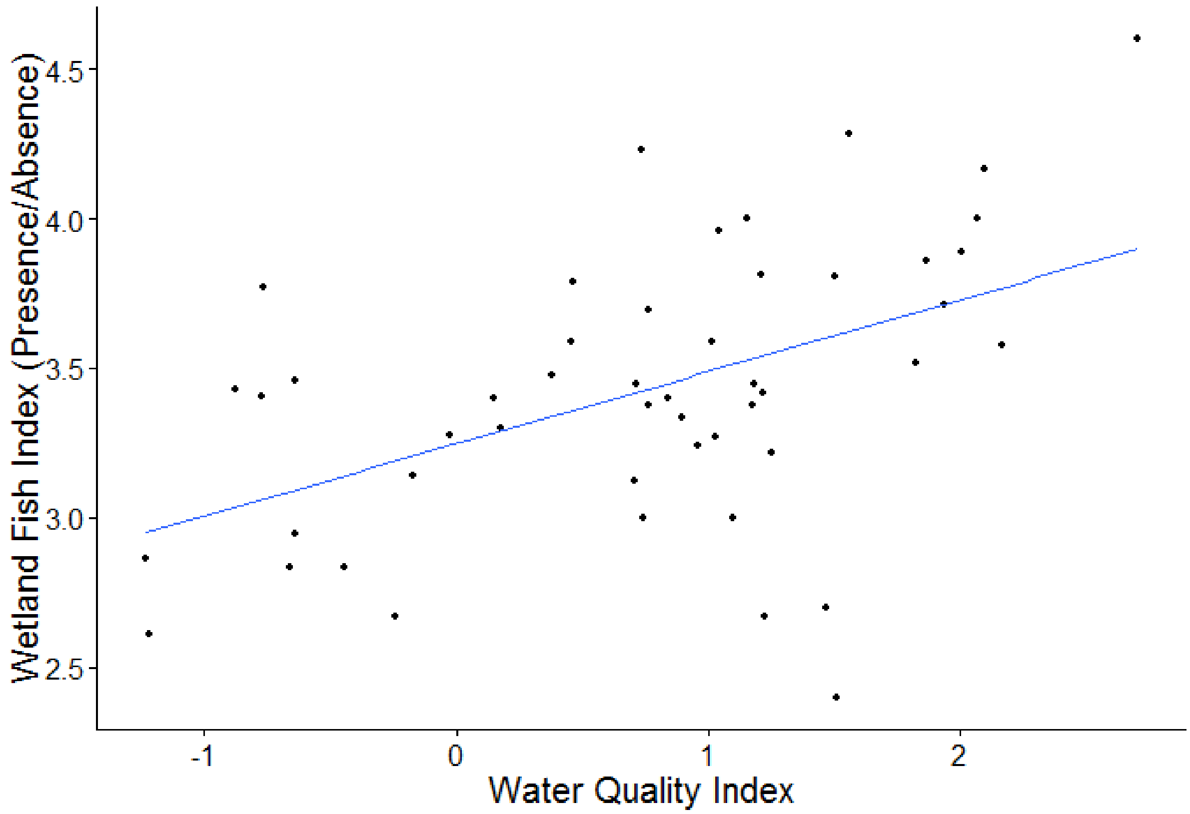

Fig. 2.2b. The relationship between the index score for WFI – PA and WQI score using the GLEI 2 test dataset. Increasing values for both the WFI and WQI indicate greater site degradation. The trendline is the least squares linear regression fitted to the data, R^2= 0.22, p = <0.001, n = 49.

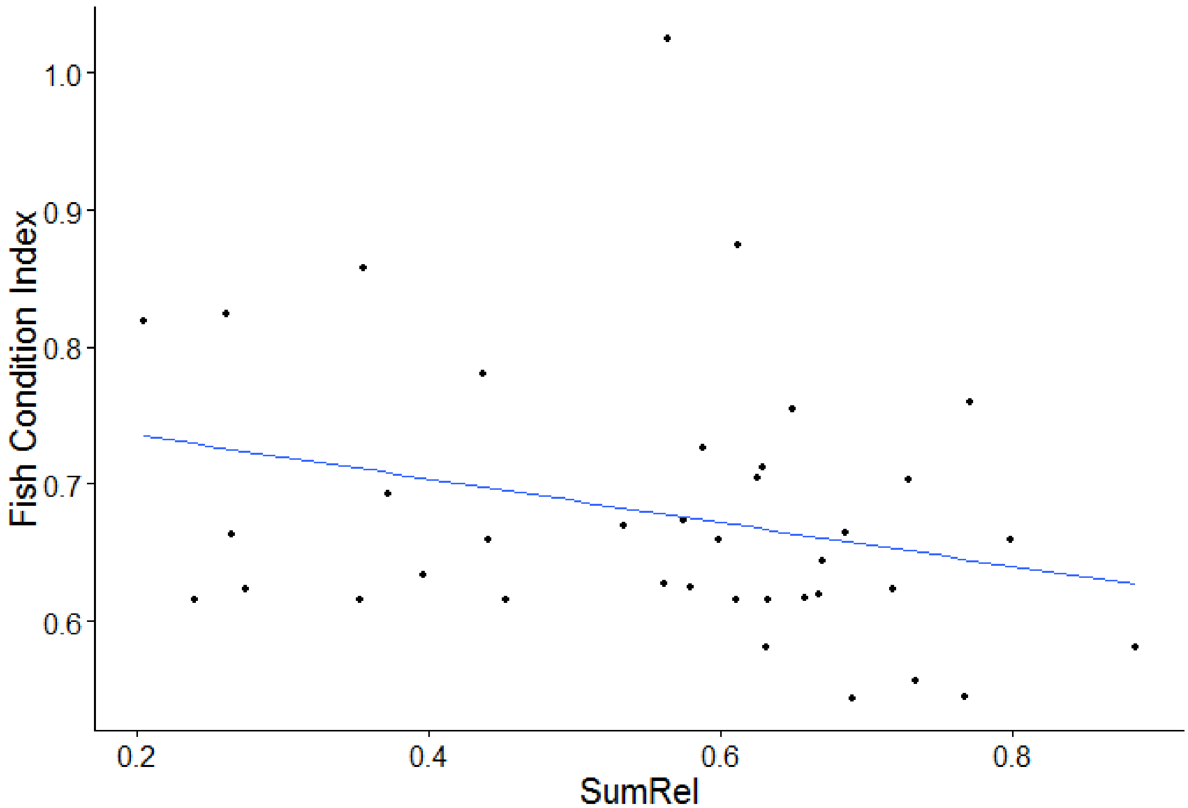

Fig. 2.2c. The relationship between the index score for FCI and SumRel score using the GLEI 2 test dataset, northern region. Increasing FCI score indicates less site degradation while increasing SumRel indicates greater degradation. The trendline is the least squares linear regression fitted to the data, $R^2 = 0.05$, $p = 0.0944$, $n = 38$.

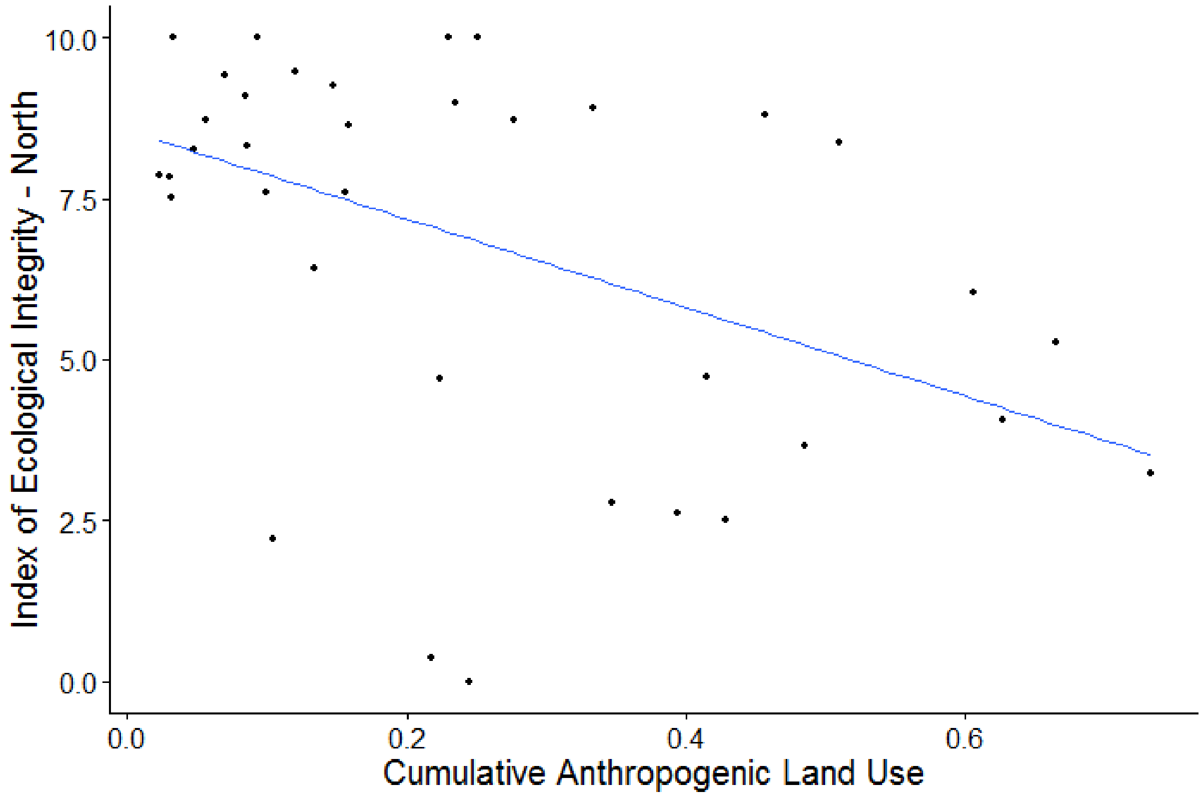

Fig. 2.2d. The relationship between the index score for IEC – North and cumulative land use using the GLEI 2, test dataset, northern region. Increasing IEC score indicates less site degradation while increasing cumulative land use indicates greater site degradation. The trendline is the least squares linear regression fitted to the data, $R^2 = 0.2$, p = <0.01, n = 38.

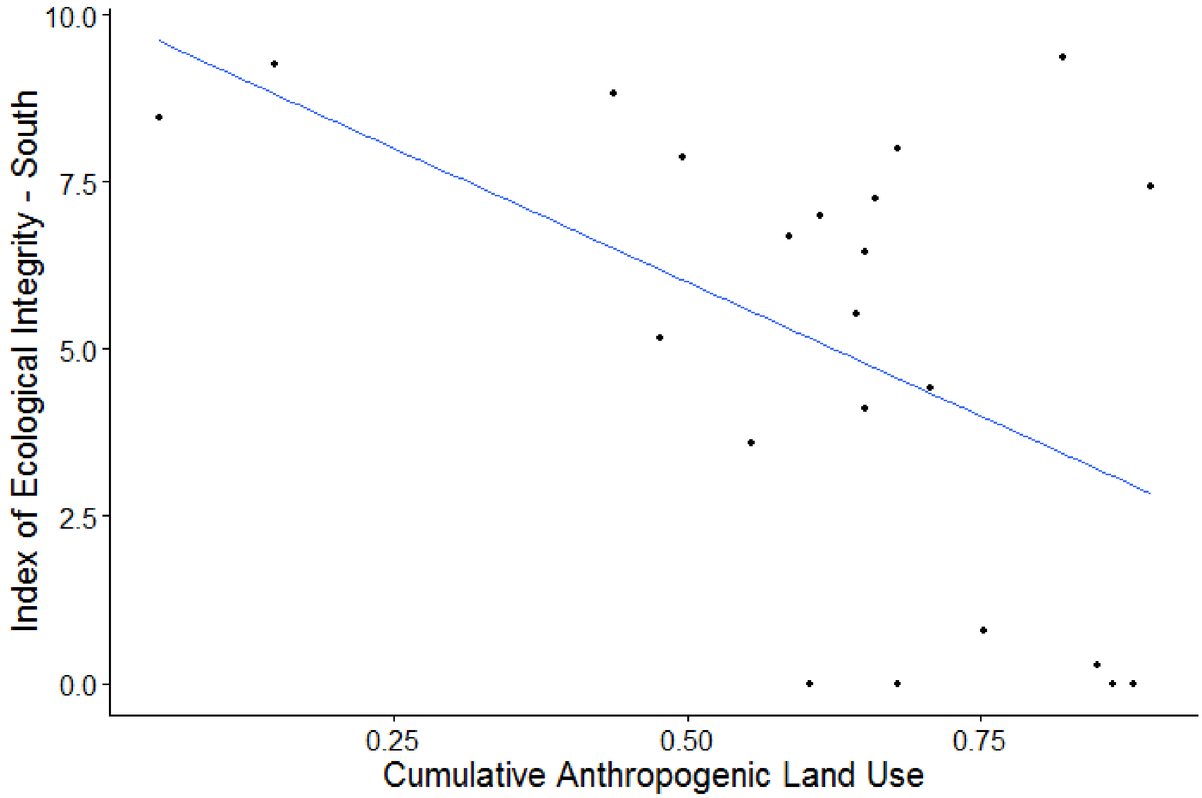

Fig. 2.2e. The relationship between the index score for IEC – South and cumulative land use using the GLEI 2 test dataset, south region. Increasing IEC score indicates less site degradation, while increasing cumulative land use indicates greater site degradation. The trendline is the least squares linear regression fitted to the data, R^2= 0.213, p = 0.018, n = 22.

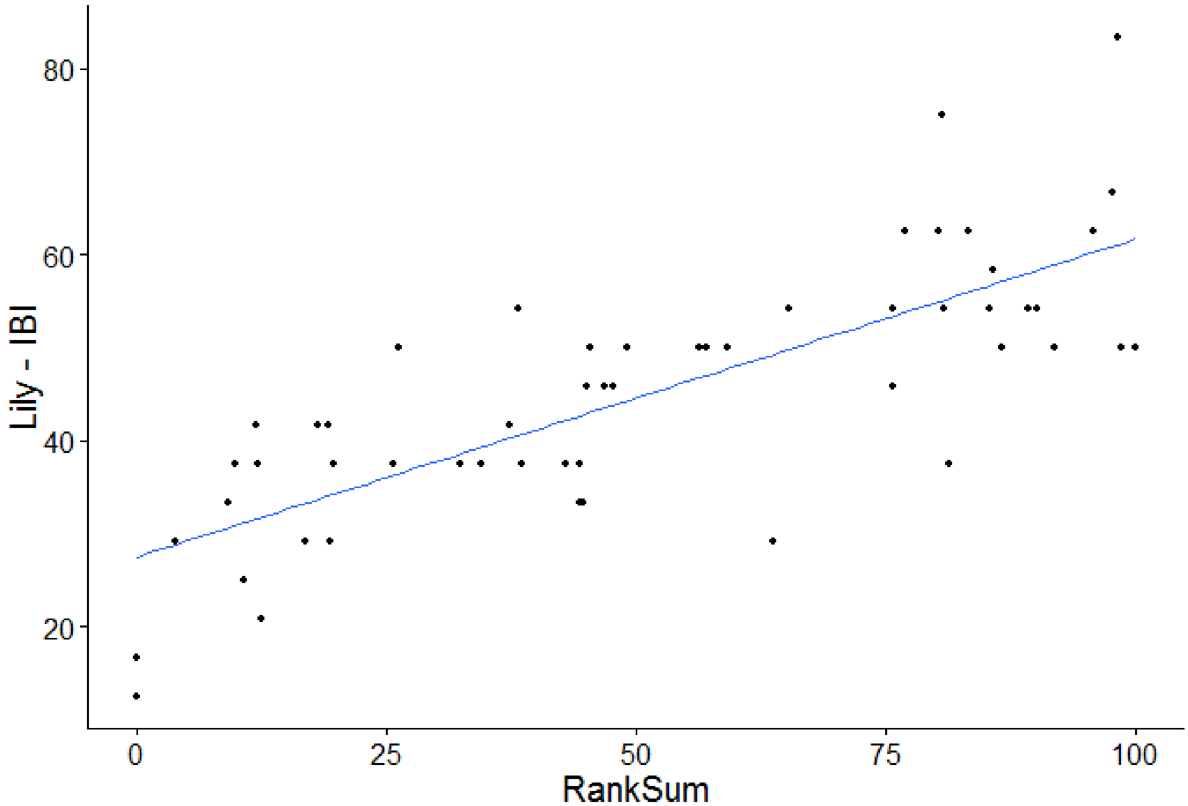

Fig. 2.2f. The relationship between the index score for Lily – IBI and the RankSum value using CWM test dataset, lily dominant zones. Increasing Lily-IBI scores and RankSum scores indicate less site degradation. The trendline is the least squares linear regression fitted to the data, R^2= 0.597, p = <0.001, n = 54.

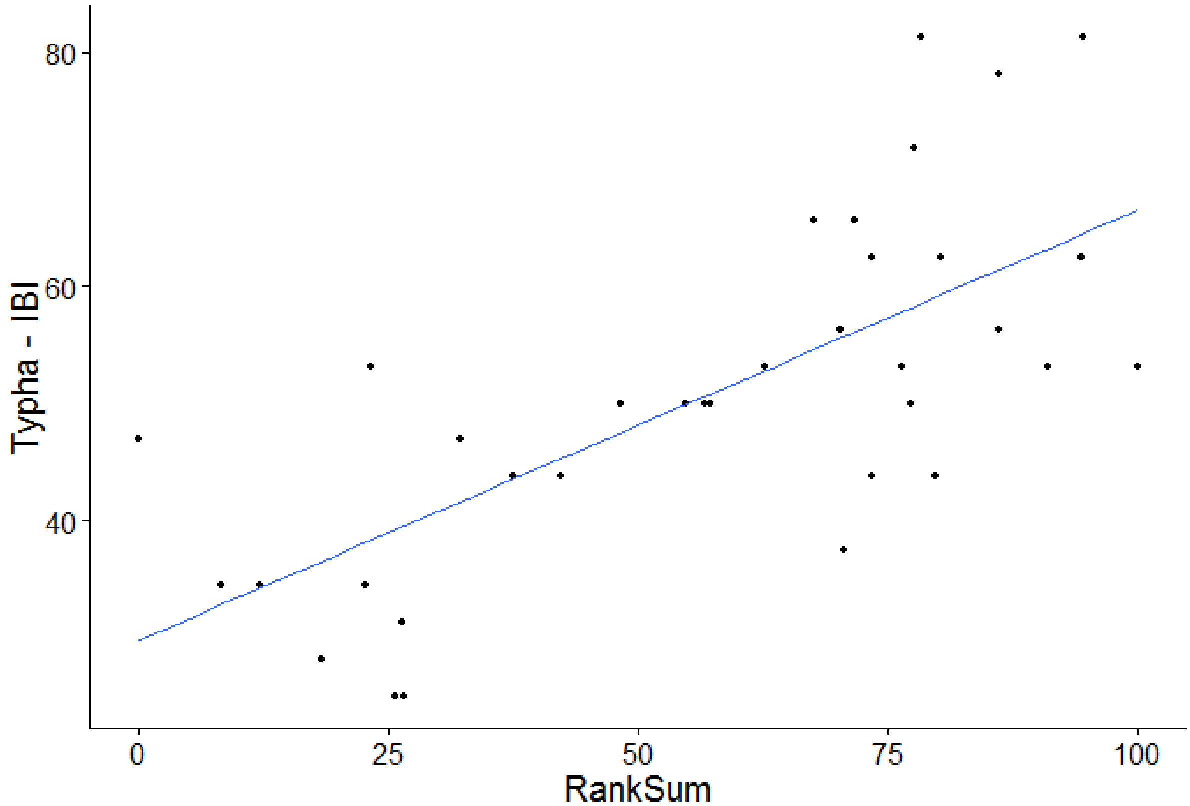

Fig. 2.2g. The relationship between the index score for Typha – IBI and the RankSum value using CWM test dataset, *Typha* dominant zones. Increasing Cattail-IBI scores and RankSum scores indicate less site degradation. The trendline is the least squares linear regression fitted to the data, R^2= 0.478, p = <0.001, n = 35.

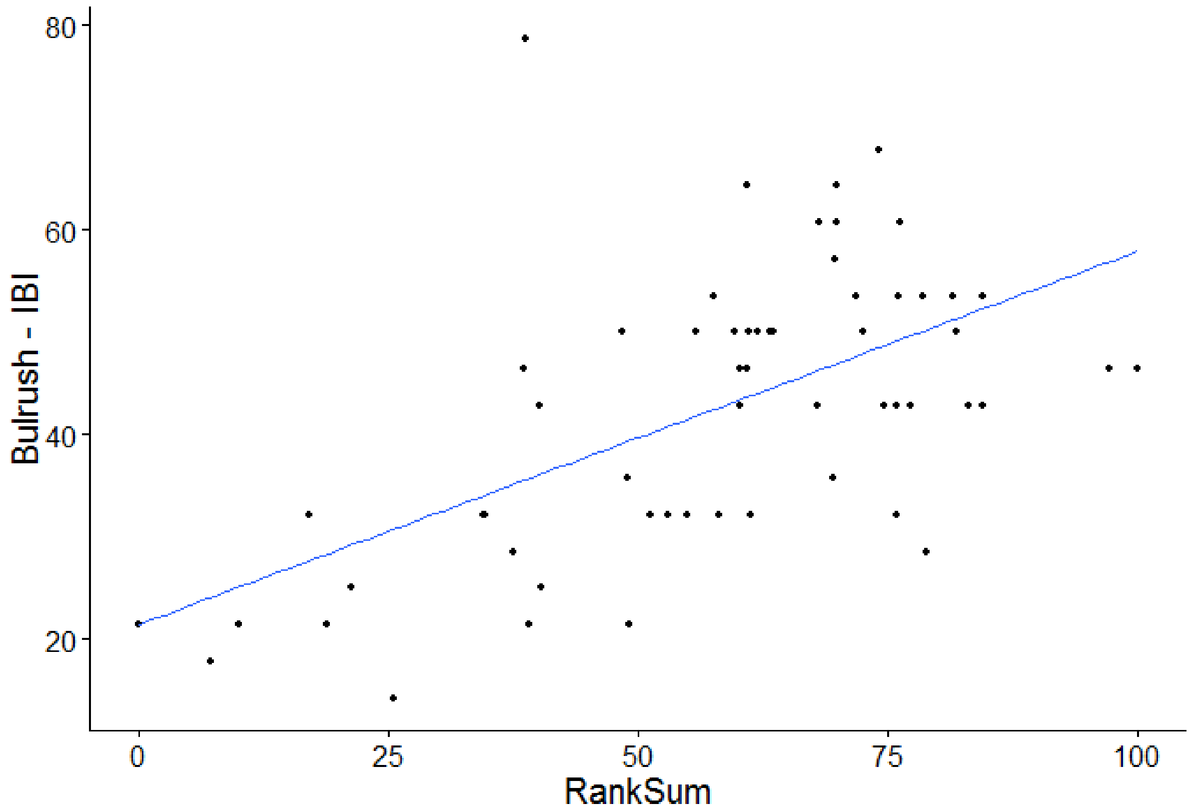

Fig. 2.2h. The relationship between the index score for Bulrush – IBI and the RankSum value using CWM test dataset, Bulrush dominant zones. Increasing Bulrush-IBI scores and RankSum scores indicate less site degradation. The trendline is the least squares linear regression fitted to the data, R^2= 0.327, p = <0.001, n = 58.

