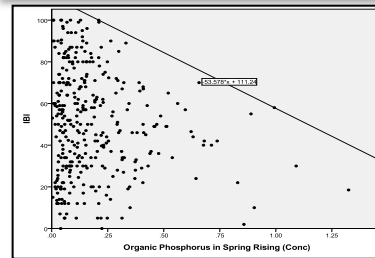
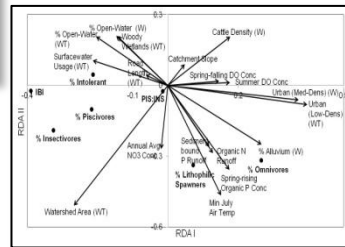
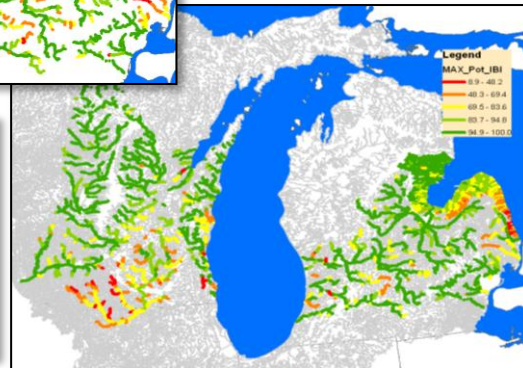
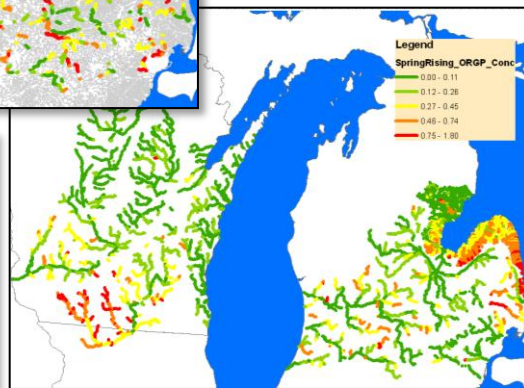
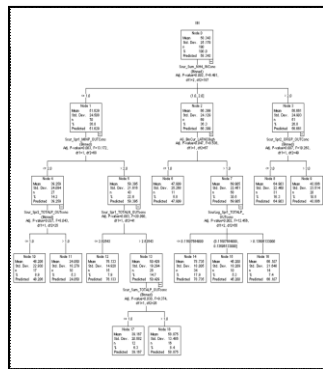
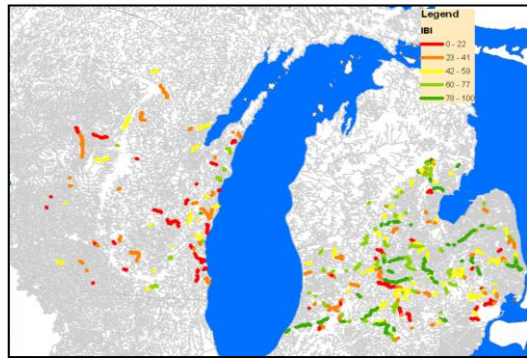




Assessing Benefits of Conservation Practices to the Biological Integrity of Agricultural Streams in MI and WI

2011 FINAL REPORT



United States Department of Agriculture
 Natural Resource Conservation Service
 Conservation Effects Assessment Project:
*A Multi-agency Effort to Quantify the Environmental Effects
 of Conservation Practices and Programs*

Assessing Benefits of Conservation Practices to the Biological Integrity of Agricultural Streams in MI and WI

FINAL REPORT

30 September 2011

Scott P. Sowa, PhD, Principal Investigator
Matthew Herbert, Data Analysis and Modeling
Layla Cole, Data Management, Analysis and Modeling
Sagar Mysorekar, GIS Mapping and Modeling
John Legge, Study Design and Project Management
Tia Bowe, GIS Mapping and Modeling
Amirpouyan Nejadhashemi, PhD, SWAT Modeling
Matt Einheuser, SWAT Modeling and Analysis
Lizhu Wang, PhD, Fish Community Data and Cooperator

The Nature Conservancy
Michigan Chapter
Lansing, MI 48906

Submitted by:
Scott P. Sowa

Research Performed Under:
Cooperative Agreement No. 8-3A75-8-66 Mod 5

Suggested Citation:

Sowa, S. P¹., M. Herbert¹, L. Cole¹, S. Mysorekar¹, J. Legge¹, T. Bowe¹, A. Nejadhashemi², M. Einheuser², and L. Wang³. 2011. Assessing benefits of conservation practices to the biological integrity of agricultural streams in MI and WI. Final Report submitted to NRCS Conservation Effects Assessment Project. 56 pp.

¹ *The Nature Conservancy*

² *Michigan State University, Department of Biosystems and Agricultural Engineering*

³ *Michigan Department of Natural Resources*

TABLE OF CONTENTS

List of Tables	iii
List of Figures	iv
Executive Summary	v
Acknowledgements	vii
Introduction	1
Study Area	4
Objective 1	10
Objective 1: Methods.....	10
Objective 1 Results and Discussion.....	13
Objective 2	18
Objective 2: Methods.....	18
Objective 2: Results and Discussion.....	28
Overall Discussion	60
Literature Cited	64

LIST OF TABLES

Table 1. Statistical analysis based on daily streamflow SWAT model outputs.....	14
Table 2. Conservation practices for Phase II.....	18
Table 3. Variance between IBI metrics explained.....	32
Table 4. Threshold values for IBI.....	51
Table 5. Threshold values for percent intolerant species	51
Table 6. Frequencies for limiting variables.....	55
Table 7. Percentages for limiting variables.....	56

LIST OF FIGURES

Figure 1. Study area for the Great Lakes CEAP Project.....	5
Figure 2. Maps showing current and pre-settlement land use.....	11
Figure 3. Map showing precipitation and temperature monitoring stations.....	12
Figure 4. Map showing USGS gauging stations.....	12
Figure 5. Maps showing modeled percent changes in several landscape factors.....	15
Figure 6. Graph showing classes of percent change in project area.....	16
Figure 7. Average daily discharge values for seven streams in project area.....	21
Figure 8. Maps showing predicted mineral phosphorous concentrations.....	21
Figure 9. Map showing the location and IBI scores for fish samples.....	23
Figure 10. Box plot showing difference in IBI between streams and rivers.....	27
Figure 11. Redundancy Analysis plot; natural variables and IBI metric scores.....	29
Figure 12. Redundancy Analysis plot; threat variables and IBI metric scores.....	30
Figure 13. Redundancy Analysis plot; water quality variables and IBI metric scores.....	31
Figure 14. Redundancy Analysis plot; natural, threat and water quality variables and IBI metric scores.....	32
Figure 15. CART model for predicting IBI in streams using SWAT data.....	34
Figure 16. CART model for predicting IBI in rivers using SWAT data.....	35
Figure 17. CART model for predicting IBI with natural landscape variables.....	36
Figure 18. CART model for predicting IBI with human induced threats.....	37
Figure 19. Wedges of IBI to watershed scale natural variables.....	39
Figure 20. Wedges of IBI to watershed scale threat variables.....	40
Figure 21. Wedges of percent intolerant to watershed scale threat variables.....	41
Figure 22. Maps of maximum potential IBI based watershed scale natural variables.....	44
Figure 23. Maps of maximum potential IBI based on watershed scale threats.....	45
Figure 24. Maps of maximum potential percent intolerant based on watershed scale groundwater index and percent impervious.....	46
Figure 25. Wedges showing upper limit of IBI based on SWAT variables.....	49
Figure 26. Wedges showing upper limit of percent intolerant based on SWAT variables.....	50
Figure 27. Maps showing which type o variable are currently most limiting either IBI or percent intolerant metrics.....	52
Figure 28. Maps showing the specific limiting variables for IBI.....	53
Figure 29. Maps showing the specific limiting variables for percent intolerant species.....	54
Figure 30. Map showing stream reaches that are limited by target disturbances.....	57
Figure 31. Maps showing the improvement capacity of each stream reach.....	58
Figure 32. Tri-plot of IBI to predicted historic organic phosphorus and percent change from historic to current levels.....	59

EXECUTIVE SUMMARY

In recent years there has been increased interest in a more thorough understanding and accounting of the benefits of conservation practices to fish and wildlife, particularly in response to the significant increase in funding for conservation programs that was authorized under the 2002 Farm Bill. In response the Conservation Effects Assessment Project (CEAP) was initiated by the NRCS, Agricultural Research Service (ARS), and Cooperative State Research, Education, and Extension Service (CSREES) to help better inform society of the likely benefits Farm Bill conservation program funding. The original goals of CEAP were to establish the scientific understanding of the effects of conservation practices at the watershed scale and to estimate conservation impacts and benefits for reporting at national and regional levels. Early CEAP investigations revealed that the cumulative benefits of NRCS conservation practices to aquatic communities is poorly understood and further scientific investigation is needed. The Great Lakes CEAP Project grew out of this realization and seeks to provide the science needed to assess and forecast the benefits of NRCS conservation practices to stream fish communities to help advance strategic conservation of freshwater biodiversity across the agricultural regions of the southern Great Lakes.