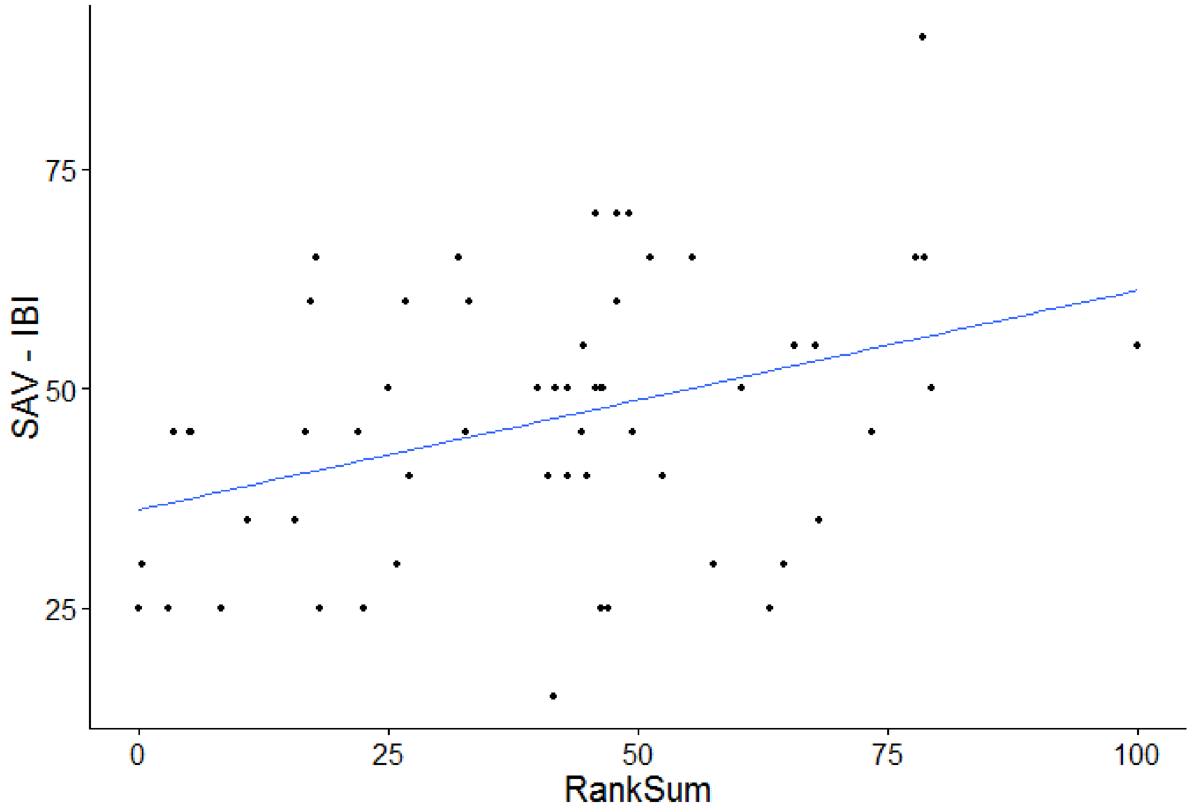

Fig. 2.2i. The relationship between the index score for the SAV – IBI and the RankSum value using CWM test dataset, SAV dominant zones. Increasing SAV-IBI scores and RankSum scores indicate less site degradation. The trendline is the least squares linear regression fitted to the data, R^2= 0.133, p = <0.01, n = 59.

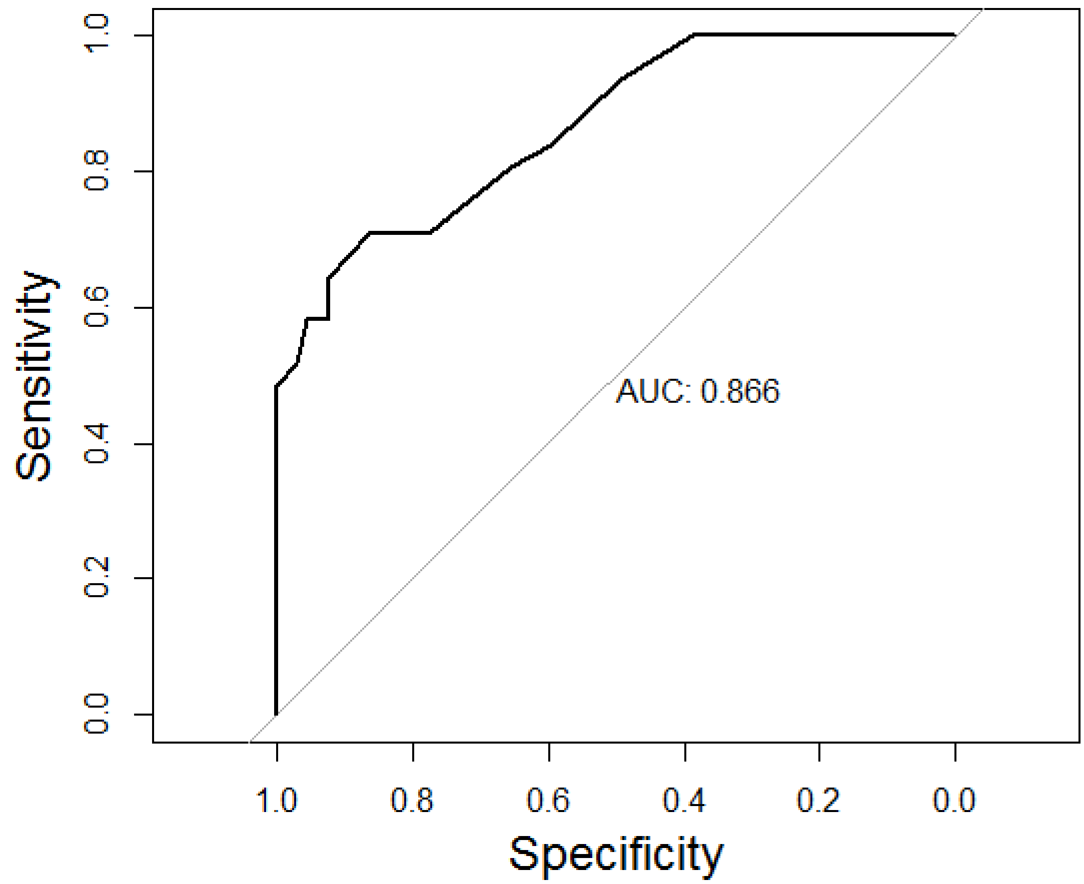

Fig. 2.3a. Overall classification accuracy of the WFI – AB using the GLEI 1 training dataset. WQI based disturbance threshold: WQI < 0. AUC = 0.866, AUC 95% CI = 0.7895 - 0.9424.

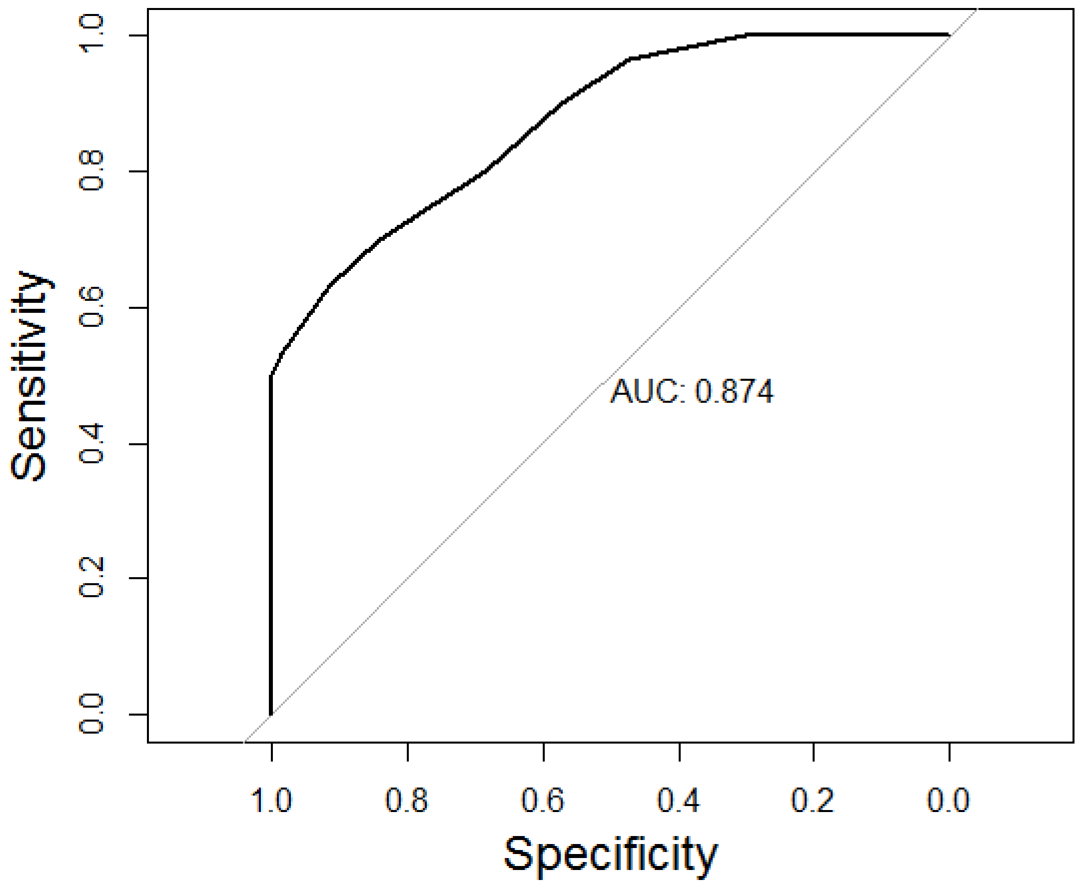

Fig. 2.3b. Overall classification accuracy of the WFI – PA using the GLEI 1 training dataset. WQI based disturbance threshold: WQI = 0. AUC = 0.8736, AUC 95% CI = 0.7998 - 0.9474

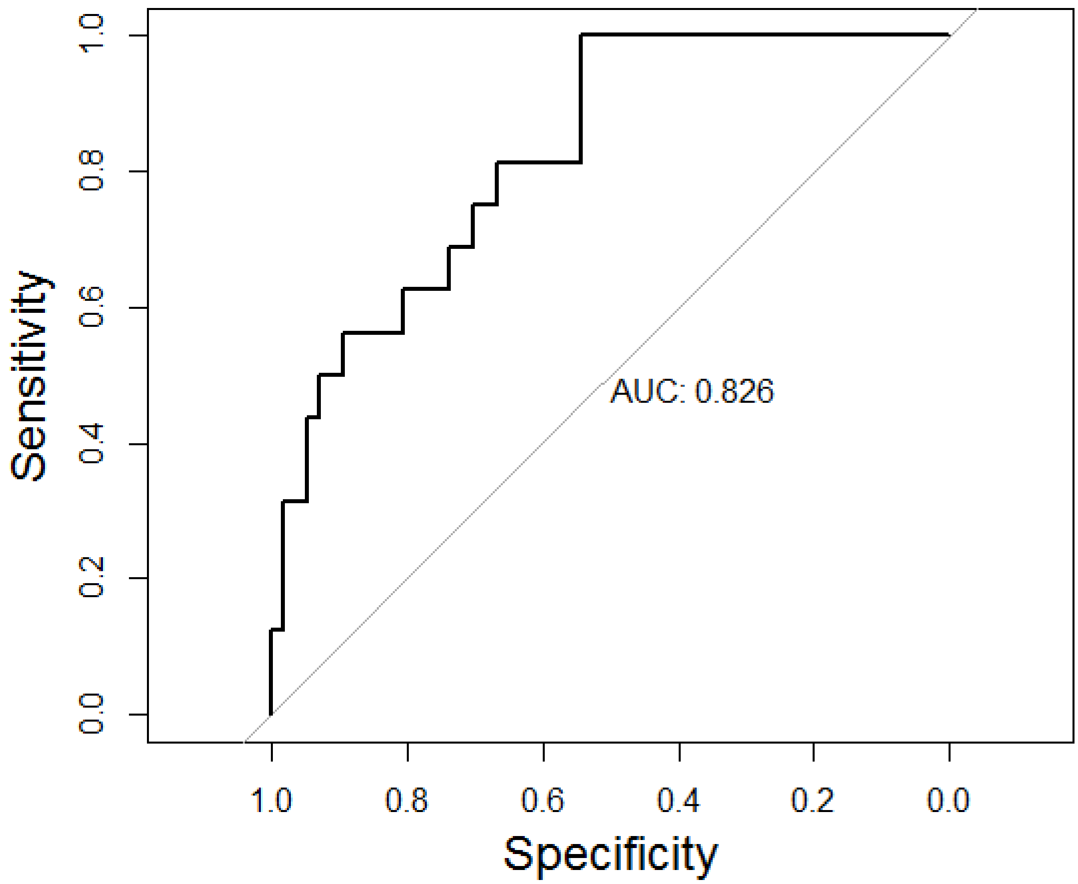

Fig. 2.3c. Overall classification accuracy of the FCI using the GLEI 1 training dataset. SumRel based disturbance threshold: SumRel = 0.745. AUC = 0.8257, AUC 95% CI = 0.7211 - 0.9303.

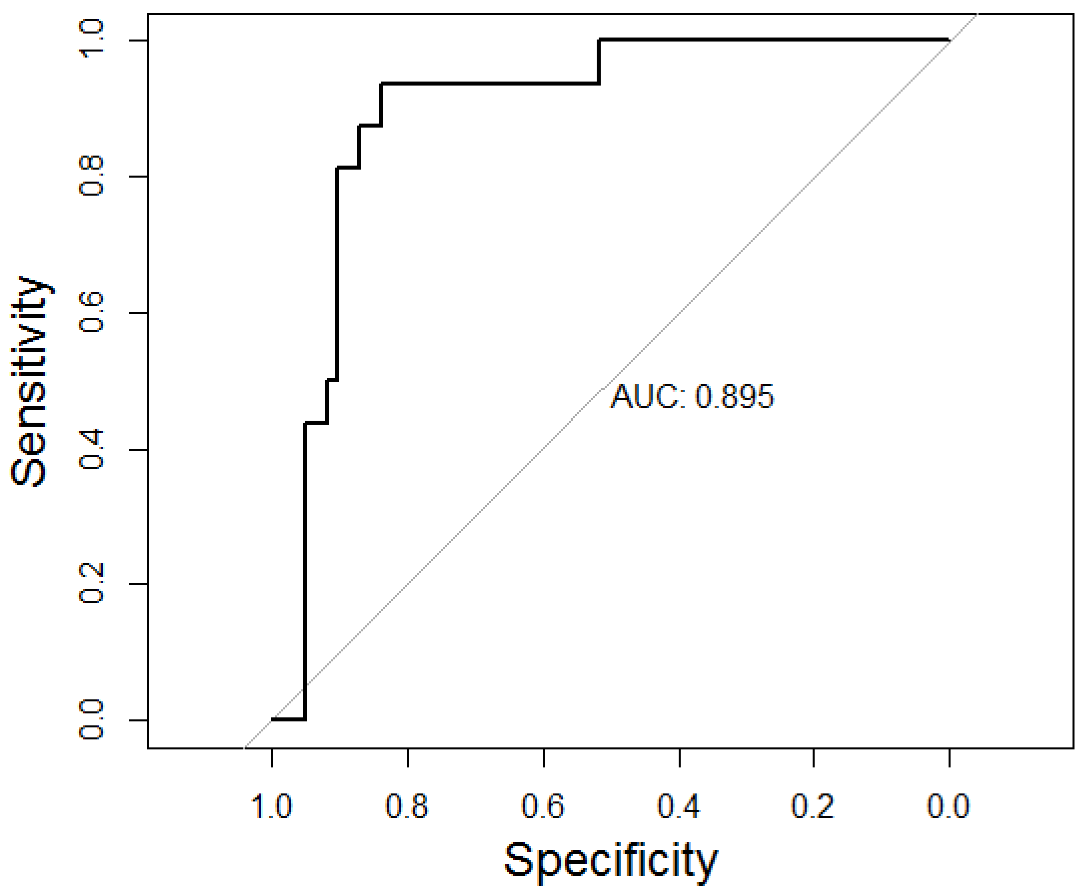

Fig. 2.3d Overall classification accuracy of the IEC-North using the GLEI 1 training dataset. Cumulative land use based disturbance threshold = 0.358. AUC = 0.8952, AUC 95% CI = 0.8166 - 0.9737.

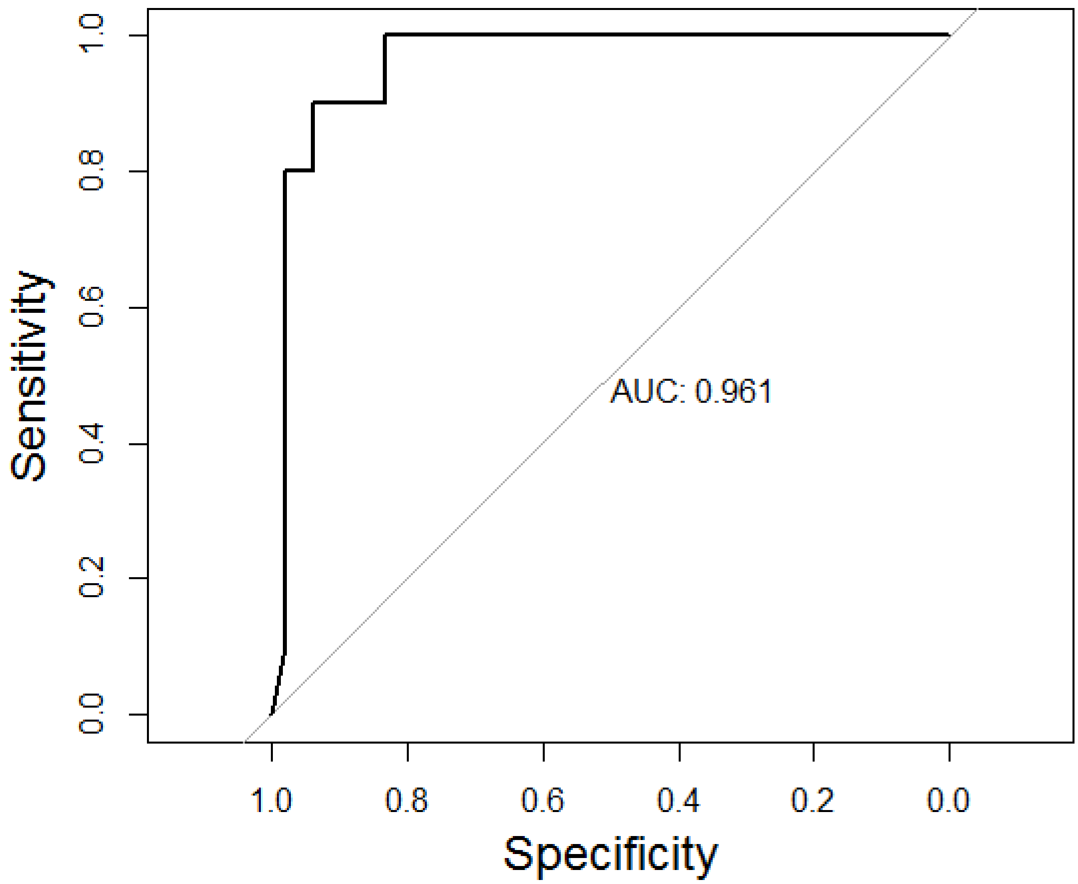

Fig. 2.3e. Overall classification accuracy of the IEC-South using the GLEI 1 training dataset. Cumulative land use based disturbance threshold = 0.755. AUC = 0.9615, AUC 95% CI = 0.9166 - 1.0000.

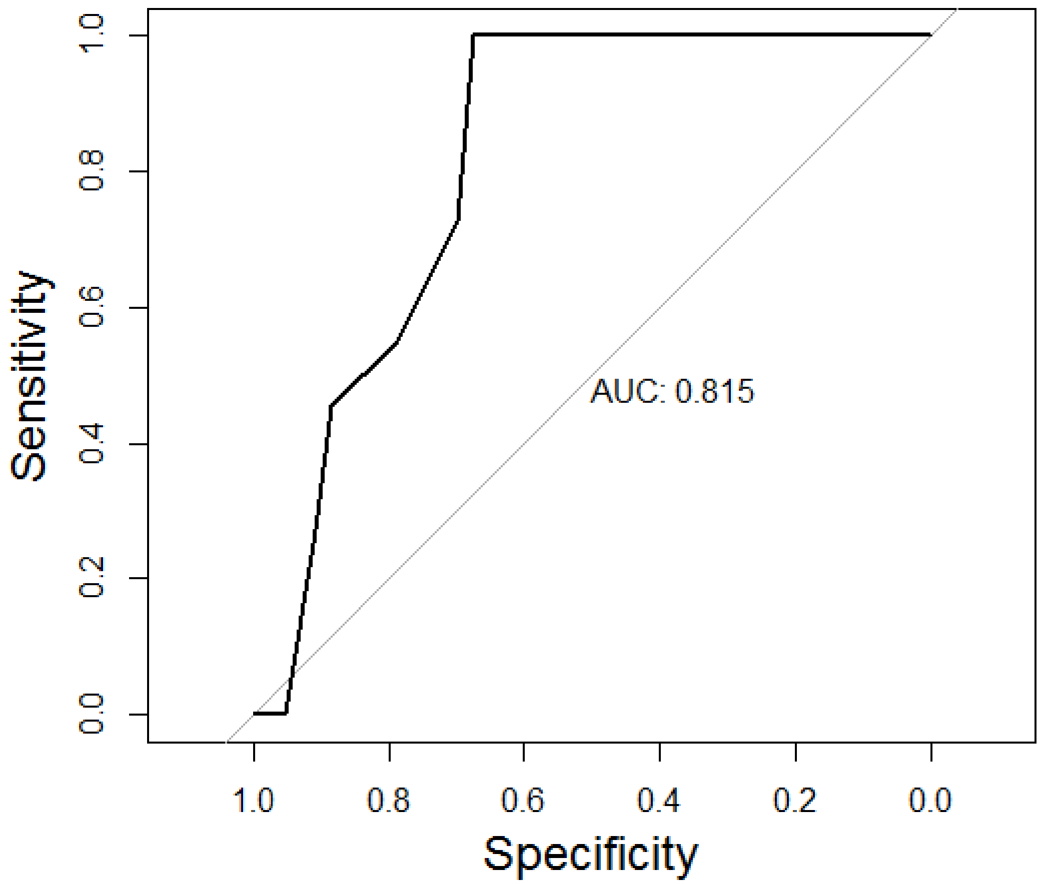

Fig. 2.3f. Overall classification accuracy of the Lily-IBI using the CWM training dataset. RankSum based disturbance threshold: RankSum = 16, AUC = 0.815, AUC 95% CI = 0.704 - 0.926.

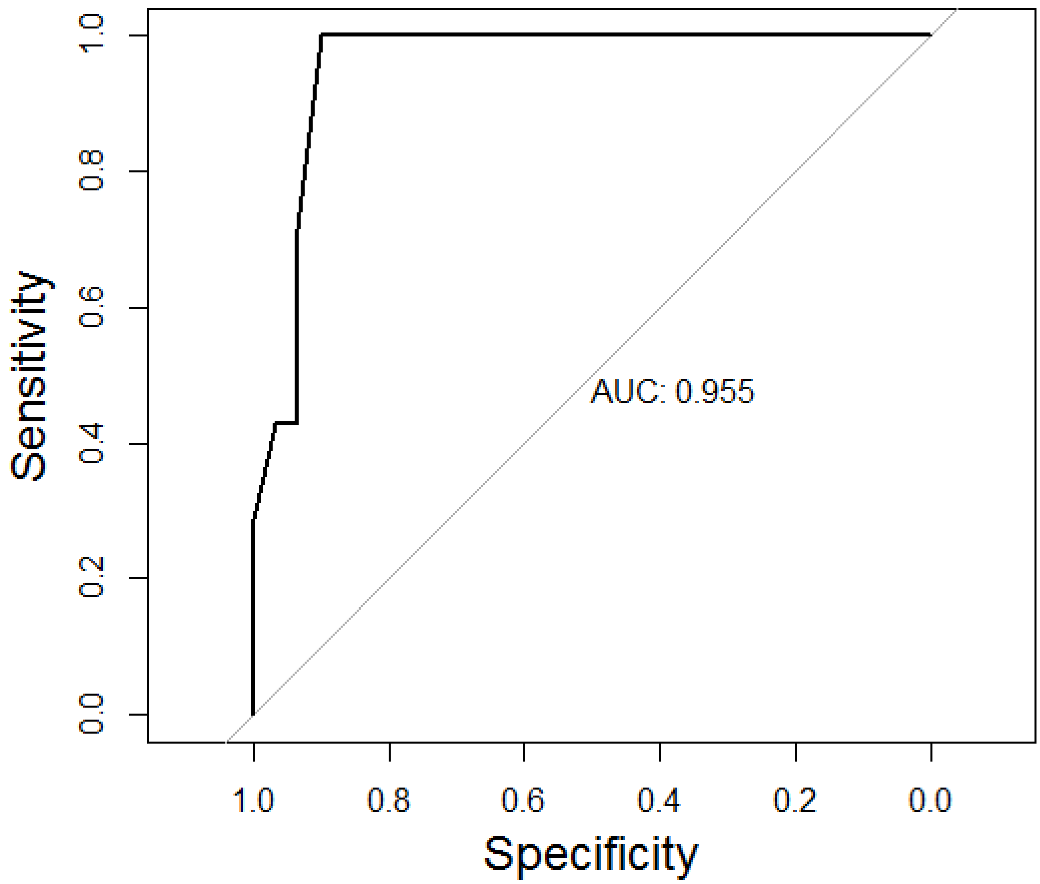

Fig. 2.3g. Overall classification accuracy of the Cattail-IBI using the CWM training dataset. RankSum based disturbance threshold: RankSum = 24. AUC = 0.955, AUC 95% CI = 0.892 – 1.

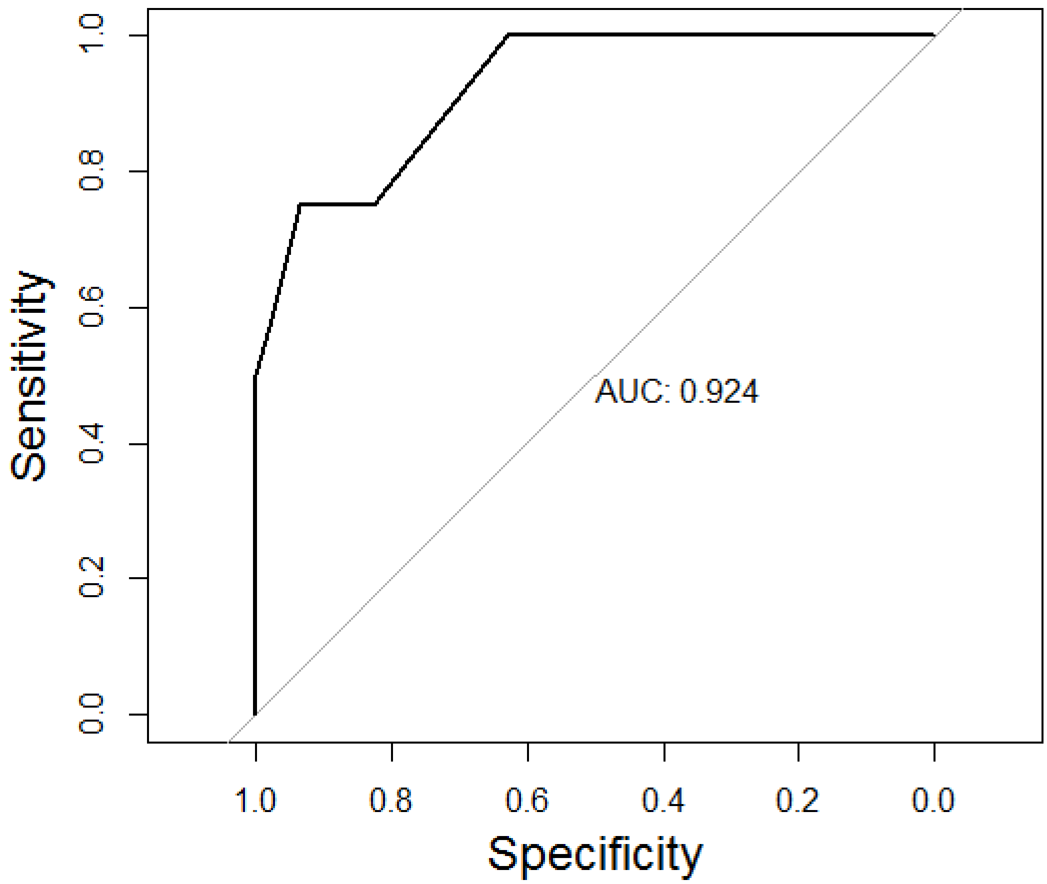

Fig. 2.3h. Overall classification accuracy of the Bulrush-IBI using the CWM training dataset. RankSum based disturbance threshold: RankSum = 41. AUC = 0.924, AUC 95% CI = 0.847 - 1.

.

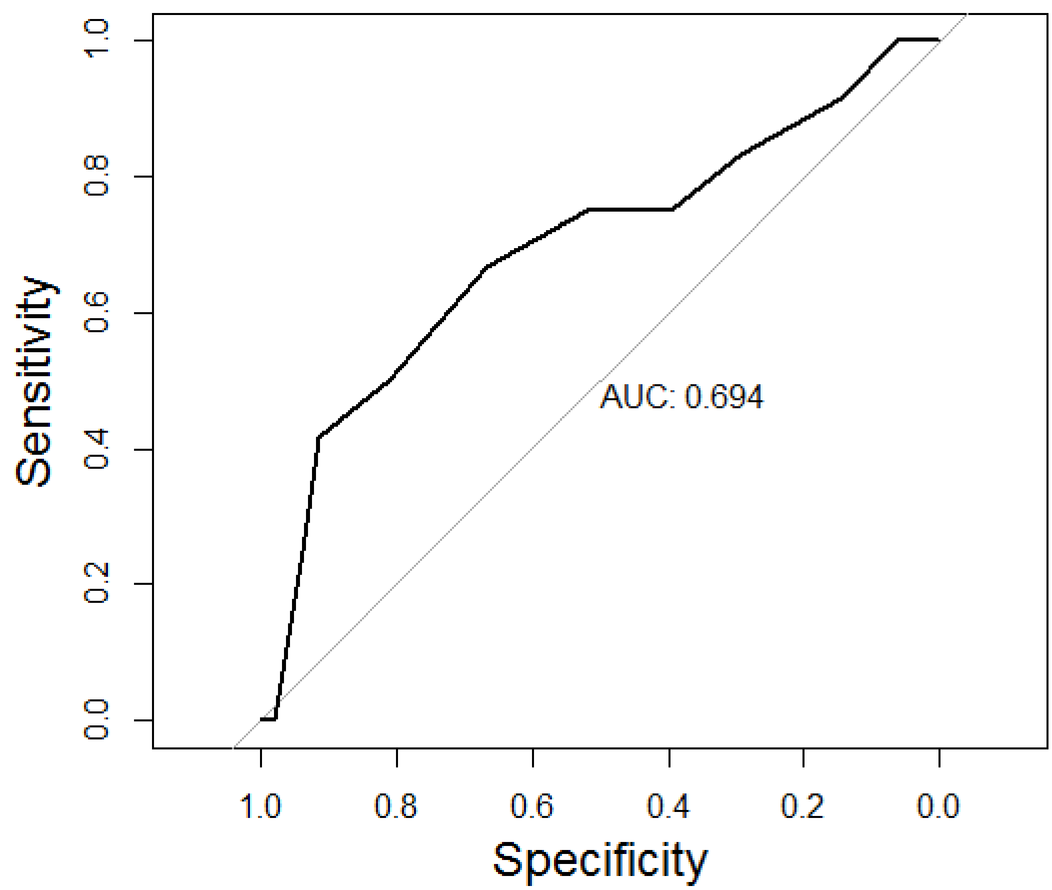

Fig. 2.3h. Overall classification accuracy of the SAV-IBI using the CWM training dataset. RankSum based disturbance threshold: RankSum = 42. AUC = 0.694, AUC 95% CI = 0.507 - 0.882

.

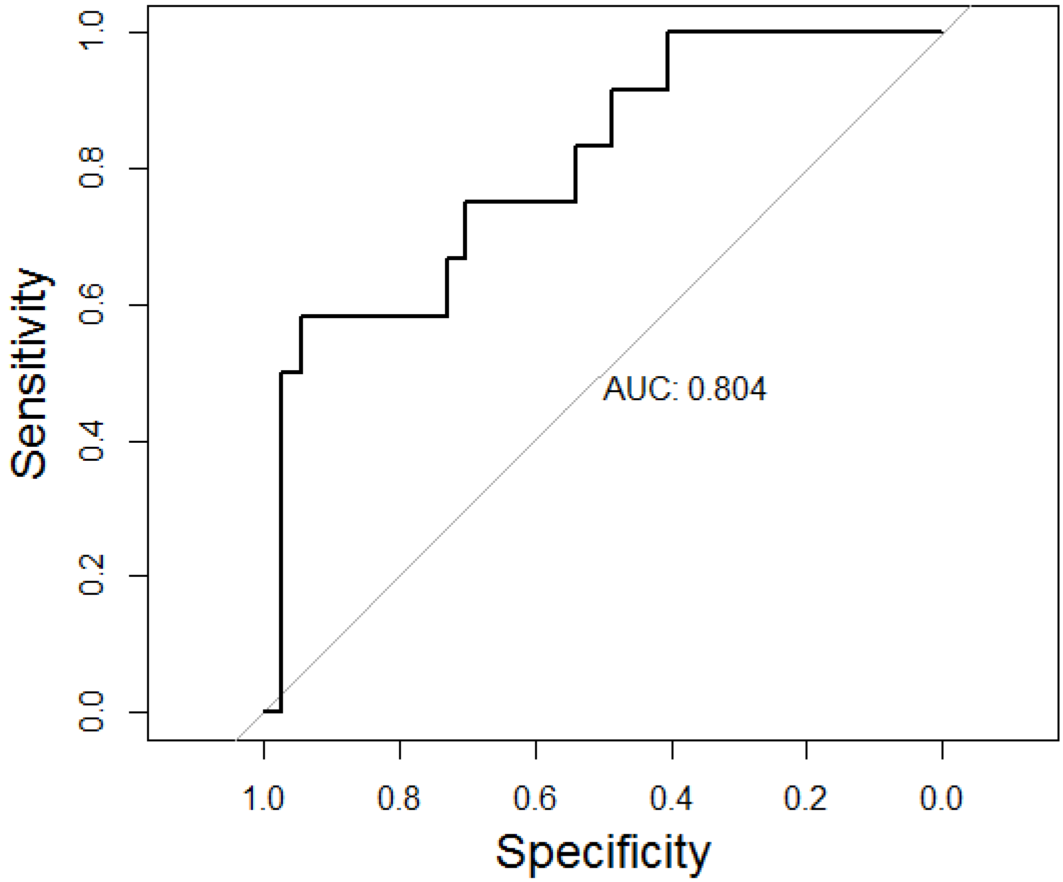

Fig. 2.4a. Overall classification accuracy of the WFI – AB using the GLEI 2 test dataset. WQI based disturbance threshold: WQI < 0. AUC = 0.8041, AUC 95% CI = 0.66 - 0.9481.

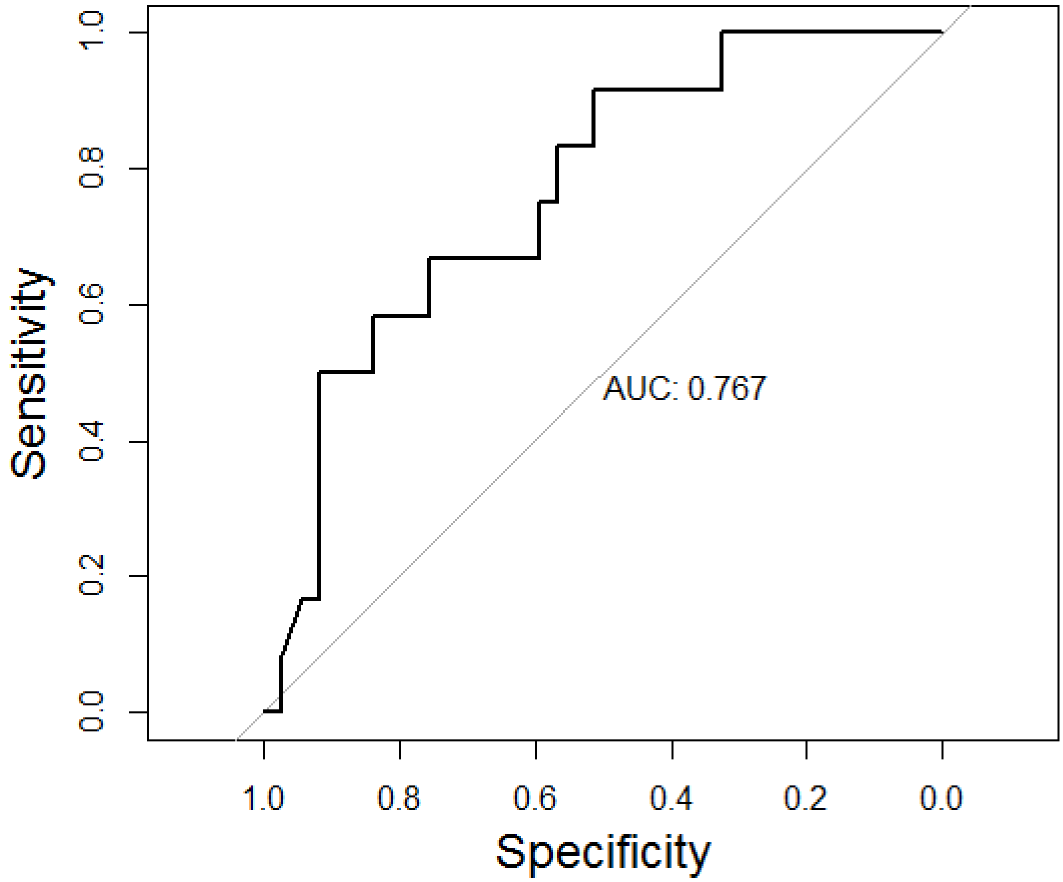

Fig. 2.4b. Overall classification accuracy of the WFI – PA using the GLEI 2 test dataset. WQI based disturbance threshold: WQI < 0. AUC = 0.7669, AUC 95% CI = 0.6182 - 0.9156.

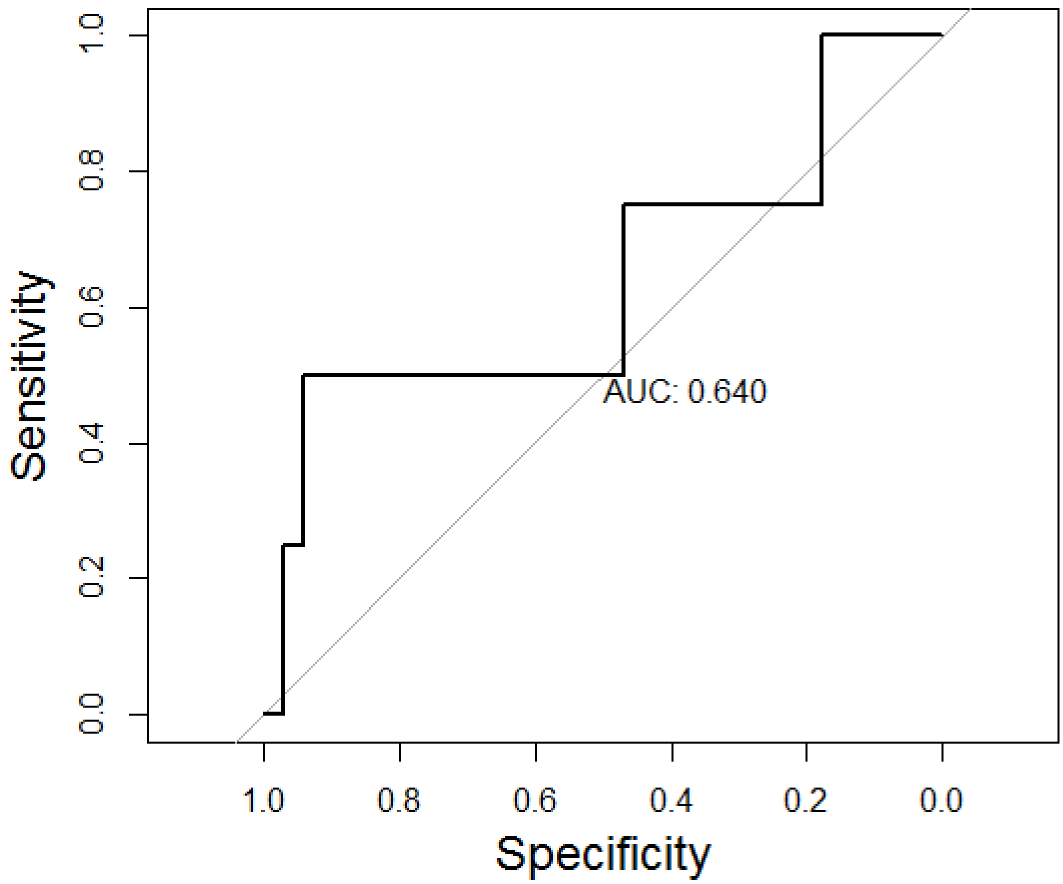

Fig. 2.4c. . Overall classification accuracy of the FCI using the GLEI 2 test dataset. SumRel based disturbance threshold: SumRel = 0.745. AUC = 0.6397, AUC 95% CI = 0.2549 - 1.0.

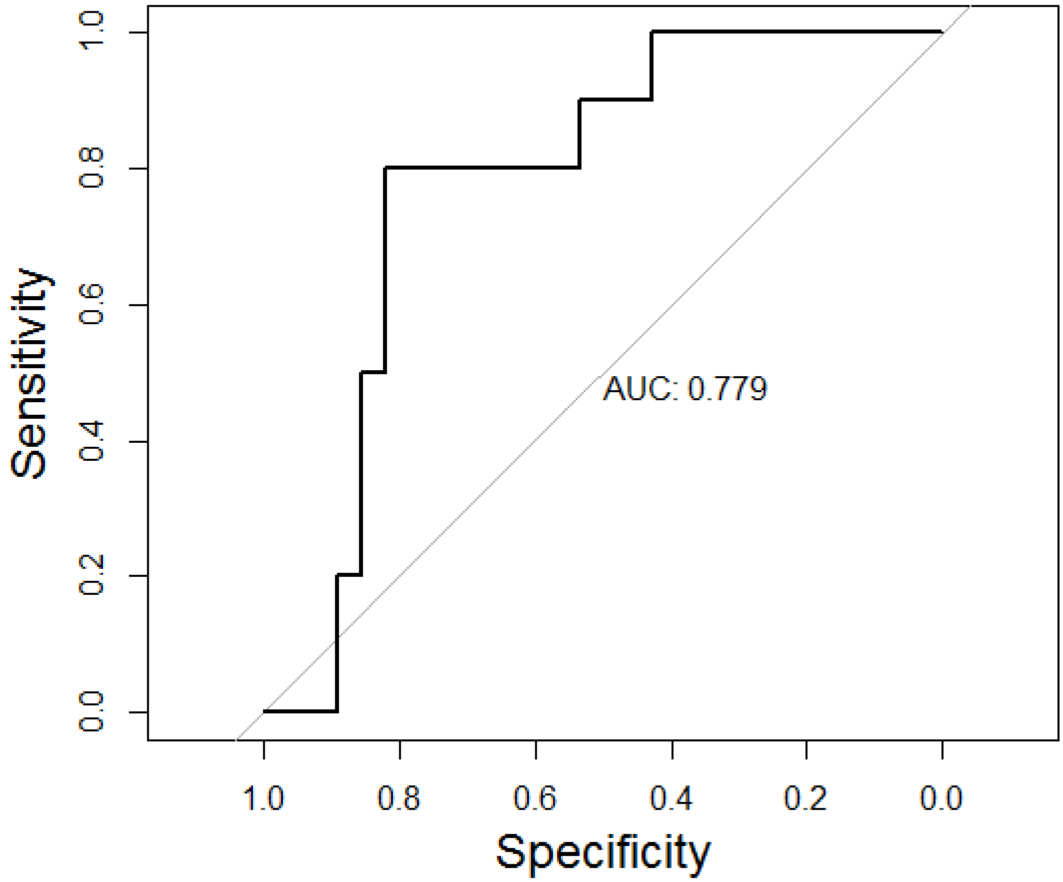

Fig. 2.4d. Overall classification accuracy of the IEC-North using the GLEI 2 test dataset. Cumulative land use disturbance threshold = 0.358. AUC = 0.7786, AUC 95% CI = 0.6222 - 0.935.