The overall goal of our project, which consists of two phases, is to provide decision makers with information to determine the limits of ecological improvement across the southern Great Lakes and models that use this information to establish realistic desired biological conditions. Phase 1 of our project, which is the focus of this report, is concentrating on using the predictive capabilities of SWAT to help generate the information needed for developing realistic biological expectations. Phase 1 consists of two primary objectives; 1) develop a fine-resolution SWAT model across the agricultural regions of the southern Great Lakes, and 2) develop models that predict fish community metrics based on SWAT output variables and other relevant watershed and local catchment variables.

Collectively the results our project successfully demonstrated the ability develop fine resolution SWAT model predictions across a large geographic area and to quantitatively link the resulting water quality and flow variables to fish community indicators to generate spatially explicit predictions. Our ability to, in essence, extend the predictive capabilities of SWAT to biological endpoints and also incorporate constraints not addressed by SWAT or NRCS conservation practices allowed to begin developing more realistic expectations to guide strategic conservation across the project area. This will help us to achieve our objectives in Phase 2 of the Great Lakes which is seeking to develop realistic goals (expectations) for fish community conditions in priority subwatersheds of the project area and working with partners to develop detailed strategies for achieving those goals.

Demonstrating the ability to predict fish community metrics from SWAT model outputs has the potential to significantly advance strategic conservation in the Great Lakes and beyond. Our results consistently demonstrated the importance of seasonal water quality and flow parameters, particularly the spring rising period, rather than average annual

conditions, which are more typically available and thus used by scientist to elucidate relations of these parameters to biological endpoints. The detailed and spatially comprehensive data provided by SWAT and the other predictors allowed us to assess and map likely fish community conditions and thresholds beyond sampled locations. Our models and maps exhibited extreme spatial heterogeneity in biological expectations under both current and historic conditions. This finding suggests that we should not hold all streams to the same standard even within a relatively small watershed or region, which is somewhat contrary to certain methods used to establish goals for fish community endpoints in streams.

Equally important to the temporal and spatial issues described above is the fact that the SWAT model also allows you to assess past, present, and potential future conditions based on different land use, land cover and management scenarios. The demand for demonstrating the benefits of conservation, particularly to biological endpoints, has increased sharply in recent years. Monitoring program and the associated retrospective analyses are useful for addressing this demand. However, we argue that equally important to these retrospective assessments are modeling efforts that forecast the likely benefits of conservation. The ability of SWAT to forecast future instream habitat and biological conditions based on different amounts and configurations of agricultural BMPs is very appealing for conservation planning. These management scenarios provide a means of developing management alternatives needed for developing truly realistic desired conditions by allowing decision makers to simultaneously evaluate ecological benefits relative to funding needs and constraints and potentially other socioeconomic costs in terms of agricultural production, farm income, and other valued services. As stated earlier, having the ability to extend such forecasts to biological endpoints, like fish communities, provides organizations like The Nature Conservancy the ability to identify where we can make meaningful improvements in freshwater biodiversity and help secure the necessary resources and attention needed to bring about those improvements.

Despite all of the realized and potential benefits of our project we must also be mindful of its limitations. We address these limitations by offering suggestions on how we might address them to significantly improve our ability to develop realistic biological expectations (goals) and forecast the likely benefits of future conservation scenarios to help develop effective strategies for achieving those goals.

ACKNOWLEDGEMENTS

We first want to thank to Charlie Rewa for his unwavering support of this project through the Wildlife Component of the NRCS CEAP. We also want to thank our key collaborators on this project who have contributed in many ways by providing critical input, review, and data needed to successfully complete this project. Specifically, we want to thank several individuals from Michigan State University, including Brad Wardynski from the Department of Biosystems and Agricultural Engineering , Jon Bartholic from the Institute of Water Research, Phanikumar Mantha and Chaopeng Shen from the Department of Civil and Environmental Engineering, and Dana Infante and Arthur Cooper from the Department of Fisheries and Wildlife. We would also like to thank Jana Stewart from the USGS Water Resource Division for providing critical datasets. Finally, we would like to thank Susan Wallace and George Wallace, of the USDA NRCS Resources Inventory and Assessment Division, for compiling the NRCS National Conservation Planning Database data sets used in this project. Additional financial support for TNC staff working on this project was provided by The Nature Conservancy's Great Lakes Fund for Partnership in Conservation Science and Economics.

INTRODUCTION

Agriculture, through its production of food, materials for clothing and shelter, and jobs, plays an important role in improving the quality of life for people across the United States, including those residing in the Great Lakes Region. In economic terms alone the benefits of agriculture to the Great Lakes Region are immense. The 2007 Census of Agriculture reported that there were nearly 126,000 farms in the region and that the value of agricultural sales was about \$14.5 billion with about half of this total generated from crop production and the other half from livestock production. About 67 percent of the farms in the Great Lakes Region primarily raise crops, about 26 percent are primarily livestock operations, and the remaining 7 percent produce a mix of livestock and crops. The five Great Lakes also moderate the climate of coastal areas, improving production and creating microclimates that are ideal for specialty crops such as cherries, asparagus and wine grapes. These high-value specialty crops also lead to spin-off industries such as culinary festivals and beverage production that provide social benefits and further increase economic outputs and jobs related to recreation and tourism. Unfortunately, the collective benefits of agriculture can sometimes have associated costs, particularly with regard to alteration of aquatic ecosystems, which also influence people's quality of life and also highly valued by society and organizations like The Nature Conservancy.

The effects of agriculture on aquatic ecosystems and freshwater biodiversity have been extensively studied and documented. Studies have consistently shown that various practices associated with row-crop agriculture and livestock production; including vegetative clearing, soil compaction, water withdrawal, channelization, and irrigation can significantly alter flow regimes, physical habitat, energy flow, water quality and the plant and animal biota (FISRWG 2001; Richter et al. 1997; Waters 1995). Major agricultural stressors include altered flow and thermal regimes and excess nutrients and sediments which affect 55% of the impaired waters in the United States (Allan 2004; Wells 1992). Collectively these changes in habitat lead to corresponding changes in the biotic communities and many recent studies have revealed connections between increased nutrients, sediments, and pesticides with changes in biological measures of algae, invertebrate, and fish communities (Frey et al. 2011; Hambrook-Berkman et al. 2010; Wang et al. 2007; Heiskary and Markus 2003; Cuffney et al. 2000; Rankin et al. 1999). Over the years farmers and state and federal governments have developed programs, policies, and funding mechanisms, like the Food Security Act of 1985 (aka the 1985 Farm Bill) to improve the sustainability and profitability of agriculture and to also reduce the impacts of agriculture on fish and wildlife habitat.