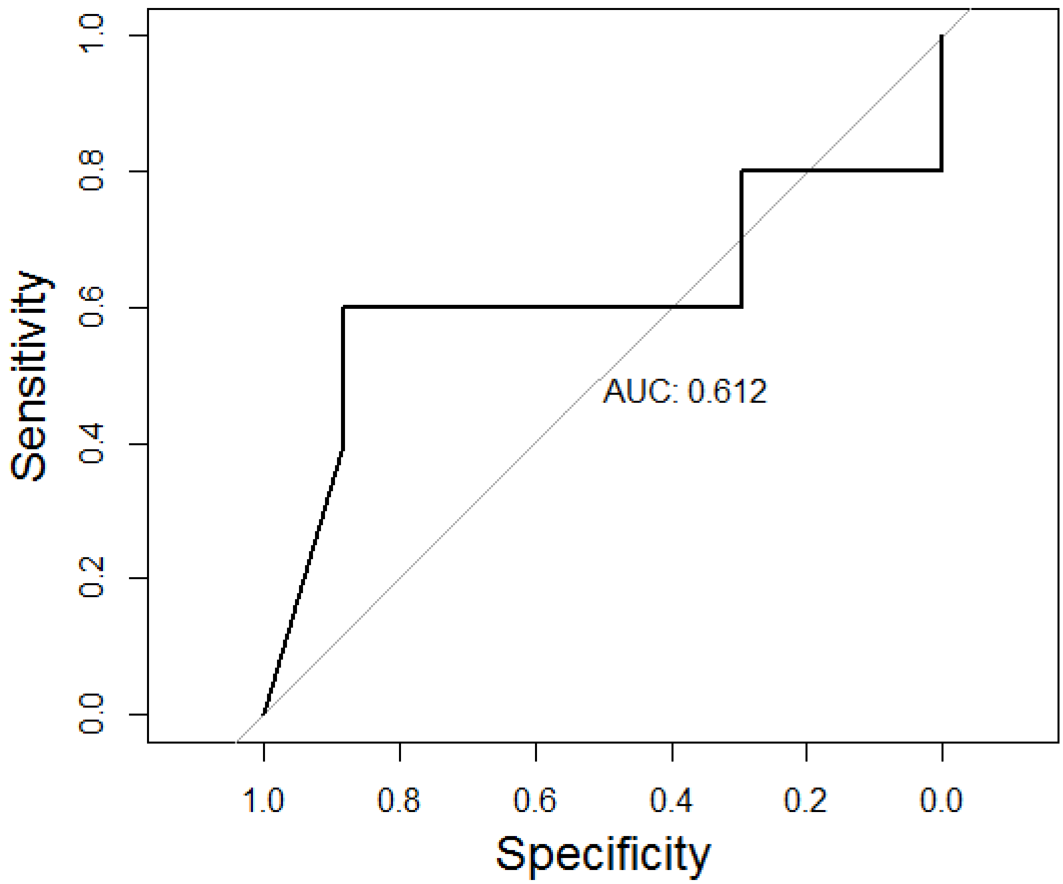

Fig. 2.4e Overall classification accuracy of the IEC-South using the GLEI 2 test dataset. Cumulative land use disturbance threshold = 0.6118. AUC = 0.6118, AUC 95% CI = 0.2189 - 1.

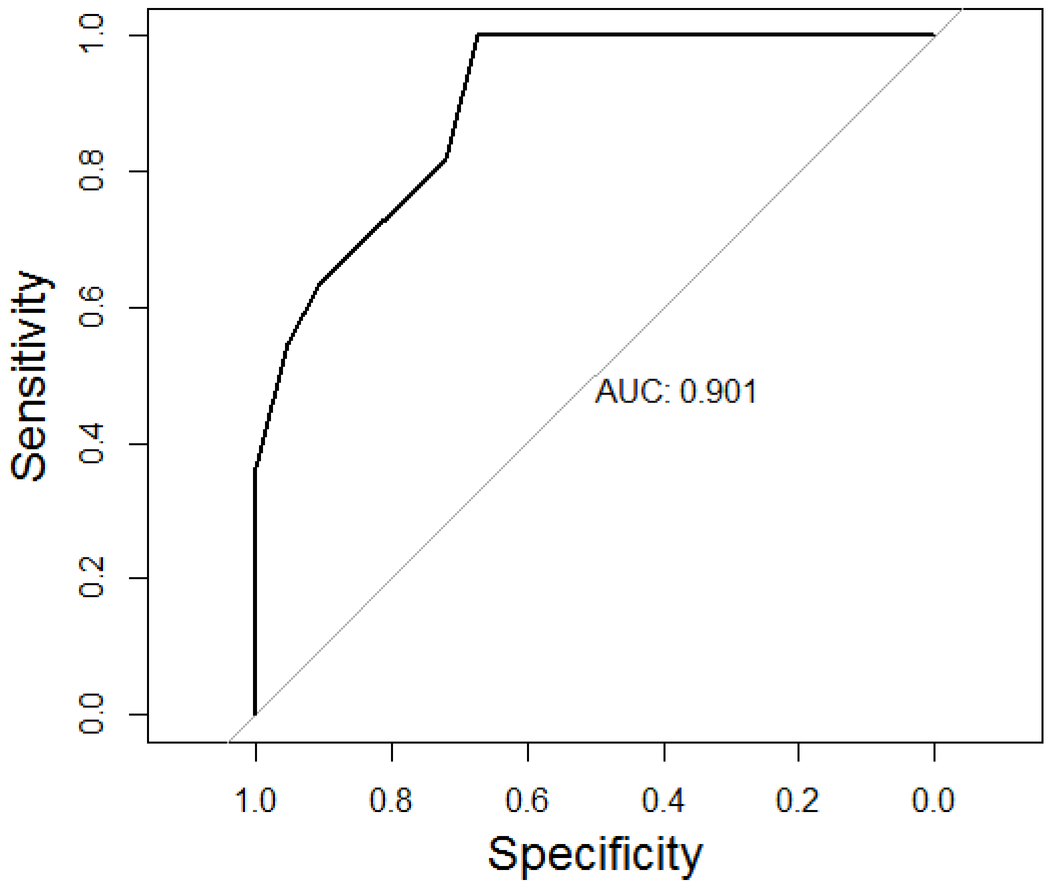

Fig. 2.4f. Overall classification accuracy of the Lily-IBI using the CWM test dataset. RankSum based disturbance threshold: RankSum = 19. AUC = 0.901, AUC 95% CI = 0.814 - 0.988.

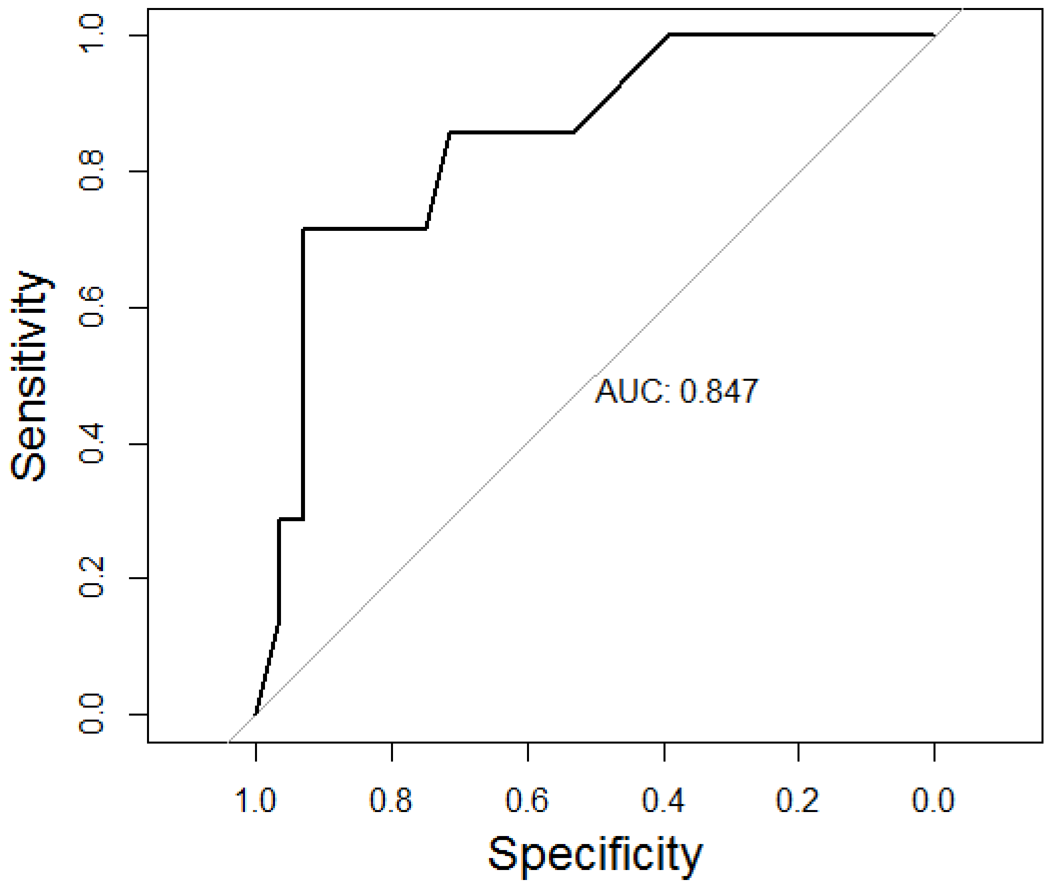

Fig. 2.4g. Overall classification accuracy of the Cattail-IBI using the CWM test dataset. RankSum based disturbance threshold: RankSum = 26. AUC = 0.847, AUC 95% CI = 0.686 – 1.

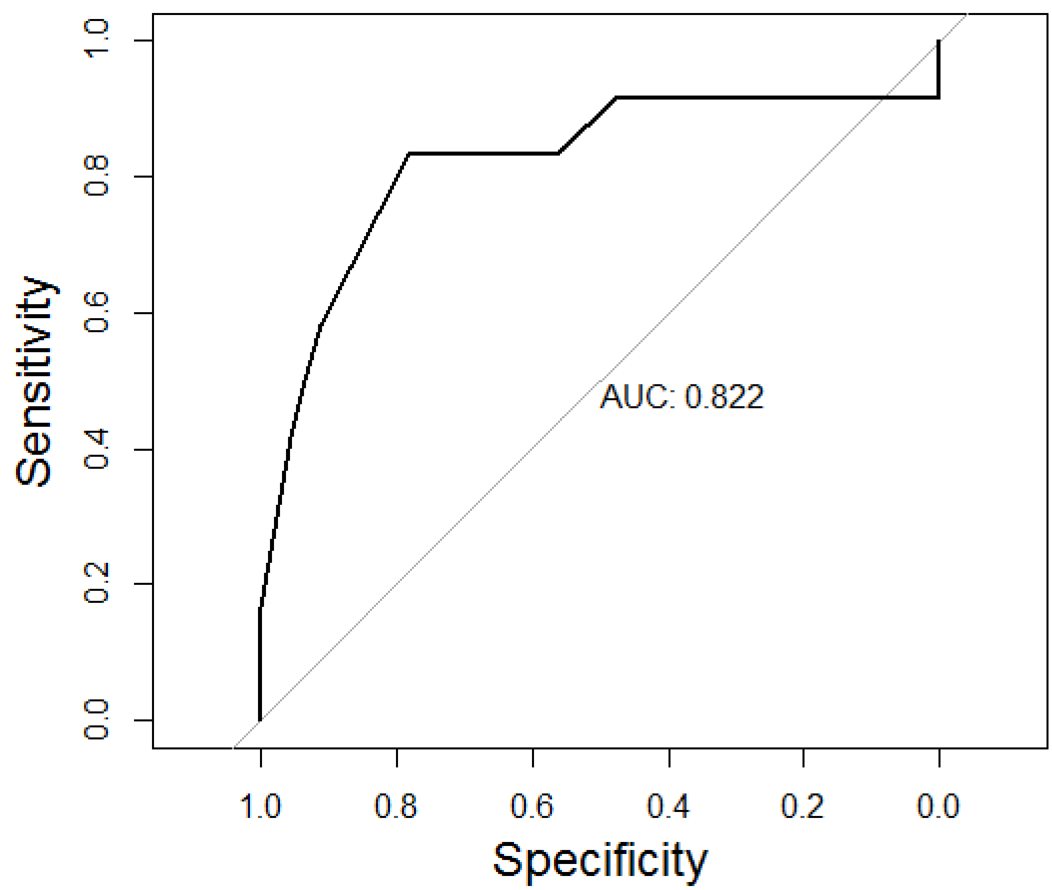

Fig. 2.4h. Overall classification accuracy of the Bulrush-IBI using the CWM test dataset. RankSum based disturbance threshold: RankSum = 39. AUC = 0.822, AUC 95% CI = 0.652 – 0.993.

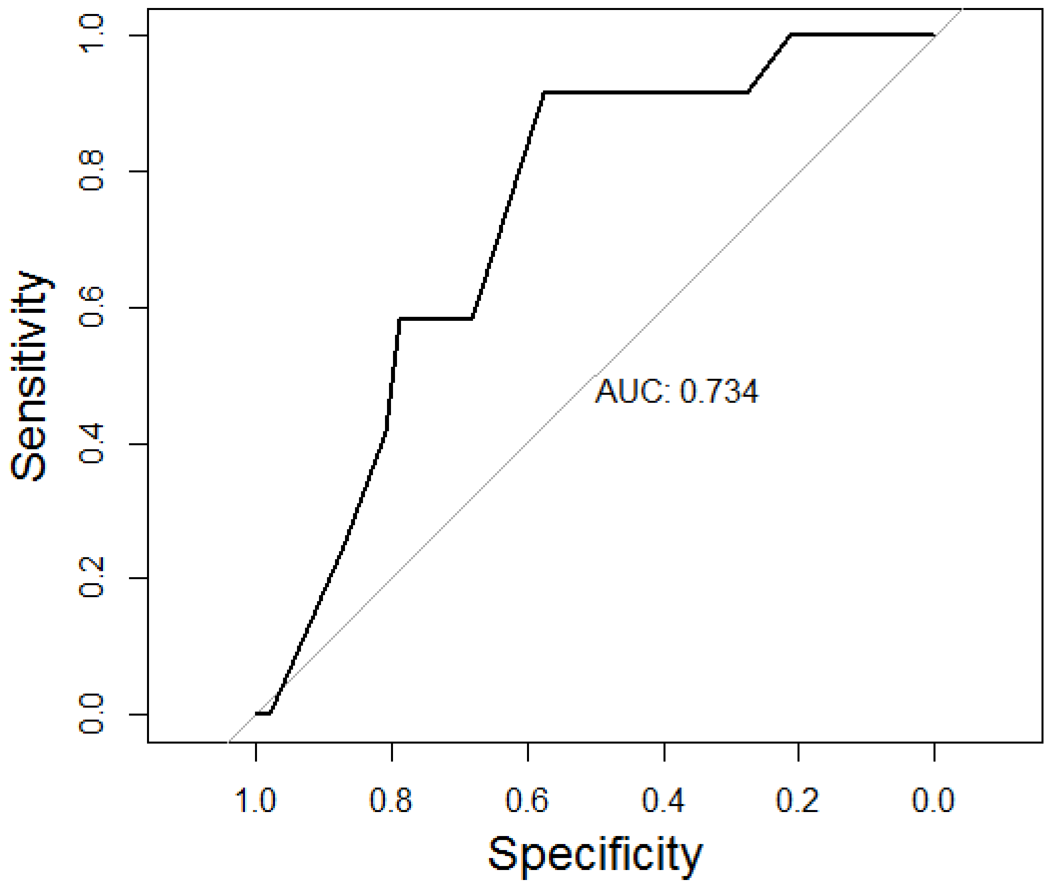

Fig. 2.4i. Overall classification accuracy of the SAV-IBI using the CWM test dataset. RankSum based disturbance threshold: RankSum = 17.5. AUC = 0.734, AUC 95% CI = 0.592 – 0.876.

References

Alberg, A. J., J. W. Park, B. W. Hager, M. V Brock, and M. Diener-West. 2004. The use of “overall accuracy” to evaluate the validity of screening or diagnostic tests. Journal of general internal medicine 19:460–5.

Aparicio, E., G. Carmona-Catot, P. B. Moyle, and E. García-Berthou. 2011. Development and evaluation of a fish-based index to assess biological integrity of Mediterranean streams. Aquatic Conservation: Marine and Freshwater Ecosystems 21:324–337.

Bailey, R. C., R. H. Norris, and T. B. Reynoldson. 2004. Bioassessment of freshwater ecosystems: using the reference condition approach. Pages 1–170 Freshwater Biology. Kluwer Acedemic Publishers.

Borja, A., S. B. Bricker, D. M. Dauer, N. T. Demetriades, J. G. Ferreira, A. T. Forbes, P. Hutchings, X. Jia, R. Kenchington, J. Carlos Marques, and C. Zhu. 2008. Overview of integrative tools and methods in assessing ecological integrity in estuarine and coastal systems worldwide. Marine pollution bulletin 56:1519–37.

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.

Burton, T. M., D. G. Uzarski, J. P. Gathman, J. A. Genet, B. E. Keas, and C. A. Stricker. 1999. Development of a preliminary invertebrate index of biotic integrity for Lake Huron coastal wetlands. Wetlands 19:869–882.

Cairns, J., P. V. McCormick, and B. R. Niederlehner. 1993. A proposed framework for developing indicators of ecosystem health. Hydrobiologia 263:1–44.

Calabro, E. J., B. A. Murry, D. A. Woolnough, G. Uzarski, and D. G. Uzarski. 2013. Application and transferability of Great Lakes coastal wetland indices of biotic integrity to high quality inland lakes of Beaver Island in northern Lake Michigan. Aquatic Ecosystem Health & Management 16:338–346.

Chow-Fraser, P. 2006. Development of the Water Quality Index (WQI) to assess effects of basin-wide land-use alteration on coastal marshes of the Laurentian Great Lakes. Pages 137–166 *in* T.P. Simon and P.M. Stewart. Indiana Biological Survey, editor. Coastal wetlands of the Laurentian Great Lakes: health, habitat and indicators.

Danz, N. P., G. J. Niemi, R. R. Regal, T. Hollenhorst, L. B. Johnson, J. M. Hanowski, R. P. Axler, J. J. H. Ciborowski, T. Hrabik, V. J. Brady, J. R. Kelly, J. A. Morrice, J. C. Brazner, R. W. Howe, C. A. Johnston, and G. E. Host. 2007. Integrated measures of anthropogenic stress in the U.S. Great lakes basin. Environmental management 39:631–47.

DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–45.

Fawcett, T. 2006. An introduction to ROC analysis. Pattern Recognition Letters 27:861–874.

Folke, C., S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. S. Holling. 2004. Regime shifts, resilience, and biodiversity in ecosystem management. Annual Review of Ecology, Evolution, and Systematics 35:557–581.

Frey, D. G. 1977. Biological integrity of water - An Historical Approach. Pages 127–140 The Integrety of Water. United States Environmental Protection Agency, Washington, D.C.

Grabas, G. P., E. A. Blukacz-Richards, and S. Pernanen. 2012. Development of a submerged aquatic vegetation community index of biotic integrity for use in Lake Ontario coastal wetlands. Journal of Great Lakes Research 38:243–250.

Groffman, P. M., J. S. Baron, T. Blett, A. J. Gold, I. Goodman, L. H. Gunderson, B. M. Levinson, M. a. Palmer, H. W. Paerl, G. D. Peterson, N. L. Poff, D. W. Rejeski, J. F. Reynolds, M. G. Turner, K. C. Weathers, and J. Wiens. 2006. Ecological thresholds: The key to successful environmental management or an important concept with no practical application? Ecosystems 9:1–13.

Guntenspergen, G. R., editor. 2014. Application of threshold concepts in natural resource decision making. Springer New York, New York, NY.

Hand, D. J., and R. J. Hill. 2001. A simple generalisation of the area under the ROC curve for multiple class classification problems. Machine Learning 45:171–186.

Hawkins, C. E. 2006. Quatifying biological integrity by taxonomic completeness: its itility in regional and global assessments. Ecological Applications 16:1277–1294.

Hering, D., A. Borja, J. Carstensen, L. Carvalho, M. Elliott, C. K. Feld, A.-S. Heiskanen, R. K. Johnson, J. Moe, D. Pont, A. L. Solheim, and W. Van de Bund. 2010. The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. The Science of the Total Environment 408:4007–19.

Holm, E., Mandrak, Nicholas E, and M. E. Burridge. 2010. The ROM field guide to freshwater fishes of Ontario. Royal Ontario Museum, Toronto.

Hondorp, D. W., E. F. Roseman, and B. A. Manny. 2014. An ecological basis for future fish habitat restoration efforts in the Huron-Erie Corridor. Journal of Great Lakes Research 40:23–30.

Host, G. E., T. N. Brown, T. P. Hollenhorst, L. B. Johnson, and J. J. H. Ciborowski. 2011. High-resolution assessment and visualization of environmental stressors in the Lake Superior basin. Aquatic Ecosystem Health & Management 14:376–385.

Hughes, R. M., P. R. Kaufmann, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fisheries and Aquatic Sciences 55:1618–1631.

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. 1991. Biological Integrity: A Long-Neglected Aspect of Water Resource Management. Ecological Applications 1:66.

Karr, J. R., and D. R. Dudley. 1981. Ecological perspective on water quality goals. Environmental Management 5:55–68.

Liu, C., M. White, and G. Newell. 2011. Measuring and comparing the accuracy of species distribution models with presence-absence data. Ecography 34:232–243.

Metz, C. E. 1978. Basic principles of ROC analysis. Seminars in Nuclear Medicine 8:283–298.

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.

Mouton, A. M., B. De Baets, and P. L. M. Goethals. 2010. Ecological relevance of performance criteria for species distribution models. Ecological Modelling 221:1995–2002.

Niemi, G. J., J. R. Kelly, and N. P. Danz. 2007. Environmental indicators for the coastal region of the North American Great Lakes: Introduction and prospectus. Journal of Great Lakes Research 33:1–12.

Niemi, G. J., E. D. Reavie, G. S. Peterson, J. R. Kelly, C. a. Johnston, L. B. Johnson, R. W. Howe, G. E. Host, T. P. Hollenhorst, N. P. Danz, J. J. H. Ciborowski, T. N. Brown, V. J. Brady, and R. P. Axler. 2011. An integrated approach to assessing multiple stressors for coastal Lake Superior. Aquatic Ecosystem Health & Management 14:356–375.

Obuchowski, N. A., S. Karlik, and C. Reinhold. 2005. Fundamentals of clinical research for radiologists - ROC analysis. American Journal of Roentgenology 184:364–372.

Park, S. H., J. M. Goo, and C.-H. Jo. 2004. Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean Journal of Radiology 5:11–8.

R Core Team. 2014. R: A language and environment of statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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. Austral Ecology 20:198–219.

Robin, X., N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, and M. Müller. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics 12:77.

Seilheimer, T. S., and P. Chow-Fraser. 2006. Development and use of the wetland fish index to assess the quality of coastal wetlands in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 366:354–366.

Seilheimer, T. S., and P. Chow-Fraser. 2007. Application of the Wetland Fish Index to Northern Great Lakes marshes with emphasis on Georgian Bay coastal wetlands. Journal of Great Lakes Research 33:154–171.

Sierszen, M. E., J. a. Morrice, A. S. Trebitz, and J. C. Hoffman. 2012. A review of selected ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. Aquatic Ecosystem Health & Management 15:92–106.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16:1267–76.

Uzarski, D. G., T. M. Burton, M. J. Cooper, J. W. Ingram, and S. T. A. 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 31:171–187.

Uzarski, D. G., T. M. Burton, and J. A. Genet. 2004. Validation and performance of an invertebrate index of biotic integrity for Lakes Huron and Michigan fringing wetlands during a period of lake level decline. Aquatic Ecosystem Health & Management 7:269–288.

Whillans, T. H. 1982. Changes in marsh area along the Canadian shore of Lake Ontario. Journal of Great Lakes Research 8:570–577.

Whittier, T. R., R. M. Hughes, J. L. Stoddard, G. A. Lomnicky, D. V. Peck, and A. T. Herlihy. 2007. A structured approach for developing indices of biotic integrity:

Three examples from streams and rivers in the Western USA. Transactions of the American Fisheries Society 136:718–735.

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.

Wu, N., B. Schmalz, and N. Fohrer. 2012. Development and testing of a phytoplankton index of biotic integrity (P-IBI) for a German lowland river. Ecological Indicators 13:158–167.

Yates, A. G., and R. C. Bailey. 2010. Selecting objectively defined reference sites for stream bioassessment programs. Environmental monitoring and assessment 170:129–40.

CHAPTER 3

USING FISH-BASED BIOLOGICAL INDICATORS TO DIAGNOSE LAND USE-BASED ANTHROPOGENIC STRESS IN GREAT LAKES COASTAL WETLANDS

Abstract

Biological indicator models can be used to assess the effects of human disturbance on biological communities and to diagnose possible sources of stress. If an indicator is to be used to make actionable decision with respect to human impacts on wetlands, the indicator must first be calibrated with a scale of human disturbance. Land use-based measures of stress directly measure the risk of stress due to anthropogenic activity. In this study, four fish-based biological indicator models of wetland condition were evaluated to determine whether they can accurately classify sites as degraded or not-degraded based on various land use-based stresses. Additionally, I use Receiver-Operator Curve Characteristic (ROC) analysis to determine biological index score cut-points that most effectively distinguish non-degraded from degraded sites. Using data collected by the US EPA-funded Great Lakes Environmental Indicator (GLEI) and Coastal wetland Monitoring (CWM) projects, the characteristic biological condition scores derived from fish assemblages at wetlands were tested against scales of agricultural land use stress, development and population-based stress, as well as a cumulative measure combining both agricultural and development based stresses. The Lily-IBI, Wetland Fish Index and fish Index of Ecological Condition were found to be accurate classifiers of cumulative stress, while only the Lily-IBI and Wetland Fish Index were found to be accurate classifiers of agricultural stress. No biological indicators accurately classified sites based on development-based stress.

Introduction

Several fish-based biological indicator models have seen initial development for use in Great Lakes coastal wetlands (Seilheimer & Chow-Fraser 2007, Howe *et al.* 2007, Uzarski *et al.* 2007). These models are meant to indicate the condition of the biological community and quantify the amount of degradation a wetland experiences due to anthropogenic activity. While biological indicators can be used as an end-point in environmental monitoring as a measure of the condition of the biological community (Karr *et al.* 1991), many (Dale & Beyeler 2001, Cairns *et al.* 2003, Niemi *et al.* 2004, Meador *et al.* 2008, Quataert *et al.* 2001, Murphy *et al.* 2013) argue that indicators can also be useful tools in diagnosing sources of stress. For these models to be useful tools in the assessment of anthropogenic stress on wetlands, they must be evaluated to determine whether they can accurately indicate anthropogenic stress.

Quantifying Anthropogenic Impacts on Wetlands

Although biological indicators share similar goals in assessing the effect of anthropogenic activity on the biota, current indicators have been developed under different assumptions regarding how environmental condition is assessed. This is because there is no universally agreed upon measure of anthropogenic stress. Many indicators have used chemical variables related to water quality (e.g. phosphorus concentrations, or turbidity) in their initial development (Seilheimer & Chow-Fraser 2006, Grabas *et al.* 2012, Wilson & Bayley 2012, Wu *et al.* 2012). However, according to Yates and Bailey (2010) chemical based measures of anthropogenic impact show only possible *effects* of human activity, and that these measures show naturally high variability (Bailey *et al.*

2007, Lucena-Moya *et al.* 2012). Therefore, they argue that if biological indicators are to be related meaningfully to human activity, then human activity should be directly measured and not its effects (Yates and Bailey 2010).

Water chemistry-based measures of stressors are driven largely by land use. Varanka & Luoto (2012) demonstrated a strong correlation in phosphorus and nitrogen concentrations to both increasing agricultural land use and decreasing forested land. Likewise, Peterson *et al.* (2007) showed that nitrogen, specifically in Great Lakes coastal wetlands, is clearly linked to agricultural land use. Further, Chow-Fraser (2003) found a strong correlation between land cover and a composite measure of water quality in Great Lakes wetlands that included phosphorus and nitrogen concentrations as well as pH, conductivity, temperature and turbidity.

Agricultural land use and urban development are considered to be the primary axes of land use-based anthropogenic stress (Johnson *et al.* 1997, Carpenter *et al.* 1998, Wang *et al.* 2001, Allan 2004, Foley *et al.* 2005). The importance of agricultural and urban development from nonpoint sources as measures of potential anthropogenic stress was confirmed by the Great Lakes Environmental Indicators (GLEI) project in analysis of over 200 individual measures of anthropogenic land use (Danz *et al.* 2005, Danz *et al.* 2007). Using principal components analysis, they identified five primary axes based on anthropogenic land use: agriculture, atmospheric deposition, human population, land cover, and point source pollution. Each of these axes was then found to correlate strongly with either percent agricultural land cover or percent developed land cover (Bhagat *et al.* 2007, Host *et al*. in prep).

The goal of the current study was to first to determine which fish-based biological indicator models are able to accurately classify sites as degraded or non-degraded using a land use-based disturbance scale. Four biological indices were evaluated: the Wetland Fish Index – Abundance (Seilheimer & Chow-Fraser 2007), the Index of Ecological Condition (IEC, Howe *et al.* 2007), the Fish Condition Index (FCI, Bhagat *et al.*, in prep.), and the Lily-IBI (Cooper *et al.*, in prep, but based on Uzarski *et al.* 2007). Biological indicator model scores were assessed to determine whether a given model was diagnostic of scales of agricultural-based land use stress, development-based land use stress and cumulative anthropogenic stress using receiver operating characteristic curve analysis. Next, I use ROC analysis to determine the accuracy of biological indicators when either sensitivity or specificity is more heavily weighted in importance. Finally, I used ROC analysis to identify the biological index cut-point that optimally classified sites ordinated against each of the types of land use-based stresses.

Methods

For full data collection and indicator calculation procedures, see General Methods (Chapter 1).

ROC Analysis

Receiver operating characteristic (ROC) curve analysis (see Chapter 2) was conducted using the pROC package (Robin *et al.* 2011) for the R Statistical Programming Language version 3.1.2 (R Core Team 2014). In short, to assess classification accuracy, predicted site scores (biological indicator model scores) were compared to the observed site status (degraded/non-degraded) for each wetland. Observed site status was based on watershed land use (see below). ROC curves were generated based on the sensitivity and specificity of the model, calculated for all possible indicator values. The AUC (Area Under the ROC Curve) was used as a measure of overall classification accuracy.

At times though, the relative importance of sensitivity and specificity are not equal (Mapstone 1995, Field *et al.* 2007, Mudge *et* al 2012, Connors & O'Conner 2014). For example, a manager's policy decision to designate a site as a candidate for conservation may depend on an index having high specificity. In contrast, a sampling program designed to routinely monitor a small group of sites may need to immediately document when degradation has occurred. In such a case sensitivity is the important measure of effectiveness. Therefore, it may not be adequate to assess only a model's overall accuracy. The AUC can be evaluated across a limited region of the ROC curve to assess either high sensitivity or high specificity. This produces an estimate of the partial area under the curve (pAUC; Robin *et al.* 2011, Ma *et al.* 2013) giving independent measures of the model's sensitivity and specificity. A model with a high accuracy (AUC) in the 'high sensitivity- region of the curve minimizes Type II error when classifying sites. A model with a high accuracy (AUC) in the 'high specificity- region of the curve minimizes Type I error when classifying sites.

In the current study, (pAUC) values were calculated for both high sensitivity and high specificity regions of the curve. A value of 80% was defined operationally to delineate the area of the curve designated as high sensitivity or specificity. Therefore, pAUC-sensitivity is a performance criterion measuring the classification accuracy of a biological index and is defined by the partial area under the ROC curve for a given index-stress pair in which the sensitivity is at least 80%. pAUC-specificity refers to the partial area under the ROC curve for a given index-stress pair in which the specificity is at least 80%. All pAUC values were standardized as per McClish (1989) to values between 0 and 1.

Finally, indicators can be optimized by determining the meaning of a biological indicator model's output. Most indicators produce a numeric score denoting the relative condition of a biological community (Karr 1981, Howe *et al.* 2007, Seilheimer & Chow-Fraser 2007, Uzarski *et al.* 2007, Grabas *et al.* 2012). ROC analysis evaluates a model across all possible values of the indicator to assess overall classification accuracy. However, for an indicator to be easily interpretable by the end user, the biological cut-point that delineates degraded sites from non-degraded sites must be determined.

Optimal biological cut-points were calculated for each index-stress pair based on the 'closest to topleft' method. 'Closest to topleft' is defined as the minimum value of $(1 - Sensitivity)^2 + (1 - Specificity)^2$ across all points on the ROC curve. Confidence intervals for all AUC values were calculated using the DeLong method (DeLong *et al*. 1988, Robin *et al.* 2011), while confidence intervals for pAUC values and for optimal cut-points values were calculated by bootstrapping using 10 000 random permutations of the data with replacement (Robin *et al.* 2011).

Observed Site status

The observed degradation status of sites was determined using the overall land use of the watershed draining into a given site (see Chapter 2). Three scales of land use-based stresses were considered: agricultural stress, development stress, and a cumulative measure combining both agriculture and development based stress. The agricultural stress scale was measured as the percentage by area of agricultural land cover in a watershed scaled from 0 (no agricultural land in a watershed) to 1.0 (100% of the watershed used for agriculture) according to a composite map of the Great Lakes basin compiled from Canadian and US land cover data available for 2000-2001 (Ciborowski *et al*. 2012). Development stress was measured as the percentage of developed land cover in a watershed. Cumulative stress was the Euclidean distance of combined agricultural and development stress. All stress measures have been transformed to a scale 0 to 1 with 0 signifying the absence of that stress within the site's watershed and 1 signifying maximal stress for a given stress type.

Site degradation was defined operationally as the most environmentally disturbed wetlands within the Great Lakes basin. A site was degraded if it was among the most disturbed 20% for a given land use-based stress type of all sites within the entire Great Lakes (i.e. the 20% of sites with the highest agricultural land cover, developed land cover or cumulative land use-based stress). A disturbance threshold was determined for each stress type. This threshold is the disturbance scale value that delineates degraded wetlands from non-degraded wetlands. Watershed land use data collected by the GLEI project (Host *et al.* 2005, Danz *et al.* 2007, Hollenhorst *et al.* 2007) was analyzed for the entire population of wetland sites within the Great Lakes basin to determine the amount of land cover that denoted a degraded site for each stress type.

Results

Disturbance Thresholds

A total of 3488 wetland sites across the entire Great Lakes basin were analyzed. Watershed land use is summarized in Table 3.1. Agricultural stress scores ranged from 0.000 to 0.976, with a mean value of 0.213. The disturbance threshold for agricultural stress was 0.464. Development stress scores ranged from 0.000 to 0.969, with a mean value of 0.108. The disturbance threshold for development stress was 0.151. Cumulative stress ranged from 0.000 to 0.966, with a mean value of 0.288. The disturbance threshold for cumulative stress was 0.548.

ROC Analysis

The Lily-IBI was found to have the highest overall classification accuracy when tested against agriculture-based stress (Table 3.2), with an AUC of 0.851 and had a greater accuracy at almost all points on the ROC curve (Fig. 3.1a). The WFI was also found to have a high overall classification accuracy of 0.797. No indices were found to have a high level of classification accuracy when tested against development-based stress (Fig. 3.1b, Table 3.2). However, the FCI had the highest overall accuracy with an AUC of 0.661. The Lily-IBI, WFI and IEC were found to have high overall accuracy, with AUCs of 0.839, 0.830, and 0.791 respectively (Fig. 3.1c, Table 3.2).

When tested against agricultural-based stress, at high levels of sensitivity the Lily-IBI was found to have the greatest classification accuracy, followed by the WFI then IEC (Fig. 3.2a). The corrected pAUCs for these indices were 0.780, 0.726, and 0.651 respectively (Table 3.2). Accuracy was worse than random chance at high sensitivities for the FCI. When tested against agricultural-based stress, at high levels of specificity the Lily-IBI was found to have the greatest classification accuracy, followed by the WFI then

IEC (Fig. 3.2b). The corrected pAUCs for these indices were 0.698, 0.640, and 0.537 respectively (Table 3.2). Accuracy was worse than random chance at high sensitivities for the FCI.

When tested against development-based stress, at high levels of sensitivity all indices were found to have low classification accuracies. The corrected pAUCs for all indices were below 0.600 (Table 3.2), with the FCI having the greatest accuracy with a pAUC of 0.593 (Fig. 3.3a). When tested against development-based stress, at high levels of specificity again all indices were again found to have low classification accuracies (Table 3.2). The corrected pAUCs for all indices were below 0.600, with the WFI having the greatest accuracy with a pAUC of 0.550 (Fig. 3.3b).

When tested against cumulative land cover-based stress, at high levels of sensitivity the Lily-IBI was found to have the greatest classification accuracy (Fig. 3.4a), followed by the IEC then WFI. The corrected pAUCs for these indices were 0.846, 0.755, and 0.710 respectively (Table 3.2). Accuracy was worse than random chance at high sensitivities for the FCI. When tested against cumulative land cover-based stress, at high levels of specificity the WFI was found to have the greatest classification accuracy, followed by the IEC then Lily-IBI (Fig. 3.4b). The corrected pAUCs for these indices were 0.715, 0.607, and 0.5957 respectively (Table 3.2). Accuracy was worse than random chance at high sensitivities for the FCI.

Optimal biological cut-points are shown in Table 3.3 and Figure 3.5. Optimal cut-points for agriculture-based stress are 0.66, 7.48, 39.59, and 3.05 for the FCI, IEC, Lily-IBI, and WFI respectively. For Development-based stress optimal cut-points are 7.34, 43.75, and 3.11 for the FCI, IEC, Lily-IBI, and WFI respectively. Optimal cut-points for cumulative land use-based stress are 0.70, 7.48, 39.59 and 3.5 for the FCI, IEC, Lily-IBI, and WFI respectively.

Discussion

Indicators as Measures of Anthropogenic Stress

This study was undertaken to determine the relative accuracy with which selected fish-based biological indicator models distinguished wetlands that are degraded by anthropogenic stress from those that are not.

High overall classification accuracy for the Lily-IBI, WFI and IEC (Table 3.2) suggests that these indicators reflect anthropogenic stress in coastal wetlands. Despite being developed in part (Lily-IBI) or entirely (WFI) with a disturbance gradient based on water quality, these two indices showed high levels of accuracy when assessing degradation based on land use. These findings are consistent with previous studies, which have demonstrated strong correlations between water quality and land use (Johnson *et al* 1997, Chow-Fraser 2003, Morrice *et al.* 2008) and suggests that biological indices that have been calibrated with integrated water quality characteristics measured contemporaneously with biological data collections also reflect watershed-based anthropogenic pressures.

Disturbance thresholds used in this study were based on an operation definition of degradation for a wetland as being in the 20% of wetlands at most risk for stress. This definition of degradation was similar to previous studies which have used quintile based divisions of environmental condition (Coates *et al.* 2007, Hallet 2014), including the European Water Framework (Sandin & Herring 2010) which uses a scale of five quality classes (bad, poor, moderate, good, high) for all indicator. Further, the disturbance threshold values derived in this study (Table 3.1) correspond to values found in previous studies. For example, Chow-Fraser (2006) found a disturbance threshold of ~48% natural land cover when developing the Water Quality Index, which is comparable to thresholds found in the current study for agricultural land use (~46% land cover).

Indicators as Diagnostic Tools

Many researcher including Cains *et* al (1993), Dale & Beyeler (2001), and Niemi & McDonald (2004), suggest that a primary purpose of indicators is to be diagnostic of the source of degradation. In this study, biological indicator models were evaluated as measures of different types of land use-based stress to determine if they were diagnostic of these sources of stress (Table 3.2). Previous work by the Great Lakes Environmental Indicators (GLEI) project (Danz *et al*. 2007) demonstrated that human activity could be summarized according to two distinct types of pressure - one based on agricultural uses, and a second based on human urban and suburban development and associated population pressures. The Lily-IBI developed by Cooper *et al*. and the Wetland Fish Index (WFI) developed by Seilheimer & Chow-Fraser both showed high overall classification accuracy when tested against agricultural-based land use stress. However, no indicators were found to accurately classify sites when tested against development-based stress.

The failure of the 5 fish-based biological indices to accurately classify sites based on development pressures could be due to biases in the models' development disturbance gradients. WFI and Lily-IBI were developed using water quality gradient that included measures associated with nutrient loading such as total nitrogen, and total phosphorus.

These measures have been shown to be more highly associated with agricultural land use than urbanization. For example Lenat & Crawford (1998) found greater concentrations of total phosphorus, nitrate/nitrite, and ammonium in streams within agricultural catchments compared to streams in urban catchments. Likewise, Morrice *et al.* (2008) found a much higher correlation between total nitrogen concentrations and agriculture-based stress than with human population-based pressures.

Alternatively, poor classification accuracy in sites heavily affected by development-based stress may be due to the scale used to measure this stress. In the current study, all stress gradient data were based on watershed level land use. However, previous research by Wang *et al.* (2001) found that urbanization within 50 m of a stream had a much greater influence on fish communities than did urbanization measured at larger buffer sizes. Similarly, Wang *et al.* (1997) found only a weak correlation between watershed-wide urban land use and fish community integrity. Conversely, Allan *et al.* (1997) found that agricultural land use at the scale of the entire catchment had a very strong correlation with fish community integrity, while the relationship between biotic integrity and agriculture at only a 150-m buffer was low and non-significant.

Future research should focus on evaluating the relationship between current biological indices and urban land use at different spatial scales to determine whether the effects of urbanization or if these indices are only calibrated to agricultural-based stressors. However, it would not be advisable to diagnose risk of watershed-based development pressure with current indicators.

Which indicators are most sensitive/specific?

The AUC provides an unbiased measure of the overall accuracy of an index. However, overall AUC analysis assigns equal weight to both sensitivity and specificity. In practice, the sensitivity and specificity components of an index may be weighted according to the context and specific goal of a monitoring study. This study investigated whether the fish-based indices were biased towards certain types of classification error by determining partial areas under the ROC curve (pAUCs) in which either the sensitivity or specificity was weighted as important in assessment.

The pAUCs of all of the models indicated a tendency for the indices to be sensitive rather than specific (Table 3.2). For agricultural-based stress, the Lily-IBI and WFI showed moderate sensitivity accuracy. However, neither model was very specific. All of the models were uniformly insensate and nonspecific when evaluated with respect to development-based stress. When tested against the cumulative measure of land use-based stress, all indicators except the FCI were found to be moderately to highly sensitive, but were generally nonspecific.

In a recent validation test of the BEAST biological indicator model, Strachan and Reynoldson (2014) found that the model had low classification accuracy when applied to data from certain geographic regions. However, misclassification was primarily due to Type 1 error. They noted that in these cases, despite classification error, the BEAST model could still be useful because it was sensitive to detecting sites in the degraded condition. Results of the current study similarly highlight the importance of knowing a biological index’s strengths and limitations. While the Lily-IBI and IEC were both found to have high overall classification accuracy, analysis of pAUCs demonstrates that this accuracy reflects primarily the indices’ sensitivity and not their specificity. In other words, these indices are much less likely to erroneously diagnose wetland as degraded when it isn't; however, they are more likely to make the mistake of classifying a non-degraded wetland is degraded. The Wetland Fish Index was likewise found to have a high overall classification accuracy for cumulative stress. However, pAUC values were moderate for both sensitivity and specificity. Therefore, when applying the indices in the field, the end user should take care to determine the relative importance of sensitivity and specificity in their assessment of the wetland.

Optimal biological Cut-Points

ROC analysis is useful because it does not make any assumptions about the indicator cut-point that distinguishes a non-degraded site from one that is degraded. In practice though, indicator cut-point values are important information as they give meaning to a score produced by an indicator. ROC analysis tests the overall classification accuracy of a model by evaluating every possible biological score as a criterion for classifying a site as biologically degraded or not, and calculating the total area under the ROC curve. The point on the ROC curve that is closest to the top left corner of the plot corresponds to the single biological index score that has the greatest classification accuracy for the data making up the curve. This value can be the one that can most effectively classify sites as non-degraded vs. degraded when using the biological index.

In the current study optimal cut-point values were found for all indicator-stress type pairs (Table 3.3). The optimal cut-point score when measuring cumulative stress were 0.7 for the FCI, 7.48 for the IEC, 44 for the Lily-IBI and 3.05 for the WFI. A cut-point of 3.25 has previously been previously proposed by the model's authors (Cvetkovic

& Chow-Fraser 2011). The cut-points found in this study are the first that have been proposed for all other models.

Comparison to Previous Tests of these Indicators

The high accuracy of the WFI is in line with previous assessments of this indicator. Seilheimer *et al.* (2009) tested multiple biological indices of wetland condition and found the WFI to be highly correlated with the Water Quality Index ($R^2 = 0.75$), an independent measure of stress based on water chemistry. While the Water Quality Index has been shown to have a significant correlation to anthropogenic land use (Chow-Fraser 2003), my analysis is the first to directly test the WFI against a land use-based stress gradient. Moreover, all previous evaluations of the WFI (Seilheimer & Chow-Fraser 2006, Seilheimer & Chow-Fraser 2007, Seilheimer *et al.* 2009, Cvetkovic & Chow-Fraser 2011) were performed using the data that was used in the model's initial development. The current study is the first to test the WFI against a novel dataset. The Lily-IBI, IEC, and FCI have not previously been evaluated for their accuracy in classifying wetlands.

Summary

The purpose of this study was to first test biological indicator models as measures of anthropogenic stress defined specifically by land use. Overall, the Lily-IBI and WFI most accurately classified environmental degraded sites as biologically degraded. These indices were most accurate when classifying sites ordinated according to agricultural-based stress. They were less effective at diagnosing environmental impairment due to development. Future work should look to developing fish indices that reflect development and urbanization-based stresses. While the Lily-IBI was found to be a very sensitive measure of anthropogenic-based disturbance it lacked specificity in classification. The WFI was found to be moderately accurate when either sensitivity or specificity was more highly weighted. Finally, a Lily-IBI value of 44 and a WFI value of 3.05 were the single biological scores that were found to give the highest classification accuracy with respect to cumulative anthropogenic stress.

Table 3.1. Summary land use stress values for all wetlands and coastal margin sites in the Great Lakes basin (N = 3488 sites). The disturbance threshold is the amount of land use stress which separates degraded and non-degraded sites (i.e. 20% of all sites in the Great Lakes have a higher stress score than the disturbance threshold).

Stress Type	Median	Mean	Range	Disturbance Threshold
Agricultural	0.076	0.213	0 - 0.976	0.464
Development	0.063	0.108	0 - 0.969	0.151
Cumulative	0.194	0.288	0 - 0.966	0.548

Table 3.2. Classification accuracy of biological indicators measured as the Area Under the ROC Curve (AUC) for different types of land use stress. Bracketed values are 95% confidence intervals of the AUC.

Index	Stress	AUC
FCI	Agriculture	0.522 (0.151 - 0.893)
IEC	Agriculture	0.675 (0.532 - 0.818)
Lily-IBI	Agriculture	0.851 (0.746 - 0.955)
WFI	Agriculture	0.797 (0.677 - 0.917)
FCI	Development	0.661 (0.467 - 0.854)
IEC	Development	0.603 (0.448 - 0.759)
Lily-IBI	Development	0.586 (0.356 - 0.815)
WFI	Development	0.647 (0.491 - 0.82)
FCI	Cumulative	0.463 (0.084 - 0.842)
IEC	Cumulative	0.791 (0.674 - 0.908)
Lily-IBI	Cumulative	0.839 (0.727 - 0.951)
WFI	Cumulative	0.830 (0.715 - 0.944)

Table 3.3. Classification accuracy of biological indicators measured as the partial AUC (pAUC) for different types of land use stress. pAUC –Sensitivity is the AUC where sensitivity is at least 80%. pAUC –Specificity is the AUC where specificity is at least 80 %. Missing values signify an AUC of < 0.500 across the specified range of sensitivity or specificity. Bracketed values are 95% confidence intervals of the pAUC.