Passage of the 1985 Farm Bill authorized billions of dollars (US\$17 billion in 2002) for private land conservation (Gray and Teels 2006). Originally, the Farm Bill set out to reduce soil erosion from highly erodible sites and attempted to limit excess food production by idling marginal croplands (Heard et al. 2000). Since then, the Farm Bill

has evolved to administer, through the United States Department of Agriculture's Natural Resource Conservation Service (NRCS), additional programs (e.g., Wetlands Reserve Program and Environmental Quality Incentives Program) intended to improve wildlife habitat and environmental conditions in agricultural landscapes (Burger Jr. et al. 2006; Gray and Teels 2006; Heard et al. 2000). The majority of NRCS conservation practices do not directly target freshwater biodiversity conservation, but rather are intended to indirectly benefit biodiversity by improving water quality and hydrology. However, in recent years there has been increased interest in a more thorough understanding and accounting of the benefits of conservation practices to fish and wildlife, particularly in response to the significant increase in funding for conservation programs that was authorized under the 2002 Farm Bill. In response the Conservation Effects Assessment Project (CEAP) was initiated by the NRCS, Agricultural Research Service (ARS), and Cooperative State Research, Education, and Extension Service (CSREES) to help better inform society of the likely benefits Farm Bill conservation program funding (Mausbach and Dedrick 2004). The original goals of CEAP were to establish the scientific understanding of the effects of conservation practices at the watershed scale and to estimate conservation impacts and benefits for reporting at national and regional levels.

CEAP projects have mostly investigated the response of terrestrial ecosystems or species to a subset of NRCS practices (e.g., Burger Jr. et al. 2006a; Heard et al. 2000), or have targeted water quality issues by using hydrological models to assess sediment and contaminant loading in streams after conservation practice implementation (Westra et al. 2005). However, a pilot study concluded that NRCS conservation practices do have the potential to improve stream habitat conditions for a variety of aquatic species by targeting specific conservation practices to specific locations using modeled species distributions within a geographic information system (GIS) (Comer et al. 2007). The authors of this pilot study also noted that the specific or cumulative benefits of NRCS conservation practices to aquatic communities is poorly understood and further scientific investigation through a combination of a) localized, field based, watershed studies and b) geographically extensive, associative, modeling studies were needed. The Great Lakes CEAP Project grew out of this realization and seeks to provide the science needed to assess and **forecast** the benefits of NRCS conservation practices to stream fish communities to help advance **strategic conservation of freshwater biodiversity across the agricultural regions of the southern Great Lakes**.

Strategic conservation involves getting the right conservation practices to the right places in the right amount to achieve a set of realistic desired ecological and related socioeconomic conditions. There is an extensive body of science dedicated to help with identifying the right practices and places (watersheds and fields) for improving water quality conditions in agricultural landscapes (Richardson and Gatti 1999; Mishra and Singh 2007; Maringanti et al. 2009; Schilling and Wolter 2009). However, explicit, informed and realistic goals for how much conservation is needed have generally been lacking. This largely results from our inability to develop spatially-explicit linkages between biological endpoints, water quality and conservation practices.

As a result most goals focus on improvements in water quality, which are often expressed as nutrient or sediment reduction goals that are not informed by biological endpoints, or desired funding levels for specific practices or locations that are not informed by any ecological endpoints. Without these key linkages, we have lacked studies to evaluate the costs of restoration and whether our goals are even realistic. Such goals are a critical first step toward strategic conservation. However, for conservation organizations, like The Nature Conservancy with a mission to conserve biodiversity, it is difficult and often impossible to translate these goals into improvements to freshwater biodiversity.

Our project begins to develop the linkages between biological endpoints, water quality, and conservation practices, so that we can develop realistic desired conditions and begin to answer the question, “how much is enough.” These linkages will provide the foundation for making these decisions, but will necessarily need to be combined with socio-economic factors to determine whether biological endpoints are realistic. As a result, the primary goal of our project is to **provide decision makers with information to determine the limits of ecological improvement across the southern Great Lakes and models that use this information to establish realistic desired biological conditions.**

Specifically, our project seeks to help answer five key questions that guide strategic conservation, yet remain unanswered for much of the Great Lakes:

1. What are the realistic desired biological conditions for a given waterbody?
2. What are the current biological conditions?
3. Can we achieve the desired biological conditions given the existing suite of available conservation practices?, If yes, then;
4. How much of an investment will it take? And finally,
5. Which suite of conservation practices should we use and where should they be placed on the landscape in order to maximize the ecological return on our investments?

Answering these questions is fundamental to conservation efforts in agricultural landscapes. Yet, answering these questions is difficult because the return on investment differs among; a) the parameters of interest (e.g., physical, chemical, biological), conservation practices (e.g., grassed waterway vs. constructed wetland), and location (e.g., spatial variation in soil erosion potential). Fortunately, advancements in GIS technology and modeling have allowed for the development of various models and decision tools like the Soil and Water Assessment Tool (SWAT) that account for these and other interrelated factors and forecast the benefits of conservation actions on physical and chemical parameters. However, there has been very little effort to extend these modeling capabilities to biological endpoints and thus capitalize on the many benefits that model, like SWAT, offer to conservation planning by developing realistic expectations or goals and strategies for achieving those goals.

Phase 1 of the Great Lakes CEAP Project, which is the focus of this report, is concentrating on using the predictive capabilities of SWAT to help generate the information needed for developing realistic biological expectations. Phase 2 of our project is focused on using the information from Phase 1 and the scenario development and forecasting capabilities of SWAT, to develop realistic biological goals and also first cut strategies for achieving them. The specific objectives of Phase 1 of our project are:

Objective 1:

Develop a fine-resolution SWAT model across the agricultural regions of the southern Great Lakes to provide predicted values for water quality and flow variables that can be linked to existing biological sampling data of the region.

Objective 2:

Develop models that predict selected riverine biological endpoints based on SWAT output variables and other relevant watershed and local catchment variables.

STUDY AREA

The study area for this project focuses on the predominantly agricultural regions of southern Michigan and Wisconsin (Figure 1). Most of the study area falls within 4 level III ecoregions; 1) Driftless Area, 2) Southeastern Wisconsin Tills Plains, 3) Southern Michigan/Northern Indiana Drift Plains, and 4) Huron/Erie Lake Plains (USEPA 2003; Omernik 1987). For the sake of brevity these four ecoregions will be referred to as the Driftless Area, Till Plains, Drift Plains, and Lake Plains for the remainder of the report.

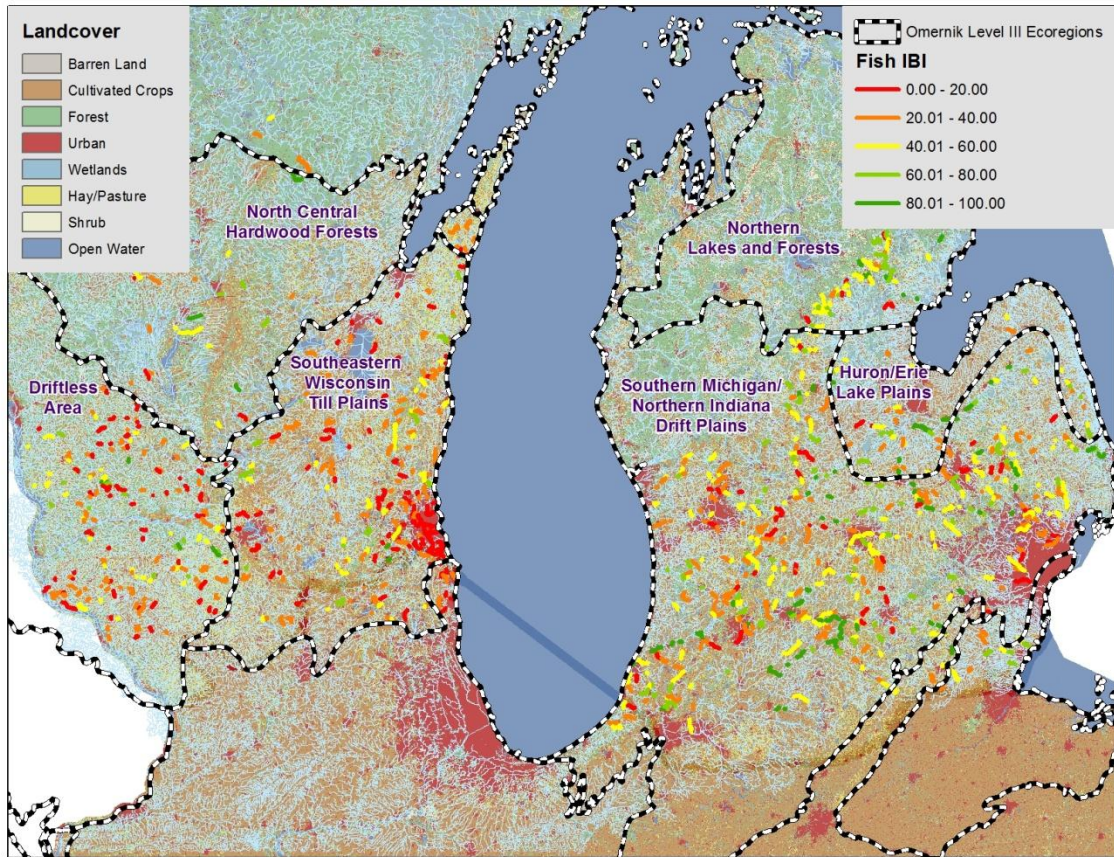


Figure 1. Study area for the Great Lakes CEAP Project showing the current land use, USEPA Level III ecoregions and the 1022 community fish samples with corresponding index of biotic integrity scores used for analysis and modeling.