Index	Stress	pAUC - Sensitivity	pAUC - Specificity
FCI	Agriculture	-	-
IEC	Agriculture	0.651 (0.534 - 0.789)	0.537 (0.451 - 0.66)
Lily-IBI	Agriculture	0.780 (0.656 - 0.912)	0.698 (0.537 - 0.888)
WFI	Agriculture	0.726 (0.617 - 0.866)	0.640 (0.512 - 0.808)
FCI	Development	0.593 (0.484 - 0.821)	0.524 (0.444 - 0.707)
IEC	Development	0.533 (0.462 - 0.663)	0.548 (0.47 - 0.667)
Lily-IBI	Development	0.526 (0.444 - 0.752)	0.528 (0.444 - 0.686)
WFI	Development	0.521 (0.449 - 0.713)	0.550 (0.463 - 0.708)
FCI	Cumulative	-	-
IEC	Cumulative	0.755 (0.597 - 0.886)	0.607 (0.509 - 0.763)
Lily-IBI	Cumulative	0.846 (0.752 - 0.929)	0.596 (0.484 - 0.839)
WFI	Cumulative	0.710 (0.542 - 0.886)	0.715 (0.579 - 0.866)

Table 3.4. Optimal cut-point scores and the sensitivity and specificity of those cut-points for biological indicators in classifying degradation due to land use-based stress. Optimal Indicator Cut-points are the biological indicator scores which provide the greatest classification accuracy. Optimal sensitivity and optimal specificity values denote the sensitivity and specificity of the optimal indicator value. Thresholds are calculated using the 'closest to topleft' method. Bracketed values are 95% confidence intervals of the AUC.

.

Index	Stress	Optimal Indicator Cut-point	Optimal Sensitivity	Optimal Specificity
FCI	Agriculture	0.66	0.750 (0.250 - 1)	0.559 (0.382 - 0.755)
IEC	Agriculture	8.42	0.938 (0.813 - 1)	0.432 (0.296 - 0.568)
Lily-IBI	Agriculture	40	0.833 (0.667 - 0.958)	0.833 (0.700 - 0.967)
WFI	Agriculture	3.16	0.875 (0.688 - 1)	0.651 (0.512 - 0.791)
FCI	Development	0.62	0.600 (0.300 - 0.9)	0.750 (0.571 - 0.893)
IEC	Development	7.34	0.667 (0.476 - 0.857)	0.615 (0.462 - 0.769)
Lily-IBI	Development	48	0.875 (0.625 - 1)	0.348 (0.217 - 0.478)
WFI	Development	3.01	0.571 (0.381 - 0.762)	0.763 (0.632 - 0.895)
FCI	Cumulative	0.7	0.500 (0 - 1)	0.706 (0.559 - 0.853)
IEC	Cumulative	7.48	0.905 (0.762 - 1)	0.718 (0.564 - 0.846)
Lily-IBI	Cumulative	44	0.957 (0.870 - 1)	0.677 (0.516 - 0.839)
WFI	Cumulative	3.05	0.762 (0.571 - 0.952)	0.816 (0.684 - 0.921)

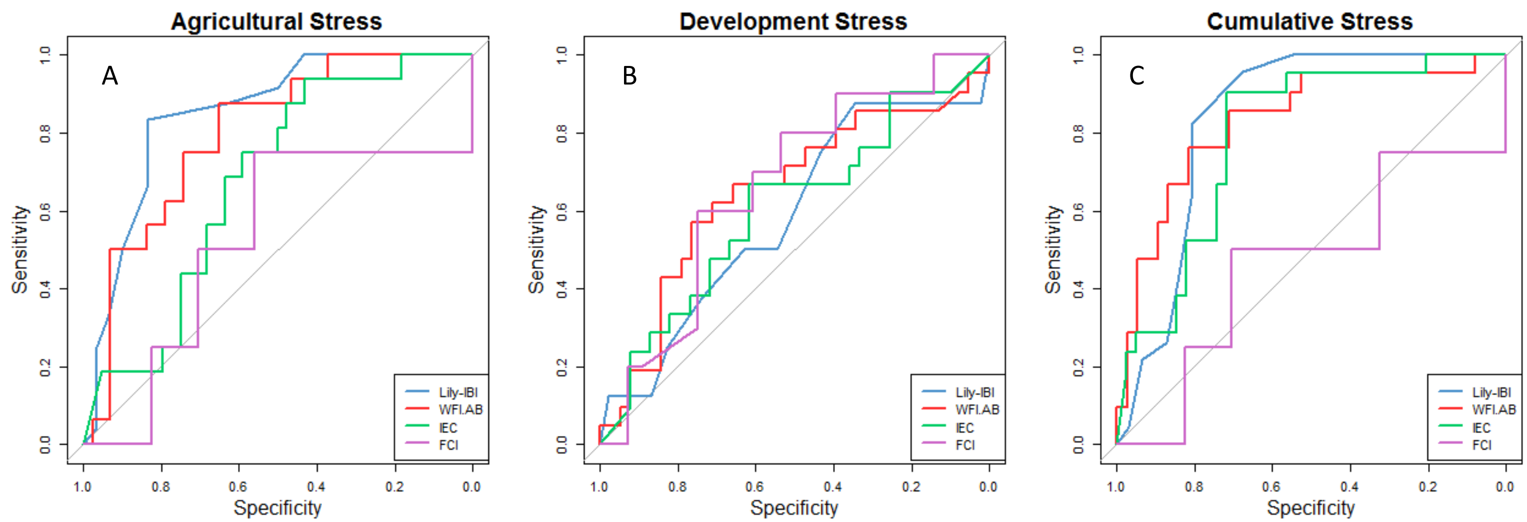

Figure 3.1. Receiver Operating Characteristic curves for Great Lakes biological indicators for agricultural-based stress (A), development-based stress (B), or cumulative stress (C). Lily-IBI = blue; WFI = red; IEC = green; FCI = purple.

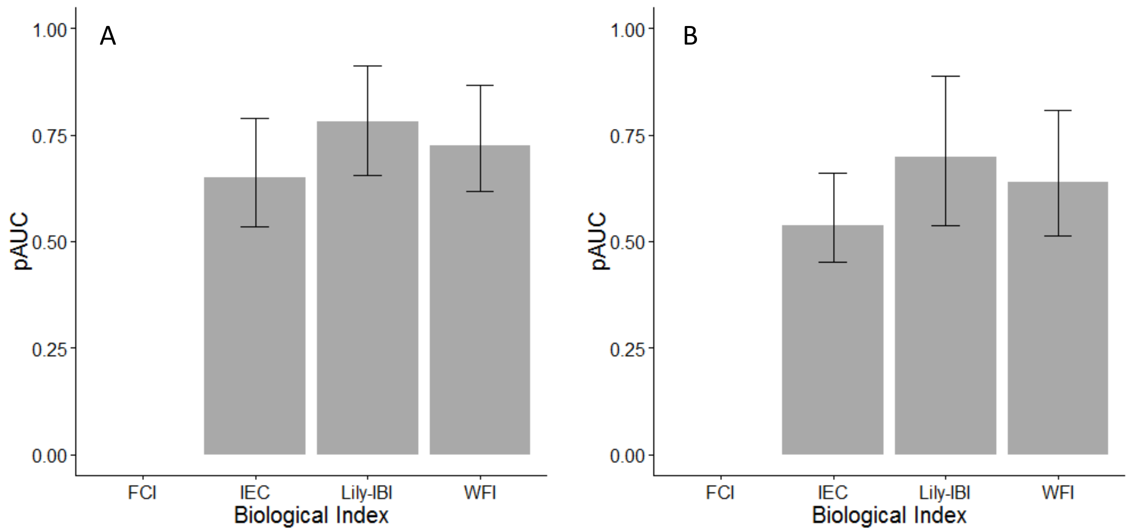

Figure 3.2. pAUC values showing the classification accuracy of biological indicator for agriculture-based land use stress when either sensitivity or specificity is weighted heavier. pAUC –Sensitivity (A) is the AUC where sensitivity is at least 80%. pAUC –Specificity (B) is the AUC where specificity is at least 80 %. Missing values signify an AUC of < 0.500 across the specified range of sensitivity or specificity. Error bars are 95% confidence intervals of the pAUC.

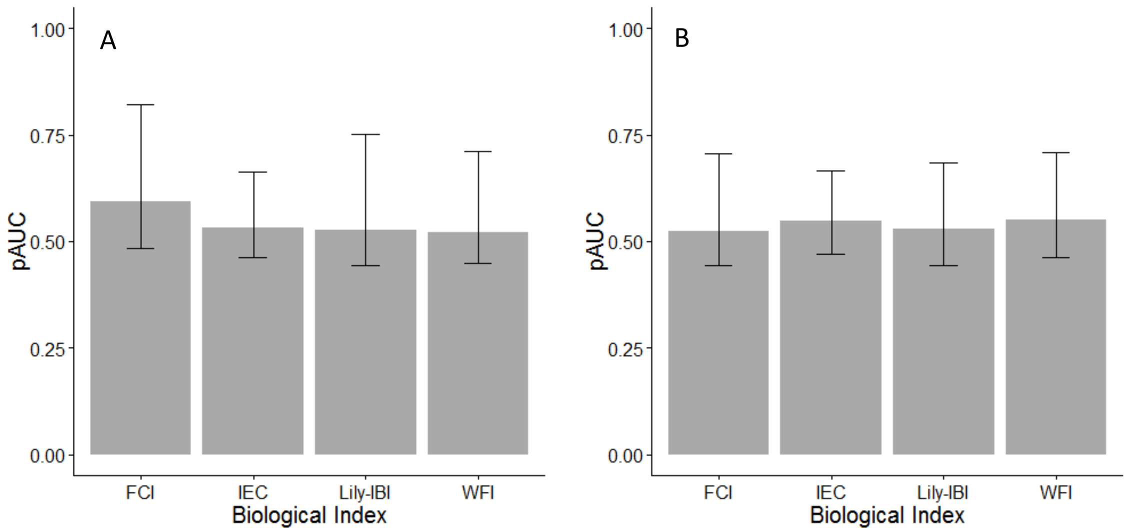

Figure 3.3. pAUC values showing the classification accuracy of biological indicator for development-based land use stress when either sensitivity or specificity is weighted heavier. pAUC –Sensitivity (A) is the AUC where sensitivity is at least 80%. pAUC – Specificity (B) is the AUC where specificity is at least 80 %. Missing values signify an AUC of < 0.500 across the specified range of sensitivity or specificity. Error bars are 95% confidence intervals of the pAUC.

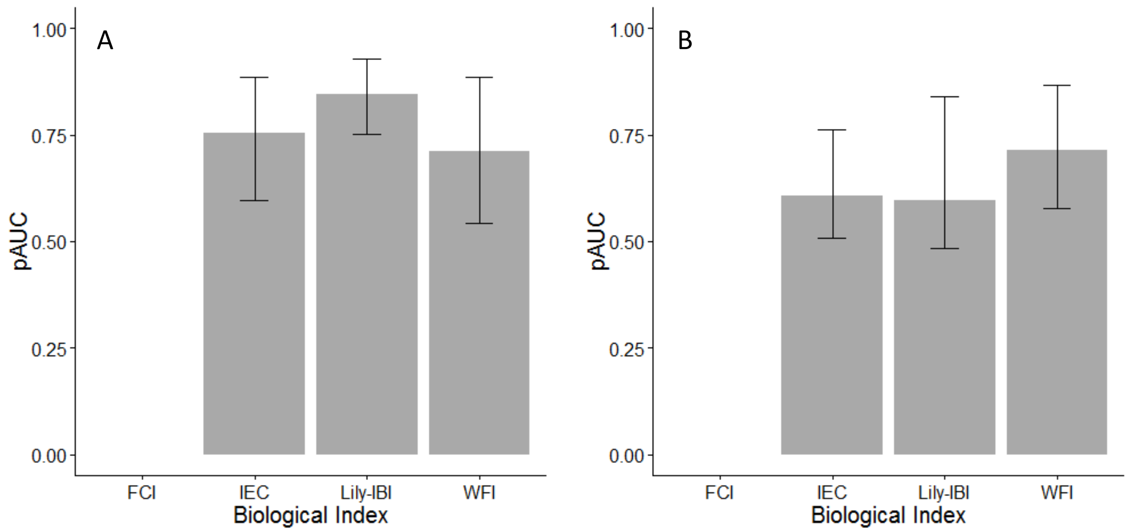

Figure 3.4. pAUC values showing the classification accuracy of biological indicator for cumulative land use-based stress when either sensitivity or specificity is weighted heavier. pAUC –Sensitivity (A) is the AUC where sensitivity is at least 80%. pAUC – Specificity (B) is the AUC where specificity is at least 80 %. Missing values signify an AUC of < 0.500 across the specified range of sensitivity or specificity. Error bars are 95% confidence intervals of the pAUC.

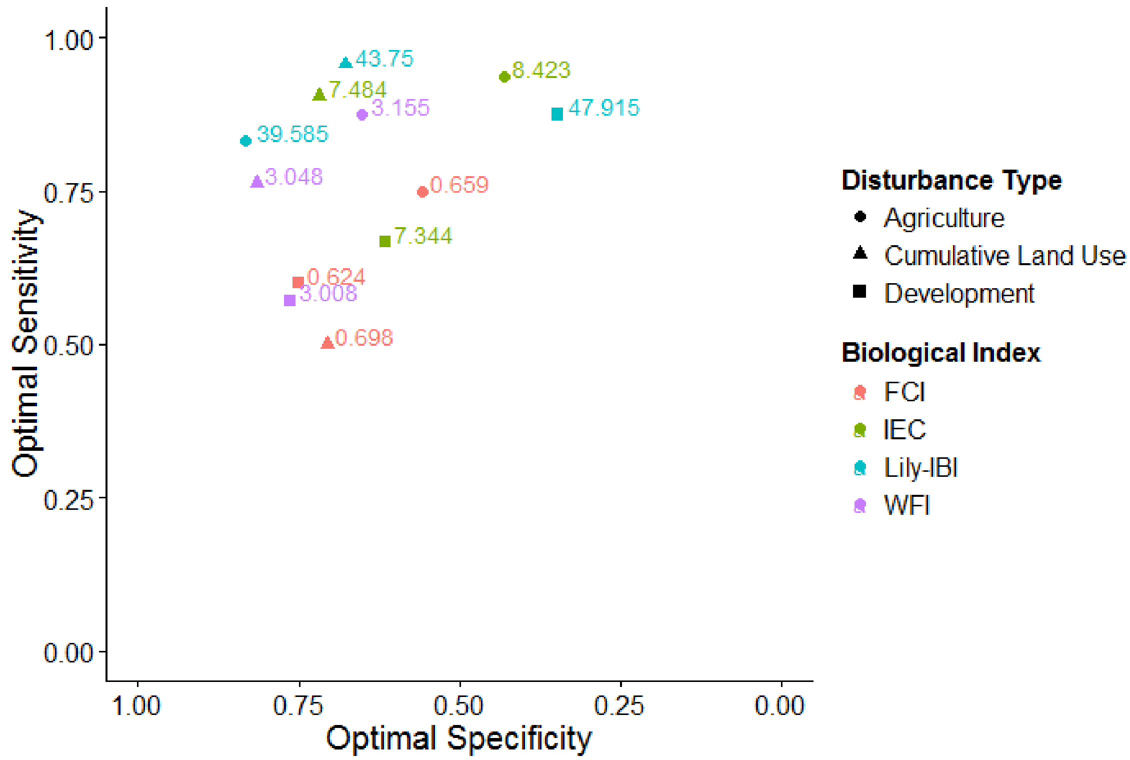

Figure 3.5. Comparison of the sensitivity and specificity of optimal biological indicator score cut-points. Cut-points are calculated using the ‘closest to topleft’ method.

References

Allan, J. D., D. L. Erickson, and J. Fay. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37:149–161.

Allan, J. D. 2004. Landscapes and riverscapes: The influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics 35:257–284.

Bailey, R. C., T. B. Reynoldson, A. G. Yates, J. Bailey, and S. Linke. 2007. Integrating stream bioassessment and landscape ecology as a tool for land use planning. Freshwater Biology 52:908–917.

Bhagat, Y. 2005. Fish indicators of anthropogenic stress at Great Lakes coastal margins: multimetric and multivariate approaches. University of Windsor.

Booth, D. B., and L. E. Reinelt. 1993. Consequences of urbanization on aquatic systems - measured effects, degradation thresholds, and corrective strategies. Pages 545–550 Proc. of Watershed '93. U.S. Environmental Protection Agency.

Cairns, J., P. V. McCormick, and B. R. Niederlehner. 1993. A proposed framework for developing indicators of ecosystem health. Hydrobiologia 263:1–44.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8:559.

Chow-Fraser, P. 2006. Development of the Water Quality Index (WQI) to assess effects of basin-wide land-use alteration on coastal marshes of the Laurentian Great Lakes. Pages 137–166 *in* T.P. Simon and P.M. Stewart. Indiana Biological Survey, editor. Coastal wetlands of the Laurentian Great Lakes: health, habitat and indicators.

Coates, S., A. Waugh, A. Anwar, and M. Robson. 2007. Efficacy of a multi-metric fish index as an analysis tool for the transitional fish component of the Water Framework Directive. Marine Pollution Bulletin 55:225–40.

Cvetkovic, M., and P. Chow-Fraser. 2011. Use of ecological indicators to assess the quality of Great Lakes coastal wetlands. Ecological Indicators 11:1609–1622.

Dale, V. H., and S. C. Beyeler. 2001. Challenges in the development and use of ecological indicators. Ecological Indicators 1:3–10.

Danz, N. P., G. J. Niemi, R. R. Regal, T. Hollenhorst, L. B. Johnson, J. M. Hanowski, R. P. Axler, J. J. H. Ciborowski, T. Hrabik, V. J. Brady, J. R. Kelly, J. A. Morrice, J. C. Brazner, R. W. Howe, C. A. Johnston, and G. E. Host. 2007. Integrated measures of anthropogenic stress in the U.S. Great lakes basin. Environmental management 39:631–47.

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.

DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–45.

Fawcett, T. 2006. An introduction to ROC analysis. Pattern Recognition Letters 27:861–874.

Field, S. A., P. J. O'Connor, A. J. Tyre, and H. P. Possingham. 2007. Making monitoring meaningful. Austral Ecology 32:485–491.

Foley, J. A., R. Defries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, F. S. Chapin, M. T. Coe, G. C. Daily, H. K. Gibbs, J. H. Helkowski, T. Holloway, E. A. Howard, C. J. Kucharik, C. Monfreda, J. A. Patz, I. C. Prentice, N. Ramankutty, and P. K. Snyder. 2005. Global consequences of land use. Science 309:570–4.

Grabas, G. P., E. A. Blukacz-Richards, and S. Pernanen. 2012. Development of a submerged aquatic vegetation community index of biotic integrity for use in Lake Ontario coastal wetlands. Journal of Great Lakes Research 38:243–250.

Graham, S. A., and I. A. Mendelssohn. 2014. Coastal wetland stability maintained through counterbalancing accretionary responses to chronic nutrient enrichment. Ecology 95:3271–3283.

Hallett, C. S. 2014. Quantile-based grading improves the effectiveness of a multimetric index as a tool for communicating estuarine condition. Ecological Indicators 39:84–87.

Hollenhorst, T. P., T. N. Brown, L. B. Johnson, J. J. H. Ciborowski, and G. E. Host. 2007. Methods for generating multi-scale watershed delineations for indicator development in Great Lake coastal ecosystems. Journal of Great Lakes Research 33:13–26.

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 26:5325–5342.

Howe, R. W., R. R. Regal, J. Hanowski, G. J. Niemi, N. P. Danz, and C. R. Smith. 2007. An index of ecological condition based on bird assemblages in Great Lakes coastal wetlands. Journal of Great Lakes Research 33(sp3):93–105.

Johnson, L., C. Richards, G. Host, and J. Arthur. 1997. Landscape influences on water chemistry in Midwestern stream ecosystems. Freshwater Biology 37:193–208.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21–27.

Karr, J. R. 1991. Biological Integrity: A Long-Neglected Aspect of Water Resource Management. Ecological Applications 1:66.

Lenat, D. R., and J. K. Crawford. 1994. Effects of land use on water quality and aquatic biota of three North Carolina Piedmont streams. Hydrobiologia 294:185–199.

Lucena-Moya, P., C. Gómez-Rodríguez, and I. Pardo. 2012. Spatio-temporal variability in water chemistry of Mediterranean coastal lagoons and its management implications. Wetlands 32:1033–1045.

Mapstone, B. D. 1995. Scalable decision rules for environmental impact studies: Effect size, type I, and type II errors. Ecological Applications 5: 401–410.

Ma, H., A. I. Bandos, H. E. Rockette, and D. Gur. 2013. On use of partial area under the ROC curve for evaluation of diagnostic performance. Statistics in Medicine 32:3449–58.

McClish, D. K. 1989. Analyzing a Portion of the ROC Curve. Medical Decision Making 9:190–195.

Meador, M. R., D. M. Carlisle, and J. F. Coles. 2008. Use of tolerance values to diagnose water-quality stressors to aquatic biota in New England streams. Ecological Indicators 8:718–728.

Metz, C. E. 1978. Basic principles of ROC analysis. Seminars in Nuclear Medicine 8:283–298.

Morrice, J. A, N. P. Danz, R. R. Regal, J. R. Kelly, G. J. Niemi, E. D. Reavie, T. Hollenhorst, R. P. Axler, A. S. Trebitz, A. M. Cotter, and G. S. Peterson. 2008. Human influences on water quality in Great Lakes coastal wetlands. Environmental Management 41:347–57.

Mudge, J. F., L. F. Baker, C. B. Edge, and J. E. Houlahan. 2012. Setting an optimal alpha that minimizes errors in null hypothesis significance tests. PLoS ONE 7(2): e32734.

Murphy, J. F., J. Davy-Bowker, B. McFarland, and S. J. Ormerod. 2013. A diagnostic biotic index for assessing acidity in sensitive streams in Britain. Ecological Indicators 24:562–572.

Niemi, G. J., and M. E. McDonald. 2004. Application of ecological indicators. Annual Review of Ecology, Evolution, and Systematics 35:89–111.

Niemi, G., D. Wardrop, R. Brooks, S. Anderson, V. Brady, H. Paerl, C. Rakocinski, M. Brouwer, B. Levinson, and M. McDonald. 2004. Rationale for a new generation of indicators for coastal waters. Environmental Health Perspectives 112:979–986.

Peterson, G. S., M. E. Sierszen, P. M. Yurista, and J. R. Kelly. 2007. Stable nitrogen isotopes of plankton and benthos reflect a landscape-level influence on Great Lakes coastal ecosystems. Journal of Great Lakes Research 33:27–41.

Quataert, P., P. Verschelde, J. Breine, G. Verbeke, E. Goetghebeur, and F. Ollevier. 2011. A diagnostic modelling framework to construct indices of biotic integrity: A case study of fish in the Zeeschelde estuary (Belgium). Estuarine, Coastal and Shelf Science 94:222–233.

Robin, X., N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, and M. Müller. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics 12:77.

Sandin, L., and D. Hering. 2004. Comparing macroinvertebrate indices to detect organic pollution across Europe: a contribution to the EC Water Framework Directive intercalibration. Hydrobiologia 516:55–68.