Climate

The climate of the entire project area is typical of the upper Midwest with large annual and daily fluctuations. However, the climate of the Drift Plains and Lake Plains are much more strongly influenced by a Maritime Tropical air mass, with lake-effect snows and year-long moderation of temperatures from Lake Michigan and Lake Huron (Albert *et al.* 1986, Denton 1985, Eichenlaub 1979, Eichenlaub *et al.* 1990). The growing season is relatively similar across all four ecoregions, ranging from 142 to 184 days (Hole and Germain 1994). Compared to the Driftless Area and Till Plains, both the Drift and Lake Plains have more warm humid air masses from the Gulf of Mexico and fewer cold dry air masses of continental origin. Average annual precipitation is 32 to 34 inches, and average annual snowfall ranges from 36 inches in the south to approximately 44 inches in the north (Wendland *et al.* 1992).

Geology

The Drift Plain is underlain by Paleozoic bedrock deposited in marine and near-shore environments, including sandstone, shale, limestone, and dolomite (Dorr and Eschman 1984). This Paleozoic bedrock was deposited in an intercratonic basin, known as the Michigan basin, which was occupied by marine waters from the Silurian through

Pennsylvanian Periods. Mississippian and Devonian bedrocks are nearest the surface in the south and along the Great Lakes shorelines; Pennsylvanian bedrock is near the surface in the north (at the center of the Michigan basin). Bedrock exposures are few and small. At the eastern edge of the Drift Plain near Lake Erie, Devonian limestone bedrock is often within 5 feet of the surface and is locally exposed along streams. Local exposures of Mississippian shale, sandstone, and limestone occur within the Lake Plain ecoregion, closer to Saginaw Bay, but glacial lacustrine deposits can also be as deep as 300 feet on the inland portions of the lake plain (Albert 1994). Over the rest of the Drift Plain, 100 to 400 feet of loamy glacial drift cover the bedrock (Akers 1938), but very localized outcrops of Pennsylvanian sandstone do occur along the Grand River and its tributaries (Dorr and Eschman 1984).

Within the Till Plain ecoregion, the glacial drift covering the bedrock is generally less than 50 feet thick, except on the eastern edge where it can range from 100 to 200 feet thick (Trotta and Cotter 1973). The predominant bedrocks are Silurian dolomite to the east along Lake Michigan, and Ordovician dolomite in the central and western parts of the ecoregion (Ostrom 1981, Morey *et al.* 1982). Some limestone, sandstone, and shale are present in both of these bedrocks. Undifferentiated Devonian marine deposits are localized along the Lake Michigan shoreline. Cambrian sandstone, with some dolomite and shale, is along the far western edge of the subsection. Precambrian quartzite is localized in the west and Precambrian rhyolite, granite, and diorite are localized west of Lake Winnebago (Morey *et al.* 1982).

The geological history of the Driftless Area accounts for its distinctive physiographic features, including bedrock dominance. During the Paleozoic Period (ca. five hundred million years before present), layers of sediment and shells from marine organisms were deposited in seas, which covered the region. While retreating glaciers in adjacent regions buried topographical features in glacial drift, erosion in the unglaciated Paleozoic Plateau produced a dissected landscape with deep channels in a bedrock-dominated terrain. Stream erosion has dissected the landscape leaving more resistant rock types, such as sandstones and carbonates, in high cliffs and bluffs above the gentler slopes and waterways of the more erodible shales. The oldest layer exposed at the EFMO is the Jordan sandstone, which formed during the Cambrian period. This layer is seen along the base of the east facing bluffs and is an important aquifer for the area. Overlying the Jordan sandstone is the Prairie Du Chien formation of dolomite limestone (Mg Ca (CO₃)₂). This geologic stratum forms the bluffs in EFMO and the vicinity. The Mississippi River and its tributaries contain terraces and floodplain deposits, developed through a complex history of erosion and aggradation due to melt waters, scouring, and sediment deposition following the Wisconsin glaciation. Within the calcareous strata, weathering led to karst formations, including caves, sinkholes, springs, subsurface caverns, and underground and disappearing streams.

Soils

Most of the soils within the Drift Plains are calcareous and loamy, derived from underlying limestone, shale, and sandstone. Glacial till deposits are primarily loams, silt loams, and clay loams. Lacustrine soils are silt- and clay-rich; lacustrine sands are often banded with silt or clay. The outwash plains of the interlobate regions are largely comprised of sands, often containing abundant gravel. Most of the soils are classified as Alfisols, including Aqualfs and Udalfs, but there are also Aquepts, Aquolls, and Psamments (USDA Soil Conservation Service 1967).

A silt-loam cap of loess, about 2 feet thick, covers the soils of most of the Till Plains ecoregion, but there are also clay soils developed from glaciolacustrine deposits and sand soils developed from outwash deposits. Soils derived from the loess are silt loam at the surface, but subsoils are generally calcareous loam (till) or calcareous sand and gravel outwash (Hole and Germain 1994). The Driftless Area is covered with thin loess soils that create a well-drained landscape.

Landforms

Wisconsinan-age glacial and postglacial landforms cover the entire land surface of the project area. The glacial landforms include lake plains, outwash plains, ground moraines, and end moraines. The Lake Plains ecoregion is characterized by broad, flat, lacustrine plains that occur along all of the Great Lakes and extend more than 50 miles inland along the Lake Huron shoreline at Saginaw Bay within the Lake Plain ecoregion of our project area. Within the Drift Plains, sand dunes form a 1- to 5-mile band along much of the Lake Michigan shoreline. However, the interior of the Drift Plains consists of a relatively low plain of ground and end moraines, with narrow outwash channels throughout. The Driftless Area is the most dissected region in the project area, comprised by rolling hills and bluff outcroppings, exposed bedrock ridges, and deeply carved river valleys.

Potential Natural Vegetation

Most of the Drift and Lake Plains regions were forested (Albert 1994). Oak savanna was probably the most prevalent in the Drift Plains, followed by oak-hickory forest and beech-sugar maple forest. However, the Drift Plains is the only region of Michigan that originally supported large areas of tallgrass prairie, which was concentrated in the sandy interlobate area in the southwestern part of the state. The Lake Plains also contained large areas of wet prairie along the margins of Lake Erie, Lake St. Clair, and Lake Huron. Wetlands were also prevalent in both the Drift and Lake Plains and included extensive marshes, fens, and swamp forests (Comer *et al.* 1993a, 1993b).

Bur oak openings (savannas), oak forest, and tallgrass prairie were predominant in the western part of the Till Plains, but sugar maple-basswood forest was common to the east where there is greater fire protection because of dissected topography and numerous

kettle lakes of this region (Albert 1994). The prevailing directional trend of features, such as drumlin ridges and adjacent wetlands, helped determine the dominant vegetation within the Till Plains. On some southwest-northeast trending drumlin fields, tallgrass prairie and savanna were dominant; whereas north-south-trending drumlins served as fire barriers and allowed sugar maple-basswood forests to dominate.

Prior to European settlement the vegetation in the Driftless Area consisted of bluestem-dominated tallgrass prairies and oak savannas on ridgetops and dry upper slopes, and sugar maple (*Acer saccharum*), oaks (*Quercus* spp.), and basswood (*Tilia americana*) along cooler, moister slopes. Marsh and floodplain forests, as well as wet and mesic prairies were also common on river floodplains. Prairie occurred primarily on the broader ridge tops or steep slopes with south or southwest aspects.

Natural Disturbances

Fire was a key process for maintaining oak savannas and tallgrass prairies in all four regions. However, large windthrows were also frequently documented in the late 1800's in the Government Land Office (GLO) survey notes covering the Lake Plains region. This suggests that wind also likely played an important role as a natural disturbance serving to reset the succession cycle for natural vegetation and help maintain patches of early successional states.

Current Land use and Vegetation

Most of the project study area is farmed for row crops and collectively the Drift, Till, and Lake Plains regions represent the most heavily farmed sections in Michigan and Wisconsin. Almost all the original tallgrass and wet prairies have been converted to farmland (Albert 1994; USEPA 2003). The oak savannas have become forests as a result of fire suppression and some of the heaviest urban, industrial, and residential development in Michigan and Wisconsin has occurred in our project area, especially along the Great Lakes shorelines.

Not surprisingly, agriculture plays an important role in the economy of the region. The 2007 Census of Agriculture reported that there were nearly 126,000 farms in the Great Lakes Region and the value of the associated agriculture sales from these farms was about \$14.5 billion. About 67 percent of the farms primarily raise crops, about 26 percent are primarily livestock operations and the remaining 7 percent produce a mix of crops and livestock. More specifically, land use in Till Plains is mostly cropland, but the crops are largely forage and feed grains to support dairy operations, rather than corn and soybeans for cash crops (USEPA 2003). The Drift Plains is less agricultural than the flat agricultural Lake Plain to the east. Feed grain, soybean, and livestock farming as well as woodlots, quarries, recreational development, and urban-industrial areas are common in the Drift Plains. Today, most of the Lake Plains region has been cleared and artificially

drained and contains highly productive farms producing corn, soybeans, livestock, and vegetables; urban and industrial areas are also extensive.

Stream Habitat and Fish Communities

Stream habitat and quality have been moderately to severely altered across the project area due to a variety of human activities, including by channelization, ditching, tiling, and other agricultural activities. Altered hydrologic and thermal regimes, increased sediment and nutrient inputs, and loss of instream physical habitat are all primary concerns for streams in the project area. Specifically, land clearing, ditching, tiling, impoundments, and impervious surfaces have all collectively led to significant alteration of the hydrology of the project area. Many streams presently exhibit higher peak flows and lower base flows than they did prior to these activities. Fertilizer and manure applications along with point source discharges have led to increased nutrient concentrations and loads of many streams and receiving waters. Studies have shown clear relations between these habitat alterations and various biological measures of stream health, including fish communities (Rankin et al. 1999; Wang et al. 2007). Both Rankin et al. (1999) and Wang et al. (2007) found significant reductions in percent intolerant fishes and overall index of biotic integrity values with increased nutrient and sediment concentrations within streams of the southern Great Lakes.

Objective 1:

Develop a fine-resolution SWAT model across the agricultural regions of the southern Great Lakes to provide predicted values for water quality and flow variables that can be linked to existing biological sampling data of the region.

**Note: Objective 1 was carried out by a companion project that was jointly funded by TNC and NRCS CEAP (Coop Agreement: 68-7482-10-513). The principal investigator for this project was Dr. Amirpouyan Nejadhashemi, a faculty member within the Department of Biosystems and Agricultural Engineering at Michigan State University. A more detailed description of this work can be found in the following paper: Nejadhashemi, A., C. Shen, B. J. Wardynski, P. Mantha. 2010. Evaluating the Impacts of Land Use Changes on Hydrologic Responses in the Agricultural Regions of Michigan and Wisconsin. ASABE Paper 1008770, Pittsburgh, PA.*

OBJECTIVE 1 METHODS

Description of the Soil and Water Assessment Tool (SWAT) Model

SWAT is a physically based, computationally efficient, watershed scale, continuous-time model that operates on daily time step and was developed by Dr. Jeff Arnold at the United States Department of Agriculture (USDA) Agricultural Research Service (ARS). The model “*was developed to predict the impact of land management practices on water, sediment, and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over long periods of time*” (Neitsch et al., 2000). SWAT is mostly comprised of weather, hydrology, soil characteristics, plant growth, nutrients, pesticides, and land management components (Gassman et al. 2007). To allow for better estimate of impact of varying soil and land use types on hydrology, in SWAT, a watershed is divided into number of subwatersheds or subbasins. The subbasins are further divided into hydrologic response units (HRUs) based on similar land cover, soil, slope, and management combinations.

Hydrology components of SWAT include canopy storage, infiltration, redistribution, evapotranspiration, lateral subsurface flow, surface runoff, ponds, tributary channels, and return flow. A daily water budget in each HRU is calculated based on daily precipitation, runoff, evapotranspiration, percolation, and return flow from subsurface and groundwater flow (Nelson et al., 2006). In SWAT the surface runoff is calculated using either: The SCS curve number procedure ((USDA Soil Conservation Service, 1972)) or the Green & Ampt infiltration method ((Green and Ampt, 1911)1911). In addition, peak runoff rate is calculated with a modified rational method. SWAT estimates daily potential evapotranspiration using one of the three methods requiring varying inputs: Penman-Monteith, Hargreaves, or Priestly-Taylor. SWAT uses a kinematic storage model developed by Sloan et al. (1983) to estimate lateral flow. The groundwater system in SWAT consists of shallow and deep aquifers, which are calculated using empirical and analytical techniques (Neitsch et al., 2005). In SWAT, water is routed through the

channel network using the variable storage routing method (Williams, 1969) or the Muskingum River routing method (Chow et al., 1998).

SWAT Model Inputs

Data required for this study were acquired from various sources. For the current land use, 2001 National Land Cover Data (NLCD 2001) was used (Figure 2a). Pre-settlement land uses datasets (around early to mid 1800) were obtained from 1) Michigan Natural Features Inventory (MNFI, <http://web4.msue.msu.edu/mnfi/data/veg1800.cfm>). 2) Wisconsin Department of Natural Resources (http://dnr.wi.gov/maps/gis/documents/orig_vegetation_cover.pdf). 3) the Institute of Natural Resource Sustainability at the University of Illinois at Urbana-Champaign. These pre-settlement land cover maps were reclassified to the NLCD 2001 classes to provide consistency between land cover maps, which was then incorporated into the model for further analysis (Figure 2b).

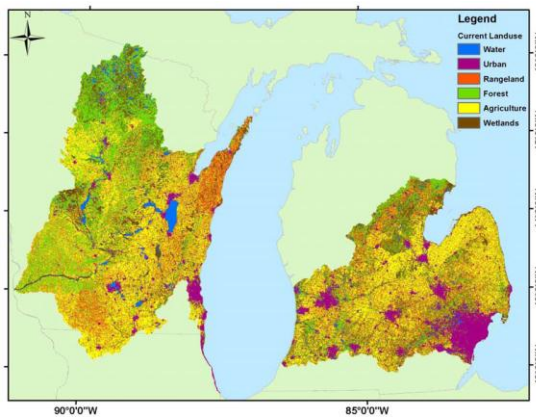


Figure 2a. Current land use map.

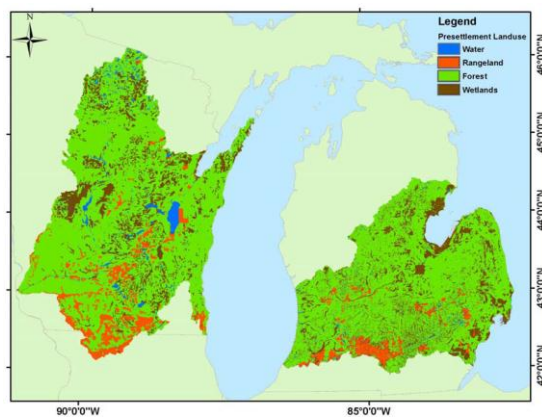


Figure 2b. Pre-settlement land use map.

The soil data was obtained from State Soil Geographic State Base (STATSGO) at the resolution of 1- by 2-degree topographic quadrangle units. USGS 1:250,000-scale Digital Elevation Model Grid (DEMG) at three arc second (100 m) resolution was obtained for the study area (<http://seamless.usgs.gov/>). National Hydrography Dataset (NHD; www.horizon-systems.com/nhdplus/) was used to improve hydrologic segmentation and subwatershed boundary delineation (Winchell et al., 2007). Daily precipitation records along with minimum and maximum temperature were acquired from 195 precipitation stations and 158 temperature stations within and around the study area (Figure 3) for 19 years (1990 - 2008). Eight different US Geological Survey (USGS) gauging stations were used for the SWAT model calibration and validation. At least nineteen years of daily stream flow records are available for each station (Figure 4).