Dos Santos, D. A., C. Molineri, M. C. Reynaga, and C. Basualdo. 2011. Which index is the best to assess stream health? Ecological Indicators 11:582–589.

Seilheimer, T. S., and P. Chow-Fraser. 2006. Development and use of the wetland fish index to assess the quality of coastal wetlands in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 366:354–366.

Seilheimer, T. S., and P. Chow-Fraser. 2007. Application of the Wetland Fish Index to Northern Great Lakes marshes with emphasis on Georgian Bay coastal wetlands. Journal of Great Lakes Research 33:154–171.

Seilheimer, T. S., T. P. Mahoney, and P. Chow-Fraser. 2009. Comparative study of ecological indices for assessing human-induced disturbance in coastal wetlands of the Laurentian Great Lakes. Ecological Indicators 9:81–91.

Strachan, S. A., and T. B. Reynoldson. 2014. Performance of the standard CABIN method: comparison of BEAST models and error rates to detect simulated degradation from multiple data sets. Freshwater Science 33:1225–1237.

Uzarski, D. G., T. M. Burton, M. J. Cooper, J. W. Ingram, and S. T. a. 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 31:171–187.

Vanacker, M., A. Wezel, V. Payet, and J. Robin. 2015. Determining tipping points in aquatic ecosystems: The case of biodiversity and chlorophyll α relations in fish pond systems. Ecological Indicators 52:184–193.

Varanka, S., and M. Luoto. 2012. Environmental determinants of water quality in boreal rivers based on partitioning methods. River Research and Applications 28:1034–1046.

Wang, L., J. Lyons, P. Kanehl, and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28:255–266.

Wang, L., J. Lyons, P. Kanehl, and R. Gatti. 1997. Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. Fisheries 22:6–12.

Wilson, M. J., and S. E. Bayley. 2012. Use of single versus multiple biotic communities as indicators of biological integrity in northern prairie wetlands. Ecological Indicators 20:187–195.

Wu, N., B. Schmalz, and N. Fohrer. 2012. Development and testing of a phytoplankton index of biotic integrity (P-IBI) for a German lowland river. Ecological Indicators 13:158–167.

Yates, A. G., and R. C. Bailey. 2010. Selecting objectively defined reference sites for stream bioassessment programs. Environmental monitoring and assessment 170:129–40.

CHAPTER 4

GENERAL DISCUSSION

The goal of this thesis was to assess the performance of fish-based indicator models for Great Lakes coastal wetlands. As these models are meant to indicate the biological condition of wetland communities, performance was assessed based on the accuracy of indicator models in classifying wetlands as being in the degraded or non-degraded environmental condition based on their biological attributes.

The purpose of biological indicator models is to determine the condition of a biological community and provide a measure of the wetland's overall environmental quality. A community is in good condition is if it is capable of "supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitat of the region" (Karr & Dudley 1981). Importantly, good condition of a community is defined by the 'naturalness' of its habitat (Herring *et al*. 2003, Stoddard *et al*. 2006, Hawkins *et al*. 2010) and, therefore poor condition is defined by an 'un-natural' habitat, that is, one that has been impacted by human activity. The accuracy of a model as an indicator of biological condition can therefore be assessed by comparing it to an independent measure of the level of human impact.

Independent measures of human impact are not widely agreed upon. For example Lougheed & Chow-Fraser (2002), Seilheimer & Chow-Fraser (2006), Croft & Chow-Fraser (2007) use water quality as a measure of human impact. However, Bailey *et al*. (2007), and Yates & Bailey (2010) argue that these measures are often confounded with natural variation and are therefore not reliable measures of human impact. They argue that human activity should be measured directly in the form of land use. This thesis took two approaches to how human impact should be measured. In Chapter 2, human impact was measured using the same disturbance scale used in the original development of the index, while in Chapter 3, human impact was explicitly measured as direct land use. Interestingly, in both chapters the Wetland Fish Index and Lily-IBI, which were originally calibrated against scales related to water quality were found to accurately classify communities that were identified as degraded according to land use criteria.

Receiver operating characteristic curve analysis was used to assess classification of biological indicators in this thesis. A large body of research has been established around the use of ROC as a test of classification accuracy (Hanley & McNeil 1982, McClish 1989, Obuchowski 1994, 2005, 2006, Obuchowski *et al*. 1998, 2004, Walter 2005, Fawcett 2006) and it has been applied to clinical medicine (Zweig & Campbell 1993), radiology (Obuchowski 2003, Park *et al*. 2008), psychiatry (Streiner & Cairney 2007), machine learning (Hand & Hill 2001, Fawcett 2006) and species distribution modelling (Fielding and Bell 1997, Mouton *et al*. 2010, Lui *et al*. 2011). However, only recently has ROC curve analysis been used to assess ecological indicators (Dos Santos *et al*. 2001, Connors & Cooper 2014). The current study was the first to apply ROC analysis in the assessment of biological indicators within the Great Lakes.

The Reference Condition Approach (RCA) is commonly used to classify site condition whereby the minimally disturbed sites are deemed ‘reference’, while more impacted sites are considered non-reference (Hughes *et al*. 1986, Stoddard *et al* 2006, Hawkins *et al*. 2010). However, Palmer *et al*. (2005) recommended that indicators should be used to assess change of condition away from the degraded state in addition to change of condition as it approaches the Reference Condition. In this thesis, assessment of sites was based on a measure of the most stressed condition of wetlands in the Great Lakes and therefore assesses the accuracy of biological indicators in classifying communities as being in the degraded state. I found that the Wetland Fish Index and the Lily-IBI were able to accurate classify communities as degraded with respect to each model's original definition of degradation (Chapter 2) as well as a definition of degradation based on land use (Chapter 3).

Through the Great Lakes Water Quality Agreement (GLWQA; IJC 1978), the governments of both Canada and the United States established Areas of Concern (AOCs) as locations in which there was especial need to "restore and maintain the chemical, physical and biological integrity of the Waters of the Great Lakes", with AOCs including many coastal wetlands (IJC 2005). AOCs have experienced high levels of environmental degradation and have been targeted for restoration through the implementation of AOC-specific Remedial Action Plans. Biological indices can play an important role in monitoring as they can quantify changes in the condition of biotic communities (Karr & Chu 1997, Niemi *et al.* 2007) providing valuable information to managers regarding the success of restoration initiatives.

Further, biological indices, may have be highly sensitive (resulting in a low Type 2 classification error rate, if the null hypothesis is that the site is 'not degraded'), highly specific (low Type 1 classification error rate), or a combination of both (Fielding and Bell 1997). A Type 2 error (i.e. not detecting the degraded condition) can be costly (Rapport & Whitford 1999, Zedler 2000, Suding *et al.* 2004, Standish *et al.* 2014) when indices are used as monitoring tools. According to Rapport & Whitford (1999), ecosystems, and the Great Lakes in particular, are often resistant to rehabilitation as mechanisms of degradation can often cause further degradation themselves. For example, a degraded environment becomes vulnerable to non-native species, who themselves make the reestablishment of native species more difficult (Rapport and Whitford 1999). In effect, degraded communities can become ecologically resilient to change due to restoration efforts (Suding *et al.* 2004, Standish *et al.* 2014) and it is therefore desirable to detect possible degradation before large changes have occurred in the community. When evaluating indicators, it is then important to know both the sensitivity and specificity of an index and to ensure that indices are not biased toward Type 2 errors. The current study (Chapter 3) demonstrated that most fish based biological indices were generally more sensitive than specific and therefore less likely to commit a Type 2 error than a Type 1 error. This means that when used in a biological monitoring context, indices evaluated in this study will likely detect the degraded state.

Change points for disturbance were set a priori to the analysis as the most disturbed condition within the Great Lakes. A degraded site was operationally defined as being among the 20% most impacted sites of all Great Lakes wetlands. This operational definition of degradation was similar to others proposed by Coates *et al*. (2007), Yates & Bailey (2010), and Hallet (2014), who used proportional thresholds of 20 %, 25 %, and 10 % respectively to define degraded sites. However, using a definition of degraded based purely on human activity may not necessarily correspond to levels of degradation that significantly affect biological communities.

Recent work by Kovalenko *et al*. (2014) could be used to set threshold levels of disturbance that specifically reflect stress levels at which changes in biological communities are observed. Kovalenko *et al*. (2014) used Threshold Indicator Taxon Analysis (TITAN, Baker & King 2010) to find threshold points on an anthropogenic disturbance gradient at which large changes in the biological community's composition occurred. Transitions were observed across multiple taxa (including fish) between reference/non-reference communities and non-reference/degraded communities. Future research should assess the ability of biological indicator models to classifying wetlands according to the disturbance thresholds determined by Kovalenko *et al*. relative to the performance of operationally defined changepoints such as those that I used. An indicator that is able to accurately classify wetlands in this way could be used to monitor sites and assess their risk of large negative shifts in community composition.

A second implication of the work of Kovalenko *et al*. (2014) is that fish community changes occur not at a single threshold of human impacts, but at two; one threshold at which sensitive species are reduced in abundance or extirpated and a second one at which disturbance tolerant species appear or become dominant. Therefore, binary classification of communities, as either reference/non-reference as seen in the Reference Condition Approach or as degraded/non-degraded as in this current study, may not be the ideal method of site evaluation. An indicator that has been shown to accurately classify biological communities at both thresholds would provide better information when assessing sites. An extension of ROC analysis called Multiclass ROC analysis (Hand & Hill 2001, Robin *et al*. 2011) uses the same principles of traditional ROC analysis except that multiple classes of disturbance (thresholds) can be applied to the analysis. Future work should look to assess the classification accuracy of biological indicators using both types of thresholds identified by Kovalenko *et al*.

GENERAL SUMMARY

Biological indicator models are a powerful tool for monitoring and managing biological communities. Individual, water chemistry-based indicators do not necessarily give a complete view of the cumulative effect of human impact on the biota. Reliable biological indicators model the community as a whole to give a value that is indicative of the community's condition. The desired biological condition can then be used as an endpoint in decision making by managers and policy makers. Reliable biological indicators will predictably reflect human impact on the community. In this thesis the criterion for reliability was the accuracy of the model in classifying wetlands as being degraded, with degradation being defined by the greatest amount of land alteration in contributing watersheds in the Great Lakes.

The Cooper-IBIs (particularly the Lily-IBI) and Wetland Fish Indices exhibited the highest classification accuracy. This accuracy was demonstrated when indicators were validated with the same disturbance gradient used in their development (Chapter 2), or with a novel disturbance gradient based purely on a direct measure of human impact (Chapter 3). The Lily-IBI was found to have the highest overall classification accuracy (Chapters 2 & 3). However, this index's accuracy was found to be weighted much more heavily on its sensitivity to degradation rather than its specificity. The Wetland Fish Index, while slightly lower in overall classification accuracy, did not exhibit a large difference between its sensitivity and specificity. Finally, optimal thresholds were found for all indices, which can be used in their future application (Chapter 3).

References

Baker, M. E., and R. S. King. 2010. A new method for detecting and interpreting biodiversity and ecological community thresholds. Methods in Ecology and Evolution 1:25–37.

Coates, S., A. Waugh, A. Anwar, and M. Robson. 2007. Efficacy of a multi-metric fish index as an analysis tool for the transitional fish component of the Water Framework Directive. Marine Pollution Bulletin 55:225–40.

Connors, B. M., and A. B. Cooper. 2014, July 22. Determining decision thresholds and evaluating indicators when conservation status is measured as a continuum. Conservation Biology 28:1626–35.

Davies, T. D., and J. K. Baum. 2012. Extinction risk and overfishing: reconciling conservation and fisheries perspectives on the status of marine fishes. Scientific Reports 2:561.

Fawcett, T. 2006. An introduction to ROC analysis. Pattern Recognition Letters 27:861–874.

Hallett, C. S. 2014. Quantile-based grading improves the effectiveness of a multimetric index as a tool for communicating estuarine condition. Ecological Indicators 39:84–87.

Hand, D. J., and R. J. Hill. 2001. A simple generalisation of the area under the ROC curve for multiple class classification problems. Machine Learning 45:171–186.

Hanley, J. A., and B. J. McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36.

Hawkins, C. P., J. R. Olson, and R. A. Hill. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society 29:312–343.

Hoffmann, M., C. Hilton-Taylor, A. Angulo, M. Böhm, T. M. Brooks, S. H. M. Butchart, K. E. Carpenter, *et al*. 2010. The impact of conservation on the status of the world's vertebrates. Science 330:1503–9.

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.

International Joint Commission. 1978. Great Lakes Water Quality Agreement. http://www.ijc.org/en_/Great_Lakes_Water_Quality.

International Joint Commission. 2005. A Guide to the Great Lakes Water Quality Agreement: background for the 2006 governmental review. Page 23.

Karr, J. R., and E. W. Chu. 1997. Biological monitoring: Essential foundation for ecological risk assessment. Human and Ecological Risk Assessment: An International Journal 3:993–1004.

Karr, J. R., and D. R. Dudley. 1981. Ecological perspective on water quality goals. Environmental Management 5:55–68.

Kovalenko, K. E., V. J. Brady, T. N. Brown, J. J. H. Ciborowski, N. P. Danz, J. P. Gathman, G. E. Host, R. W. Howe, L. B. Johnson, G. J. Niemi, and E. D. Reavie. 2014. Congruence of community thresholds in response to anthropogenic stress in Great Lakes coastal wetlands. Freshwater Science 33:958–971.

Liu, C., M. White, and G. Newell. 2011. Measuring and comparing the accuracy of species distribution models with presence-absence data. Ecography 34:232–243.

Lougheed, V. L., and P. Chow-Fraser. 2002. Development and use of a zooplankton index of wetland quality in the Laurentian Great Lakes basin. Ecological Applications 12:474–486.

McClish, D. K. 1989. Analyzing a Portion of the ROC Curve. Medical Decision Making 9:190–195.

Mouton, A. M., B. De Baets, and P. L. M. Goethals. 2010. Ecological relevance of performance criteria for species distribution models. Ecological Modelling 221:1995–2002.

Niemi, G. J., J. R. Kelly, and N. P. Danz. 2007. Environmental indicators for the coastal region of the North American Great Lakes: Introduction and prospectus. Journal of Great Lakes Research 33:1–12.

Obuchowski, N. a, M. L. Lieber, and F. H. Wians. 2004. ROC curves in clinical chemistry: uses, misuses, and possible solutions. Clinical chemistry 50:1118–25.

Obuchowski, N. A. 1994. Computing sample size for receiver operating characteristic studies. Investigative Radiology 29:238–243.

Obuchowski, N. A. 2003. Receiver operating characteristic curves and their use in radiology. Radiology 229:3–8.

Obuchowski, N. A. 2005. Estimating and comparing diagnostic tests' accuracy when the gold standard is not binary. Academic radiology 12:1198–204.

Obuchowski, N. A. 2006. An ROC-type measure of diagnostic accuracy when the gold standard is continuous-scale. Statistics in medicine 25:481–93.

Obuchowski, N. A., and M. Lieber. 1998. Confidence intervals for the receiver operating characteristic area in studies with small samples. Academic Radiology 5:561–71.

Palmer, M. A., E. S. Bernhadt, J. D. Allan, P. S. Lake, G. Alexander, S. Brooks, J. Carr, S. Clayton, C. N. Dahm, J. Follstad Shah, D. L. Galat, S. G. Loss, P. Goodwin, D. D. Hart, B. Hassett, R. Jenkinson, G. M. Kondolf, R. Lave, J. L. Meyer, T. K. O'Donnell, L. Pagano, and E. Sudduth. 2005. Standards for ecologically successful river restoration. Journal of Applied Ecology 42:208–217.

Rapport, D. J., and W. G. Whitford. 1999. How ecosystems respond to stress: Comon properties of arid and aquatic systems. BioScience 49:193–203.

Rodrigues, A. S. L., J. D. Pilgrim, J. F. Lamoreux, M. Hoffmann, and T. M. Brooks. 2006. The value of the IUCN Red List for conservation. Trends in Ecology & Evolution 21:71–6.

Dos Santos, D. A., C. Molineri, M. C. Reynaga, and C. Basualdo. 2011. Which index is the best to assess stream health? Ecological Indicators 11:582–589.

Seilheimer, T. S., and P. Chow-Fraser. 2006. Development and use of the wetland fish index to assess the quality of coastal wetlands in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 366:354–366.

Sonego, P., A. Kocsor, and S. Pongor. 2008. ROC analysis: applications to the classification of biological sequences and 3D structures. Briefings in Bioinformatics 9:198–209.

Standish, R. J., R. J. Hobbs, M. M. Mayfield, B. T. Bestelmeyer, K. N. Suding, L. L. Battaglia, V. Eviner, C. V. Hawkes, V. M. Temperton, V. a. Cramer, J. a. Harris, J. L. Funk, and P. a. Thomas. 2014. Resilience in ecology: Abstraction, distraction, or where the action is? Biological Conservation 177:43–51.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16:1267–76.

Streiner, D. L., and J. Cairney. 2007. What's under the ROC? An introduction to receiver operating characteristics curves. Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie 52:121–8.

Suding, K. N., K. L. Gross, and G. R. Houseman. 2004. Alternative states and positive feedbacks in restoration ecology. Trends in Ecology & Evolution 19:46–53.

Walter, S. D. 2005. The partial area under the summary ROC curve. Statistics in Medicine 24:2025–40.

Yates, A. G., and R. C. Bailey. 2010. Selecting objectively defined reference sites for stream bioassessment programs. Environmental Monitoring and Assessment 170:129–40.

Zedler, J. 2000. Progress in wetland restoration ecology. Trends in Ecology & Evolution 15:402–407.

Zweig, M. H., and G. Campbell. 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry 39:561–77.

APPENDICES

Appendix 1. Coastal Wetland Monitoring site descriptions.

Site #	Zone	Site Name	Lake	Latitude	Longitude
5003	Bulrush	Adolphus Reach Wetland	Ontario	44.102	-76.929
630	Bulrush	Ailes Point Area Wetland #2	Huron	45.993	-84.365
1077	Bulrush	Allouez Bay Wetland	Superior	46.681	-91.982
922	Bulrush	Ashman Bay Wetland	Huron	46.495	-84.377
812	Bulrush	Baie de Wasai Wetland #1	Huron	46.467	-84.267
590	Bulrush	Cheboygan Area Wetland #2	Huron	45.655	-84.473
901	Bulrush	Chicken Island Area Wetland	Huron	46.306	-84.132
5206	Bulrush	Corisande Bay 5	Huron	45.132	-81.56
548	Bulrush	Crooked Island Wetland	Huron	45.065	-83.296
434	Bulrush	Dickenson Island Area Wetland	Erie	42.607	-82.651
619	Bulrush	Duck Bay Wetland	Huron	45.966	-84.384
515	Bulrush	East Saginaw Bay Coastal Wetland #5	Huron	43.673	-83.575
1039	Bulrush	Fish Creek Wetland #1	Superior	46.583	-90.945
637	Bulrush	Flowers Creek Wetland	Huron	45.995	-84.319
1519	Bulrush	Garden Bay Wetland	Michigan	45.772	-86.559
5408	Bulrush	Hay Bay Wetland	Huron	46.298	-83.74
7061	Bulrush	Indian Harbor Wetland	Michigan	45.799	-85.512
5510	Bulrush	Lake George 2	Huron	46.45	-84.097
7020	Bulrush	Lakeview Pond-Sandy Creek-Colwell Ponds Marsh	Ontario	43.75	-76.204
973	Bulrush	L'Anse Bay Wetland	Superior	46.749	-88.504
951	Bulrush	Laughing Whitefish River Wetland	Superior	46.524	-87.028
521	Bulrush	Linwood Area Wetland #2	Huron	43.742	-83.949
616	Bulrush	Mackinac Creek Wetland	Huron	46.002	-84.41
615	Bulrush	Mill Pond Wetland	Huron	46.007	-84.435
1681	Bulrush	Mink River Wetland	Michigan	45.241	-87.046
793	Bulrush	Munuscong Island Wetland	Huron	46.213	-84.239
792	Bulrush	Munuscong Lake Wetland #2,#3 Munuscong River Delta	Huron	46.216	-84.257
5661	Bulrush	Musky Bay Wetland 1	Huron	44.812	-79.783
494	Bulrush	Nayanguing Point Wildlife Area Wetland #2	Huron	43.845	-83.926
122	Bulrush	North Pond Area Wetland	Ontario	43.656	-76.183
776	Bulrush	Northwest Drummond Island Wetland #4	Huron	46.076	-83.69
7033	Bulrush	Oconto Marsh #2	Michigan	44.968	-87.801
1745	Bulrush	Ogontz Bay Area Wetland	Michigan	45.811	-86.774
1514	Bulrush	Ogontz Bay Wetland #3	Michigan	45.866	-86.765

920	Bulrush	Palmers Point Area Wetland #2	Huron	46.523	-84.17
917	Bulrush	Palmers Point Wetland	Huron	46.532	-84.199
778	Bulrush	Paw Point-North Scott Bay Wetland	Superior	46.061	-83.67
1469	Bulrush	Peshtigo River Wetland #1	Michigan	44.995	-87.672
5735	Bulrush	Pine Point Wetland 1	Ontario	44.098	-77.501
5746	Bulrush	Point Au Baril 1	Huron	45.604	-80.487
976	Bulrush	Portage River Wetland #1	Superior	46.989	-88.437
5785	Bulrush	Presquille Bay Marsh 7	Ontario	44.03	-77.71
5791	Bulrush	Quarry Island Wetland 1	Huron	44.843	-79.82
5792	Bulrush	Quarry Island Wetland 2	Huron	44.835	-79.812
790	Bulrush	Raber Bay Wetland	Huron	46.12	-84.06
791	Bulrush	Roach Point Wetland	Huron	46.168	-84.173
804	Bulrush	Sand Island Wetland	Huron	46.315	-84.203
660	Bulrush	Scammon Cove, Meade Island Wetland	Huron	45.952	-83.64
1522	Bulrush	South River Bay Wetland	Michigan	45.745	-86.626
1102	Bulrush	Spirit Lake Wetland #6	Superior	46.699	-92.195
535	Bulrush	Squaw Bay Wetland #1	Huron	44.996	-83.462
5952	Bulrush	Stokes Bay Wetland 1	Huron	44.992	-81.393
1303	Bulrush	Stony Creek Wetland	Michigan	43.571	-86.455
5963	Bulrush	Sturgeon Bay 1	Superior	48.207	-89.297
6050	Bulrush	West Shore of St. Joseph Island 1	Huron	46.178	-84.053
6051	Bulrush	West Shore of St. Joseph Island 2	Huron	46.167	-84.024
461	Bulrush	Wildfowl Bay Wetland	Huron	43.872	-83.344
5013	Lily	Anderson Creek	Huron	46.331	-83.977
1866	Lily	Bay View Wetland	Erie	41.459	-82.808
1070	Lily	Bibon Lake-Flag River Wetland	Superior	46.784	-91.387
5098	Lily	Black Creek Wetland	Ontario	43.946	-77.063
7052	Lily	Braddock Bay	Ontario	43.308	-77.72
1152	Lily	Dead River Wetland	Superior	46.579	-87.402
7027	Lily	East Sodus	Ontario	43.263	-76.94
7024	Lily	Floodwood Pond	Ontario	43.727	-76.194
637	Lily	Flowers Creek Wetland	Huron	45.995	-84.319
1325	Lily	Galien River Wetland	Michigan	41.805	-86.73
1896	Lily	Halfway Creek Wetland	Erie	41.744	-83.472
5407	Lily	Hay Bay Marsh 8	Ontario	44.155	-76.909
1863	Lily	Hemming Ditch Wetland	Erie	41.435	-82.655
1438	Lily	Henderson Point Wetland	Michigan	44.848	-87.556
7053	Lily	Irondequoit Bay Wetland	Ontario	43.167	-77.528
1584	Lily	Kenyon Bay Wetland	Michigan	46.054	-85.198
999	Lily	Lac LaBelle Wetland	Superior	47.378	-87.978
7020	Lily	Lakeview Pond-Sandy Creek-Colwell Ponds Marsh	Ontario	43.75	-76.204