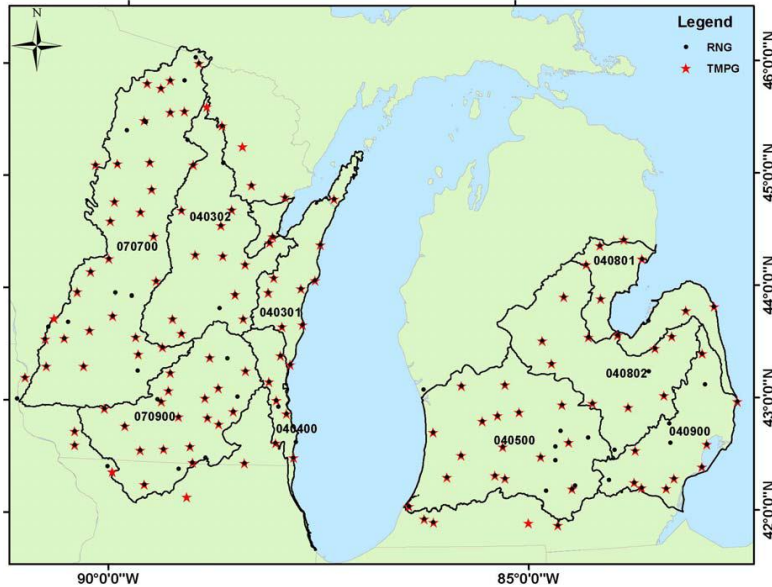


Figure 3. Precipitation (RNG) and temperature (TMPG) monitoring stations used for obtaining SWAT model input data.

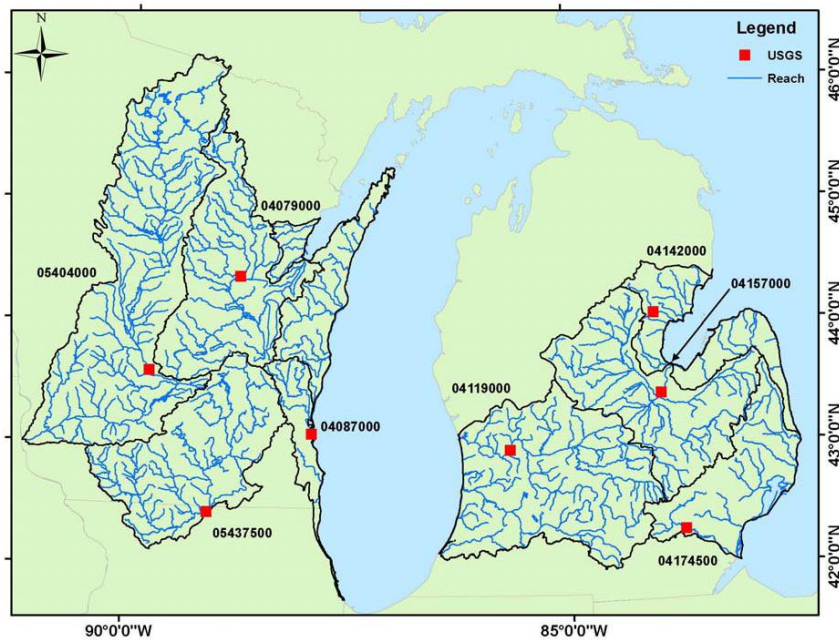


Figure 4. USGS gauging stations used for SWAT model calibration.

Sensitivity Analysis, Model Calibration and Validation Procedures

Sensitivity analysis is used to explain how the variation in the output of a model can be attributed to different sources of variation in the model input. The sensitivity analysis helps to determine parameters that controls watershed characteristics, understand behavior of the system being modeled, and to evaluate applicability of the model. Model calibration is an iterative process that compares simulated and observed data of interest through parameter evaluation. Validation extends calibration to ensure that the calibrated

model adequately represents variables and conditions affecting model results. The goal of validation is to conclude that the model is able to predict field observations for time periods separate from the calibration period (Donigan Jr. 2002).

In this study the sensitivity analysis concerning daily flow rate was performed on 42 different SWAT parameters on the nine HUC 8 digit watersheds for current and pre-settlement land uses. Eight USGS gauging stations with daily stream flow from 1990 – 2008 were used for the SWAT model calibration and validation. We plotted the average annual precipitation from 1990 to 2008 for the study area to identify the simulation period, for calibration and validation, in which a broad range of climatological conditions are represented (the figure is not shown here). We selected the period of 2002-2007 for the model calibration and validation because this period includes dry, wet, and normal climate conditions based on long term average precipitation records. Year 2002 was selected as the model warm-up year.

Due to lack of various long term weather data for mid-1800s the pre-settlement scenario was set up using current climatological data (1990 – 2008) to compare the results of land use changes in the region while eliminating the climatological difference. In addition, the same adjustments were made to the calibration parameters under pre-settlement scenario as they were under current land use scenario to minimize a possible bias caused by calibration process.

OBJECTIVE 1 RESULTS AND DISCUSSION

Sensitivity Analysis, Model Calibration and Validation Results

Among 42 parameters that were used for sensitivity analysis, 15 parameters were selected for further investigation. These parameters directly or indirectly influence the daily flow rate and overall ranked higher than others. Two criteria (mean and median) were selected to identify the most influential parameters, which affect daily flow rates. Among the study parameters, a significant shift in overall ranking was observed in *Cn2* (initial SCS curve number for moisture condition II), *Sol_Z* (depth from soil surface to bottom of layer), *Rchrg_Dp* (deep aquifer percolation fraction), and *Canmx* (maximum canopy storage).

To evaluate satisfactory model performances on daily basis we used following criteria: E_{NS} 0.20 and $R^2 > 0.4$ (Pouyan et al, 2010). Study results obtained from the SWAT model calibration, validation, and combined statistical analysis (Table 1) demonstrates that the model performance in all watersheds can be considered as satisfactory.

Table 1. Statistical analysis based on daily streamflow SWAT model outputs.

Watershed	Parameter	Uncalibrated Statistics	Calibration Statistics (2003-2005)	Validation Statistics (2006-2007)	Overall Statistics (2003-2007)
040302	NSE	-4.42	0.76	0.59	0.73
	RMSE	73.50	13.60	9.07	16.40
	R ²	0.016	0.80	0.73	0.75
040301 & 40400	NSE	-0.68	0.82	0.68	0.78
	RMSE	18.65	7.02	5.74	9.07
	R ²	0.20	0.82	0.71	0.78
070700	NSE	-1.01	0.40*	0.46**	0.45***
	RMSE	62.69	81.07*	96.87**	126.32***
	R ²	0.08	0.62*	0.56**	0.60***
070900	NSE	-8.76	0.74	0.70	0.74
	RMSE	285.70	35.64	30.57	46.95
	R ²	0.09	0.80	0.71	0.77
040801	NSE	-2.46	0.29	0.48	0.40
	RMSE	15.06	4.71	4.17	6.29
	R ²	0.17	0.47	0.55	0.50
040802	NSE	-1.38	0.77	0.83	0.80
	RMSE	206.70	48.31	35.68	60.06
	R ²	0.11	0.77	0.83	0.80
040900	NSE	-1.87	0.69	0.71	0.72
	RMSE	17.26	3.95	3.69	5.41
	R ²	0.20	0.74	0.76	0.77
040500	NSE	-2.68	0.80	0.84	0.80
	RMSE	167.56	31.62	20.46	37.7
	R ²	0.11	0.81	0.84	0.82

* Period of calibration 1994-1996, * * Period of validation 1992-1993,

** * Period of overall model performance 1992-1996

Basin-Wide Impacts of Land Use Changes

Basin-wide impacts of land use changes on hydrologic characteristics are presented in Figures 5 and 6. In general, the basin was divided into to three major classes. 1) positive high: if percent change in hydrologic characteristics is equal or more than 10% of the original value; 2) modest: if percent change in hydrologic characteristics is between -10% to 10% of the original value and; 3) negative high: if percent change in hydrologic characteristics is equal or less than -10% of the original value (Figure 5).

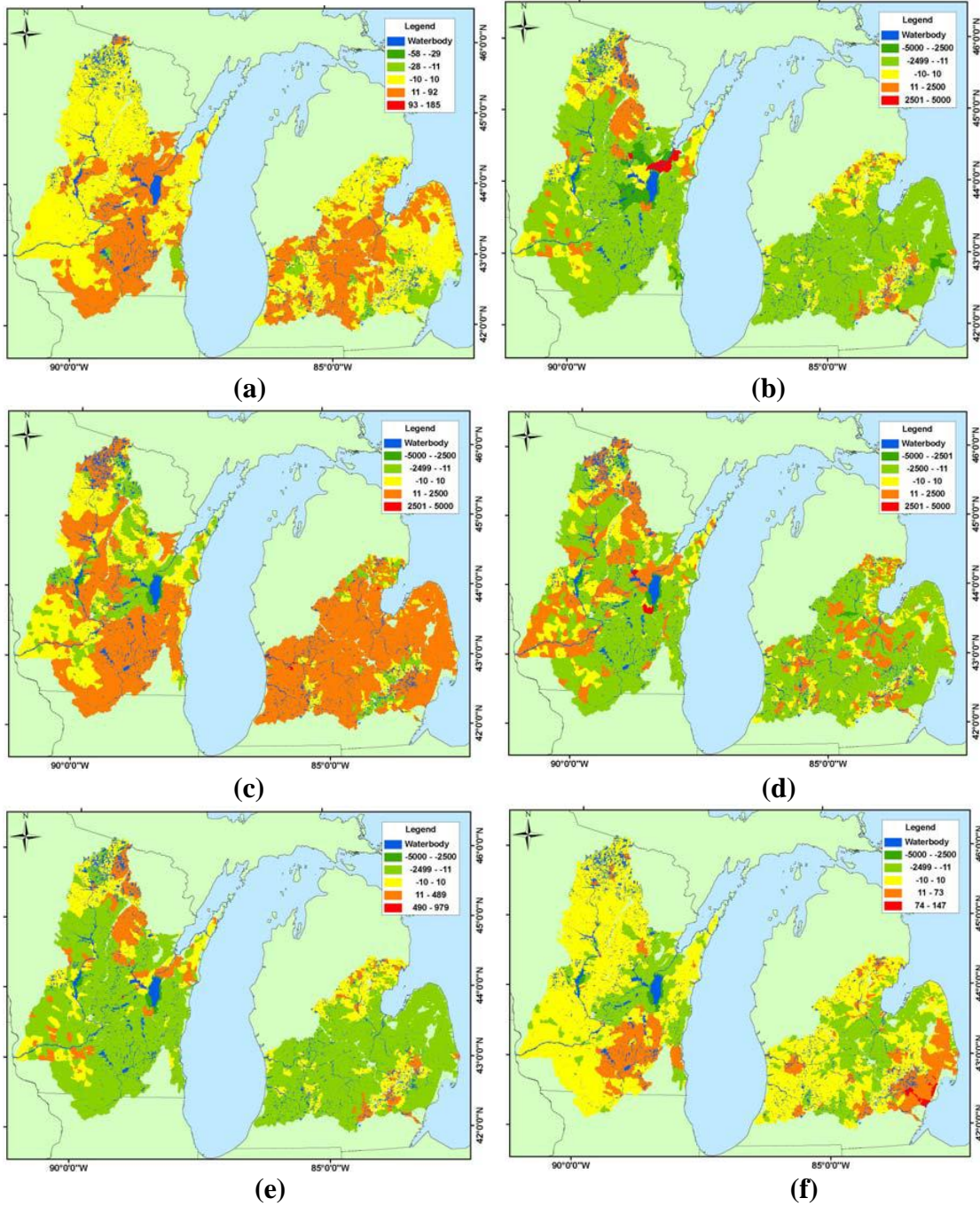


Figure 5. Modeled percent changes resulting from land use change: (a) actual evapotranspiration; (b) recharge entering aquifers; (c) surface runoff; (d) lateral flow contribution to streamflow; (e) groundwater contribution to streamflow; and (f) water yield. *Note: Values > 5000% or < - 5000% are reported as \pm 5000%.*

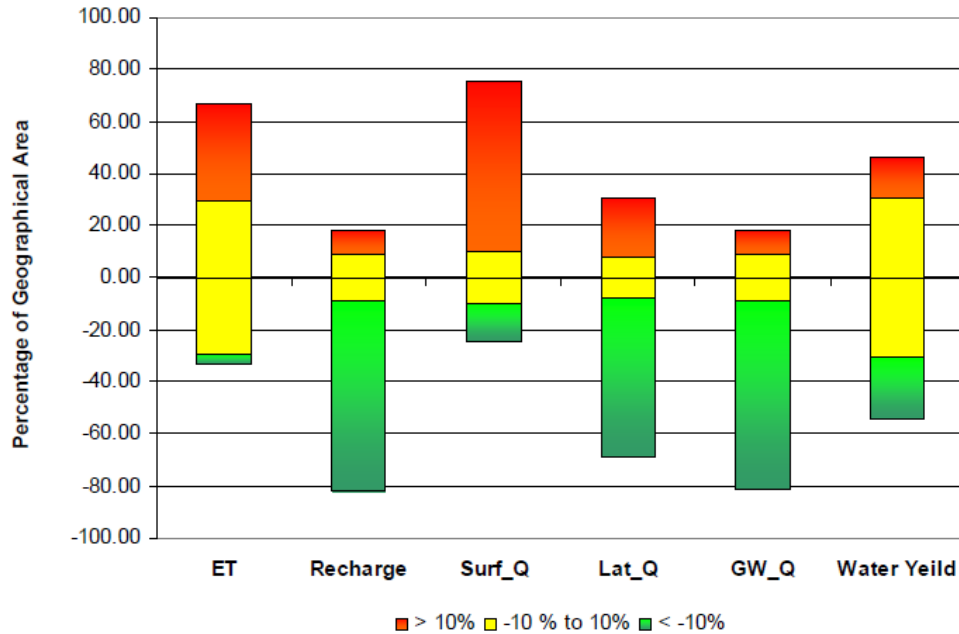


Figure 6. Percentage of project area falling into 3 change classes of: a) positive high, b) modest, or C) negative high classes; (ET) actual evapotranspiration; (Recharge) recharge entering aquifers; (Surf_Q) overland flow contribution to streamflow; (Lat_Q) lateral flow contribution to streamflow; (GW_Q) baseflow contribution to streamflow; and water yield

Figures 5a and 6 demonstrate that percent change in evapotranspiration is modest in the majority of the basin, particularly in the northwest region of the study area in which forested lands are generally preserved. In addition, decreases in evapotranspiration can be observed especially in heavily populated areas such as Detroit (MI) and Milwaukee (WI). Regarding recharge to aquifers and baseflow, more than 70% of the study area is classified as negative high. This can be attributed to conversion of forestlands to agricultural lands that have lower recharge potentials (Figures 5b, and 5e). Overland flow contribution to streamflow (*Surf_Q*) was increased in majority of the region in comparison to pre-settlement scenario. In fact, more than 65% of the study area is classified as positive high concerning overland flow which can be explained by vast expansion of agricultural lands in the region. The majority of the region experiences modest changes in water yield, while about 15% of region is classified as positive high and 24% is classified as negative high. The positive high region mostly corresponds to urbanization and the negative high region is mostly associated to conversion of wetlands, rangeland and forested areas to agricultural production.

Collectively the results demonstrate that the hydrology of the Great Lakes region have been altered due to major land use change from pre-settlement conditions over the past 150 years. More specifically the results demonstrate that at the basin-level, modest changes in evapotranspiration and water yield, significant increases in overland flow generation, and significant decreases in recharge, baseflow, and lateral flow in the majority of the basin were observed. Land use changes such as urbanization, deforestation, and reforestation have and continue to affect groundwater-surface water interactions and associated instream physical habitat, water quality, and flows. The focus of objective 2 is to use these data to assess the relation of fish community metrics to these modeled historic and current instream conditions.