1282	Lily	Little Manistee River Wetland	Michigan	44.208	-86.267
123	Lily	Little Sandy Creek Marsh	Ontario	43.637	-76.163
5573	Lily	Lynde Creek Marsh	Ontario	43.856	-78.962
5601	Lily	Maskinonge Bay 2	Huron	46.342	-84.086
1847	Lily	Mentor Marsh	Erie	41.734	-81.31
5634	Lily	Mill Creek Wetland	Erie	42.31	-81.911
1928	Lily	Monroe City Area Wetland	Erie	41.9	-83.363
7062	Lily	Monroe Dikes A	Erie	41.907	-83.361
523	Lily	Nayanguing Point Wildlife Area Wetland #3	Huron	43.861	-83.922
1933	Lily	North Maumee Bay Area Wetland	Erie	41.761	-83.456
122	Lily	North Pond Area Wetland	Ontario	43.656	-76.183
1849	Lily	Old Woman Creek Wetland	Erie	41.375	-82.512
1888	Lily	Ottawa National Wildlife Refuge Wetland	Erie	41.624	-83.213
1904	Lily	Otter Creek Wetland	Erie	41.847	-83.417
5718	Lily	Parrott Bay Wetland 2	Ontario	44.221	-76.691
1859	Lily	Plum Brook Area Wetland #2	Erie	41.428	-82.629
1870	Lily	Port Clinton Wetland	Erie	41.492	-82.951
7050	Lily	Radio Tower Bay	Superior	46.654	-92.214
116	Lily	Ramona Beach Marsh	Ontario	43.532	-76.222
5818	Lily	Roberts Island Wetland	Huron	44.858	-79.835
28	Lily	Salmon Creek	Ontario	43.31	-77.741
804	Lily	Sand Island Wetland	Huron	46.315	-84.203
5873	Lily	Sawguin Creek Marsh 5	Ontario	44.143	-77.322
1703	Lily	Seagull Bar Area Wetland	Michigan	45.078	-87.585
119	Lily	South Pond Wetland #1	Ontario	43.62	-76.187
780	Lily	South Scott Bay Area Wetland	Huron	46.048	-83.687
5988	Lily	Tobies Bay Wetland	Huron	44.847	-79.788
6025	Lily	Waupoos Creek Swamp 1	Ontario	43.983	-77.026
6053	Lily	Westside Beach Marsh	Ontario	43.888	-78.681
1898	Lily	Woodtick Penninsula Wetland	Erie	41.768	-83.44
7048	SAV	40th Ave West	Superior	46.74	-92.15
5008	SAV	Amherst Bar Wetland 1	Ontario	44.186	-76.623
130	SAV	Black Pond-Little Stony Creek Marsh	Ontario	43.796	-76.221
5103	SAV	Blessington Creek Marsh 1	Ontario	44.171	-77.317
7052	SAV	Braddock Bay	Ontario	43.308	-77.72
51	SAV	Buck Pond	Ontario	43.28	-77.674
1830	SAV	Buckthorn Island Wetland	Ontario	43.061	-78.988
7026	SAV	Buttonwood Creek	Ontario	43.298	-77.73
5151	SAV	Carnachan Bay Wetland 2	Ontario	44.076	-77.027
1475	SAV	Cedar River Wetland #1	Michigan	45.409	-87.352

167	SAV	Chaumont River Mouth Wetland	Ontario	44.067	-76.151
1089	SAV	Clough Island Wetland #1	Superior	46.71	-92.187
5187	SAV	Collingwood Harbour Marsh 5	Huron	44.505	-80.229
5235	SAV	Detroit River Marshes	Erie	42.203	-83.1
66	SAV	East Bay Wetland	Ontario	43.277	-76.902
23	SAV	East Creek Wetland	Ontario	43.338	-77.796
7027	SAV	East Sodus	Ontario	43.263	-76.94
1039	SAV	Fish Creek Wetland #1	Superior	46.583	-90.945
7024	SAV	Floodwood Pond	Ontario	43.727	-76.194
187	SAV	Fox Creek Marsh	Ontario	44.059	-76.296
8	SAV	Golden Hill State Park Wetland	Ontario	43.37	-78.478
5374	SAV	Greater Cataraqui Marsh	Ontario	44.266	-76.466
5401	SAV	Hay Bay Marsh 2	Ontario	44.167	-76.953
10	SAV	Johnson Creek Wetland	Ontario	43.366	-78.261
1651	SAV	Kalamazoo River Wetland	Michigan	42.64	-86.146
1437	SAV	Keyes Creek Wetland	Michigan	44.829	-87.57
112	SAV	Little Salmon River Marsh	Ontario	43.521	-76.253
1063	SAV	Little Sand Bay Wetland	Superior	46.948	-90.883
5541	SAV	Long Point Wetland 3	Erie	42.579	-80.296
1457	SAV	Long Tail Point Wetland #2	Michigan	44.593	-87.984
5573	SAV	Lynde Creek Marsh	Ontario	43.856	-78.962
62	SAV	Maxwell Bay Wetland	Ontario	43.269	-77.026
5634	SAV	Mill Creek Wetland	Erie	42.31	-81.911
199	SAV	Mud Bay Marsh #2	Ontario	44.084	-76.306
122	SAV	North Pond Area Wetland	Ontario	43.656	-76.183
7033	SAV	Oconto Marsh #2	Michigan	44.968	-87.801
989	SAV	Oskar Area Wetland	Superior	47.184	-88.639
163	SAV	Perch River Wetland	Ontario	43.998	-76.077
5735	SAV	Pine Point Wetland 1	Ontario	44.098	-77.501
5736	SAV	Pine Point Wetland 2	Ontario	44.105	-77.493
1096	SAV	Pokegama River Wetland	Superior	46.676	-92.144
5785	SAV	Presquille Bay Marsh 7	Ontario	44.03	-77.71
116	SAV	Ramona Beach Marsh	Ontario	43.532	-76.222
1494	SAV	Rapid River Wetland	Michigan	45.918	-86.959
5849	SAV	Sadler Creek Wetland 6	Huron	45.049	-81.461
5849	SAV	Sadler Creek Wetland 6	Huron	45.049	-81.461
5855	SAV	Sand Bay 1	Ontario	44.15	-76.503
5869	SAV	Sawguin Creek Marsh 10	Ontario	44.078	-77.309
1041	SAV	Sioux River Wetland	Superior	46.733	-90.882
119	SAV	South Pond Wetland #1	Ontario	43.62	-76.187
7051	SAV	South Pond Wetland 2	Ontario	43.58	-76.192
1522	SAV	South River Bay Wetland	Michigan	45.745	-86.626
5933	SAV	Southwest Sturgeon Bay 2	Huron	45.616	-80.459

5950	SAV	Stobie Creek 1	Huron	46.331	-83.885
1523	SAV	Sucker Lake Wetland	Michigan	45.67	-86.597
1090	SAV	Tallas Island Wetland	Superior	46.717	-92.192
1941	SAV	Thompson Bay Area Wetland	Erie	42.168	-80.081
5990	SAV	Toronto Island Wetlands 2	Ontario	43.618	-79.381
6055	SAV	Wheatley West Two Creeks	Erie	42.084	-82.46
6073	SAV	Wilmot Rivermouth Wetland	Ontario	43.901	-78.598
630	Typha	Ailes Point Area Wetland #2	Huron	45.993	-84.365
1866	Typha	Bay View Wetland	Erie	41.459	-82.808
1070	Typha	Bibon Lake-Flag River Wetland	Superior	46.784	-91.387
1464	Typha	Charles Pond Wetland	Michigan	44.764	-87.939
1089	Typha	Clough Island Wetland #1	Superior	46.71	-92.187
1201	Typha	Clough Island Wetland #3	Superior	46.701	-92.183
1458	Typha	Dead Horse Bay Wetland #9	Michigan	44.627	-88.013
1152	Typha	Dead River Wetland	Superior	46.579	-87.402
434	Typha	Dickenson Island Area Wetland	Erie	42.607	-82.651
515	Typha	East Saginaw Bay Coastal Wetland #5	Huron	43.673	-83.575
1489	Typha	Escanaba River Wetland	Michigan	45.786	-87.066
5409	Typha	Hay Bay Wetland 1	Huron	45.234	-81.704
5422	Typha	Hillman Marsh	Erie	42.042	-82.5
5509	Typha	Lake George 1	Huron	46.408	-84.111
1698	Typha	Little Suamico River Area Wetland	Michigan	44.699	-87.993
1457	Typha	Long Tail Point Wetland #2	Michigan	44.593	-87.984
616	Typha	Mackinac Creek Wetland	Huron	46.002	-84.41
1281	Typha	Manistee River Wetland	Michigan	44.265	-86.233
5654	Typha	Muddy Creek	Erie	42.071	-82.472
792	Typha	Munuscong Lake Wetland #2,#3 Munuscong River Delta	Huron	46.216	-84.257
523	Typha	Nayanguing Point Wildlife Area Wetland #3	Huron	43.861	-83.922
496	Typha	Nayanguing Point Wildlife Area Wetland #5	Huron	43.922	-83.902
1904	Typha	Otter Creek Wetland	Erie	41.847	-83.417
917	Typha	Palmers Point Wetland	Huron	46.532	-84.199
777	Typha	Paw Point-North Scott Bay Wetland #1	Huron	46.071	-83.665
1465	Typha	Pensaukee River Wetland	Michigan	44.816	-87.91
988	Typha	Pilgrim River Wetland	Superior	47.105	-88.514
1862	Typha	Plum Brook Area Wetland #3	Erie	41.427	-82.639
5782	Typha	Presquille Bay Marsh 4	Ontario	44	-77.721
780	Typha	South Scott Bay Area Wetland	Huron	46.048	-83.687
1697	Typha	Suamico River Area Wetland	Michigan	44.643	-88.012
1090	Typha	Tallas Island Wetland	Superior	46.717	-92.192

5988	Typha	Tobies Bay Wetland	Huron	44.847	-79.788
1497	Typha	Whitefish River Wetland #3	Michigan	45.916	-86.945
1898	Typha	Woodtick Penninsula Wetland	Erie	41.768	-83.44

Appendix 2. Site descriptions for GLEI2 Database

Site #	Site Name	Lake	Region	Latitude	Longitude
1041	Sioux River Wetland	Superior	N	46.733	-90.882
1077	Allouez Bay Wetland	Superior	N	46.681	-91.982
1444	Atkinson Marsh	Michigan	N	44.558	-88.039
1458	Dead Horse Bay Wetland #9	Michigan	N	44.627	-88.013
1465	Pensaukee River Wetland	Michigan	N	44.816	-87.91
1469	Peshtigo River Wetland #1	Michigan	N	44.995	-87.672
1497	Whitefish River Wetland #3	Michigan	N	45.916	-86.945
1514	Ogontz Bay Wetland #3	Michigan	N	45.866	-86.765
1698	Little Suamico River Area Wetland	Michigan	N	44.699	-87.992
1703	Seagull Bar Area Wetland	Michigan	N	45.078	-87.585
1859	Plum Brook Area Wetland #2	Erie	S	41.428	-82.629
1862	Plum Brook Area Wetland #3	Erie	S	41.427	-82.639
1863	Hemming Ditch Wetland	Erie	S	41.435	-82.655
1866	Bay View Wetland	Erie	S	41.459	-82.808
1888	Ottawa Nat'l Wildlife Refuge Wetland	Erie	S	41.624	-83.213
1928	Monroe City Area Wetland	Erie	S	41.9	-83.363
5512	Lake St. Clair Marshes	Erie	S	42.418	-82.417
5541	Long Point Wetland 3	Erie	S	42.579	-80.296
5634	Mill Creek Wetland	Erie	S	42.31	-81.911
5729	Pine Bay 1	Superior	N	48.033	-89.523
5933	Southwest Sturgeon Bay 2	Huron	N	45.616	-80.459
5950	Stobie Creek 1	Huron	N	46.331	-83.885
11362	Unnamed US High-energy	Superior	N	47.873	-89.856
11365	Unnamed US High-energy	Superior	N	47.455	-91.032
11366	Unnamed US High-energy	Superior	N	47.411	-91.099
11367	Unnamed US High-energy	Superior	N	47.004	-91.689
11368	Unnamed US High-energy	Superior	N	46.789	-92.083
11371	Unnamed US High-energy	Superior	N	46.747	-91.638
11372	Unnamed US High-energy	Superior	N	46.833	-91.29
11383	Unnamed US Embayment	Superior	N	46.434	-86.634
11385	Unnamed US Embayment	Huron	N	45.998	-84.403
11389	Unnamed US High-energy	Michigan	N	45.427	-87.32
11393	Unnamed US High-energy	Michigan	N	44.813	-87.649
11397	Unnamed US High-energy	Michigan	S	42.796	-87.768
11400	Unnamed US High-energy	Michigan	S	41.825	-86.728
11418	Unnamed US Embayment	Erie	S	41.751	-83.457
11421	Unnamed US High-energy	Erie	S	42.101	-80.199
11422	Unnamed US Embayment	Erie	S	42.139	-80.115
11423	Unnamed US High-energy	Ontario	S	43.83	-76.277
11425	Unnamed US Embayment	Ontario	N	43.968	-76.114

20032	Unnamed Canadian High-energy	Huron	N	45.004	-81.229
20034	Unnamed Canadian High-energy	Huron	N	45.193	-81.322
20037	Unnamed Canadian High-energy	Huron	S	43.722	-81.724
20102	Unnamed Canadian embayment	Superior	N	48.018	-89.54
20103	Unnamed Canadian embayment	Superior	N	48.036	-89.502
20108	Unnamed Canadian embayment	Superior	N	48.19	-89.296
20130	Unnamed Canadian embayment	Huron	N	46.293	-83.786
20143	Unnamed Canadian embayment	Huron	N	45.984	-81.536
20157	Unnamed Canadian embayment	Huron	N	44.809	-79.919
20158	Unnamed Canadian embayment	Huron	N	44.81	-80.062
20160	Unnamed Canadian embayment	Huron	N	44.789	-81.081
20168	Unnamed Canadian embayment	Huron	N	44.866	-81.331
20169	Unnamed Canadian embayment	Erie	S	42.874	-79.259
20173	Unnamed Canadian embayment	Ontario	N	43.992	-77.011
20176	Unnamed Canadian embayment	Ontario	S	44.073	-77.57
20278	Unnamed Canadian High-energy	Erie	S	42.033	-82.626
20279	Unnamed Canadian High-energy	Erie	S	41.994	-82.853
20342	Unnamed Canadian High-energy	Huron	N	44.634	-81.272
20371	Unnamed Canadian High-energy	Huron	S	43.012	-82.384
20408	Unnamed Canadian High-energy	Huron	S	43.094	-82.137

Appendix 3. Final metrics and scoring system for Cooper-IBIs. From Cooper *et al.* (in review)

	Scoring[†]		
Bulrush (*Schoenoplectus* spp.)	0	1	2
Pielou's Evenness	0-0.4	>0.4-0.8	>0.8
Non-native species richness	≥2	1	0
Notropis species richness	0	>0-2	>2
Native Cyprinidae CPUE	0	>0-50	>50
Rock bass CPUE	0	>0-4	>4
White sucker CPUE	0	>0-5	>5
Smallmouth bass CPUE	<2	>2-5	>5
% Black+brown bullhead	0	>0-25	>25
Johnny darter CPUE	0	>0-0.33	>0.34
Common carp CPUE	>2	>0-2	0
% Carnivore (invertivore+piscivore+zooplanktivore)	>90	40-90	<40
% Richness of high and extra-high temperature spawners	100	>82-100	0-82
% Richness short-lived species	<20	20-60	>60
% Richness species particularly sensitive to degradation	0	>0-15	>15
Final score for zone = (sum of metrics / 28) * 100			
Cattail (*Typha* spp.)			
Pielou's Evenness	<0.50	0.5-0.75	>0.75
% Richness native species*	<60	60-<100	100
Non-native species richness*	>2	1-2	0
Notropis species richness	0	1	≥2
% Native Cyprinidae	0-20	>20-50	>50
Rock bass CPUE	0	>0-3	>3
% Black+brown bullhead*	0	>0-25	>25
% Richness benthic habitat species	>75	40-75	<30
% Vegetation habitat species	<20	20-60	>60
% Richness nest spawners	0	>0-70	>70
% Richness of high and extra-high temperature spawners	100	60-<100	<60
% Richness large and extra-large species	>40	20-40	<20
% Richness species particularly sensitive to degradation	0	>0-8	>8
Final score for zone = (sum of metrics / 32) * 100			

	Scoring†		
Water lily (*Brassenia* spp., *Nuphar* spp., *Nymphaea* spp.)	0	1	2
Pielou's Evenness	<0.5	0.5-0.75	>0.75
Non-native species richness	>2	>0-2	0
Rock bass CPUE	<2	2-6	>6
Smallmouth bass CPUE	0	>0-3	>3
% Black+brown bullhead	<5	5-30	>30
Round goby CPUE	>4	>0-4	0
Yellow perch CPUE	0	>0-10	>10
% Common carp	>3	>0-3	0
% Richness carnivore species (invertivore+piscivore+zooplanktivore)	<50	50-75	>75
% Richness vegetation spawners	<15	15-40	>40
% Richness species particularly sensitive to degradation*	0	>0-10	>10
Final score for zone = (sum of metrics / 24) * 100			
Submersed aquatic vegetation			
Pielou's Evenness	<0.5	0.5-0.75	>0.75
Non-native species richness	≥2	1	0
% Richness Centrarchidae species	<25	25-50	>50
Cyprinidae species richness	>3	2-3	0-1
Bluntnose minnow CPUE	>5	>0-5	0
Common carp CPUE	>1	>0-1	0
% Richness carnivore species (invertivore+piscivore+zooplanktivore)	<50	50-75	>75
% Benthic invertivores	<10	10-30	>30
% Vegetation spawners	<5	5-30	>30
% Richness species particularly sensitive to degradation	0	>0-15	>15
Final score for zone = (sum of metrics / 20) * 100			

Note: Catch per unit effort (CPUE) was catch net^{-1} night^{-1}. †Scores for cattail and lily metrics identified with an "" should be doubled (e.g., 0, 2, 4)*

Appendix 4. Species specific values for calculating the Wetland Fish Index. From Seilheimer & Chow-Fraser (2006).

		WFI-PA		WFI-AB	
Scientific Name	Common Name	T	U	T	U
Alosa pseudoharengus	Alewife	2	2	2	1
Ambloplites rupestris	Northern Rock Bass	1	4	2	4
Ameiurus melas	Black Bullhead	2	3	2	3
Ameiurus nebulosus	Brown Bullhead	1	3	1	2
Amia calva	Bowfin	2	4	2	4
Aplodinotus grunniens	Freshwater Drum	2	1	2	1
Carassius auratus	Goldfish	2	1	2	1
Catostomus catostomus	Longnose Sucker	3	5	3	5
Catostomus commersonii	White Sucker	1	3	2	3
Cottus bairdii	Mottled Sculpin	3	4	3	4
Cottus cognatus	Eastern Slimy Sculpin	2	4	2	4
Culaea inconstans	Brook Stickleback	2	4	2	4
Cyprinella spiloptera	Spotfin Shiner	1	2	1	1
Cyprinus carpio	Common Carp	1	2	1	1
Dorosoma cepedianum	Gizzard Shad	2	1	2	1
Esox americanus	Grass Pickerel	3	4	3	4
Esox lucius	Northern Pike	2	4	2	4
Esox masquinongy	Great Lakes Muskellunge	3	4	3	4
Etheostoma exile	Iowa Darter	3	5	3	4
Etheostoma microperca	Least Darter	3	4	3	5
Etheostoma nigrum	Johnny Darter	2	3	2	3
Fundulus diaphanus	Banded Killifish	3	4	3	4
Gasterosteus aculeatus	Threespine Stickleback	2	2	1	2
Hybognathus hankinsoni	Brassy Minnow	2	1	2	1
Ictalurus punctatus	Channel Catfish	2	1	2	1
Labidesthes sicculus	Northern Brook Silverside	2	4	2	4
Lepisosteus osseus	Longnose Gar	3	5	3	5
Lepomis cyanellus	Green Sunfish	1	1	1	1
Lepomis gibbosus	Pumpkinseed Sunfish	2	3	2	3
Lepomis macrochirus	Bluegill Sunfish	1	3	1	3
Lepomis megalotis	Longear Sunfish	3	4	3	4
Luxilus cornutus	Common Shiner	3	4	3	4
Margariscus margarita	Pearl Dace	3	4	3	4
Micropterus dolomieu	Smallmouth Bass	2	4	2	4
Micropterus salmoides	Largemouth Bass	2	3	2	3
Morone americana	White Perch	1	1	2	1
Morone chrysops	White Bass	1	1	1	1
Moxostoma anisurum	Silver Redhorse	3	5	3	5

Moxostoma breviceps	Smallmouth Redhorse	2	4	2	4
Notemigonus crysoleucas	Golden Shiner	2	3	2	3
Notropis atherinoides	Emerald Shiner	2	3	2	3
Notropis heterodon	Blackchin Shiner	3	5	3	5
Notropis heterolepis	Blacknose Shiner	2	4	2	4
Notropis hudsonius	Spottail Shiner	1	2	1	2
Notropis stramineus	Sand Shiner	1	3	1	3
Notropis volucellus	Northern Mimic Shiner	3	5	3	5
Noturus gyrinus	Tadpole Madtom	2	4	2	4
Osmerus mordax	Rainbow Smelt	3	4	3	4
Perca flavescens	Yellow Perch	2	3	2	3
Percina caprodes	Logperch	2	3	2	4
Percopsis omiscomaycus	Troutperch	3	4	2	4
Phoxinus eos	Northern Redbelly Dace	3	5	3	5
Pimephales notatus	Bluntnose Minnow	1	3	2	4
Pimephales promelas	Fathead Minnow	1	2	1	2
Pomoxis annularis	White Crappie	1	1	1	1
Pomoxis nigromaculatus	Black Crappie	2	3	2	3
Prosopium cylindraceum	Round Whitefish	3	4	3	4
Pungitius pungitius	Ninespine Stickleback	3	4	3	4
Sander vitreus	Walleye	3	4	3	4
Semotilus atromaculatus	Creek Chub	1	3	1	3
Umbra limi	Central Mudminnow	2	4	2	4

VITA AUCTORIS

NAME:	Jeffrey Buckley
PLACE OF BIRTH:	Oshawa, ON
YEAR OF BIRTH:	1985
EDUCATION:	Queen’s University, B.Sc. (Hons.) - Biology, Kingston, ON, 2007
	University of Windsor, M.Sc. – Biological Sciences, Windsor, ON, 2015