OBJECTIVE 2:

Develop models that predict selected riverine biological endpoints based on SWAT variables and other relevant watershed and local catchment variables

OBJECTIVE 2 METHODS

Selection of Conservation Practices

In Phase 2 of the Great Lakes CEAP Project we will be working with NRCS conservationists, conservation districts and other key partners to develop detailed conservation blueprints, implementation schedules and cost estimates for implementing a select subset of conservation practices within select priority subwatersheds of the Saginaw Bay watershed. Although this is a Phase 2 objective, knowing what specific practices will be used in those scenarios is critical to certain steps being taken in Phase 1 related to establishing “caps” on fish community expectations due to factors or conditions that are either not adequately addressed by SWAT and/or not adequately addressed by the selected set of conservation practices. Because we wanted to keep these scenarios realistic, we quantified the prevalence of practices implemented from 1999-2009 across the project area, using the NRCS Conservation Practice Database (USDA-NRCS, National Conservation Planning Database, October, 2009). From this analysis we selected the nine most prevalent practices across the region that also addressed the three issues of altered flows and increased nutrients and sediments that are consistently cited as the most critical stream habitat problems within our study area. We supplemented this analysis with input on which practices SWAT will be able to effectively model from Dr. Amirpouyan Nejadhashemi and with expert input on the relative benefits of less prevalent practices. Experts consistently cited the benefits and need for more wetland restoration in the project area, which are supported by numerous studies (Craft and Casey 2000, Mitsch and Day 2006), so we added two additional practices for a total of 11 practices that will be included in our SWAT conservation scenarios for Phase 2 of the project (Table 2).

Table 2. Conservation practices for which Phase II modeling will focus.

Practice Name
Nutrient Management/Waste Utilization
Conservation Crop Rotation
Filter Strip
Conservation Cover
Residue and Tillage Management, No-Till/Strip Till/Direct Seed
Mulch Till, Residue Mgt & Residue and Tillage Mgt
Residue Management, No-Till/Strip Till
Cover Crop
Pasture and Hay Planting
Wetland Creation/Restoration
Wetland - Floodplain restoration

Selection of Biological Endpoints (Response Variables)

Biologically meaningful endpoints for setting goals and guiding watershed restoration could be developed for a variety of taxonomic groups. However, fish and aquatic macroinvertebrate communities have largely been the focus of such efforts due to availability of data for these two taxa (Berkman et al. 1986; Plafkin et al. 1989). Fish assemblages have some added advantages of being more highly valued as a resource and more readily understood by the general populace when addressing conservation issues (Karr 1981). Fish also cover many trophic levels, including piscivores, herbivores, omnivores, and insectivores and have a breadth of other functional traits (e.g., modes of reproduction) that make them sensitive to a variety of habitat variables and thus sensitive to a range of human disturbances. Furthermore, fishes exhibit a range of life spans and mobility which helps to detect both long-term and broad scale disturbances to freshwater ecosystems (Karr 1981, Babour et al. 1999). For these and other reasons we selected instream fish communities to serve as the biological endpoint for our project.

As suggested above, there are many possible metrics that could be developed based on fish community composition that could serve as indicators of biological integrity or stream health. Recognizing this Karr (1981) developed the multi-metric index of biotic integrity (IBI), which integrates several individual metrics into an overall measure of stream health. The original IBI consisted of 12 metrics and was developed for the Midwest United States. Over the years the original IBI has been modified and customized to specific geographic regions, and successfully used to assess biological integrity of streams (Lyons 1992, Lyons et al. 1996, Roth et al. 1996, MDEQ 1997, Lammert and Allan 1999, Wang et al 2008). Because of its integrative nature and successful application in Midwestern streams, our analyses and modeling efforts for this project focused on the IBI.

We also used a subset of the functional guild metrics that make up the IBI due to their known sensitivity to instream habitat disturbances that are often associated with agricultural practices.

Evaluating biological communities from a functional guild perspective provides a means for identifying the primary pathways in which a particular disturbance is transmitted throughout an ecosystem (Austen et al. 1994; Merritt and Cummins 1996; Poff 1997). The specific functional guild metrics are calculated as a percentage of the overall fish community and include; percent omnivores (PCOMNINB), percent insectivores (PCINSENB), percent lithophilus spawners (PCLITHNB), and percent piscivores (PCPISVNB). Since the ratio of these individual guild metrics are often very informative due to their ability to demonstrate community dynamics (e.g., predator to prey interactions) with a single metric, we also included a piscivore to insectivore ratio (PISINSRATIO) in our analyses. Finally, we also included a metric that quantifies the percent of intolerant individuals (PCINTONB) within the sample. This metric was of interest because the binary designation of tolerant versus intolerant fish species is largely a reflection of that species sensitivity to water quality conditions, which are a primary concern in intensively agricultural landscapes, which includes our project area.

The IBI scores used in our project are calculated depending on both the size (wadeable or non-wadeable) and thermal (cold, cool, or warm) classification of the stream. Specifically, a modified procedure developed by the MDEQ (1997) was used to determine IBI scores for wadeable warmwater streams. IBI scores for wadable coldwater sites were calculated based on procedures described by Lyons et al. (1996). For cool water sites, IBI scores were calculated based on both of the preceding methods and the higher of the two IBI scores was used. For the larger, non-wadeable rivers, the IBI scores were calculated following the scoring criteria developed by Lyons et al. (2001).

Sources of Fish Community Data

The fish community dataset that provided the biological response variables for this project consisted of 1022 fish community collections that were made between 1982 and 2007 and standardized across Michigan and Wisconsin (See Figure 1). These data were provided by collaborators Li Wang of the Michigan Department of Natural Resources, Institute of Fisheries Research and John Lyons of the Wisconsin Department of Natural Resources. The dataset included IBI scores calculated for each site, based on the methods described earlier, as well as values for each of the individual component metrics.

Selection of Predictor Variables

For this objective we are trying to extend the predictive capabilities of SWAT to include biological endpoints. This would provide us with the ability to move from retrospective assessments of biological conditions to forecasting such conditions under future conservation scenarios. As a result, our analyses and modeling efforts were primarily focused on identifying relations between fish community metrics and instream habitat (water quality and flow) variables generated by SWAT. However, we also fully recognize that riverine fishes are influenced by numerous landscape and in-channel factors and processes operating at multiple spatial and temporal scales (Rabeni and Sowa 1996). Of particular interest are those natural landscape factors and human disturbances operating within the overall watershed and local catchment draining to a stream segment (Sowa et al. 2007). Watershed and local catchment metrics, like percent of a particular surficial geology or percent impervious surface, can indirectly capture

habitat patterns and processes (e.g., stream channel morphology, thermal regime, bedload movement, etc.) that are not effectively captured by discrete field samples or even modeled by complex and temporally dynamic models like SWAT. Failing to account for these factors, that often serve as higher level constraints on fish communities, could lead to erroneous expectations in Phase 2 of our project as we develop future conservation scenarios with SWAT that will not address the full suite of potential limiting factors. Consequently, to supplement the predictor variables provided by SWAT we also included a broad suite of predictor variables pertaining to overall watershed and local catchment physiography (termed *Natural Variables*) and non-agricultural human disturbances (termed *Non-Target Threat Variables*).

Sources of Predictor Variables

Water Quality and Flow Variables—All of our instream habitat variables came from a relatively detailed SWAT model developed specifically for this project and detailed earlier under Objective 1. However, we further summarized the resulting SWAT outputs in order to put them into more ecologically meaningful set of; a) seasonal and annual instream reach loadings and concentrations and b) annual local subbasin runoff and sediment and nutrient contributions. Seasons for calculating the seasonal data were assembled based upon a visual assessment of seasonal hydrologic patterns for seven gaged streams from across the study area (Figure 7). From this assessment we identified four distinct seasons, which we called: Spring Rising (January 15-March 15), Spring Falling (March 15-May 15), Summer Falling (May 15-August 15), and Fall-Winter Stable (August 15-January 15). Water quality variables included numerous flow, nutrient, and sediment variables calculated for total loadings and concentrations, at annual and seasonal time-scales. These loadings and concentrations were calculated under both current and pre-settlement land cover (Figure 8). We also quantified the difference and percent change between pre-settlement and current data for each variables. These partitions of the data resulted in a total of 1,121 water quality and flow predictor variables.

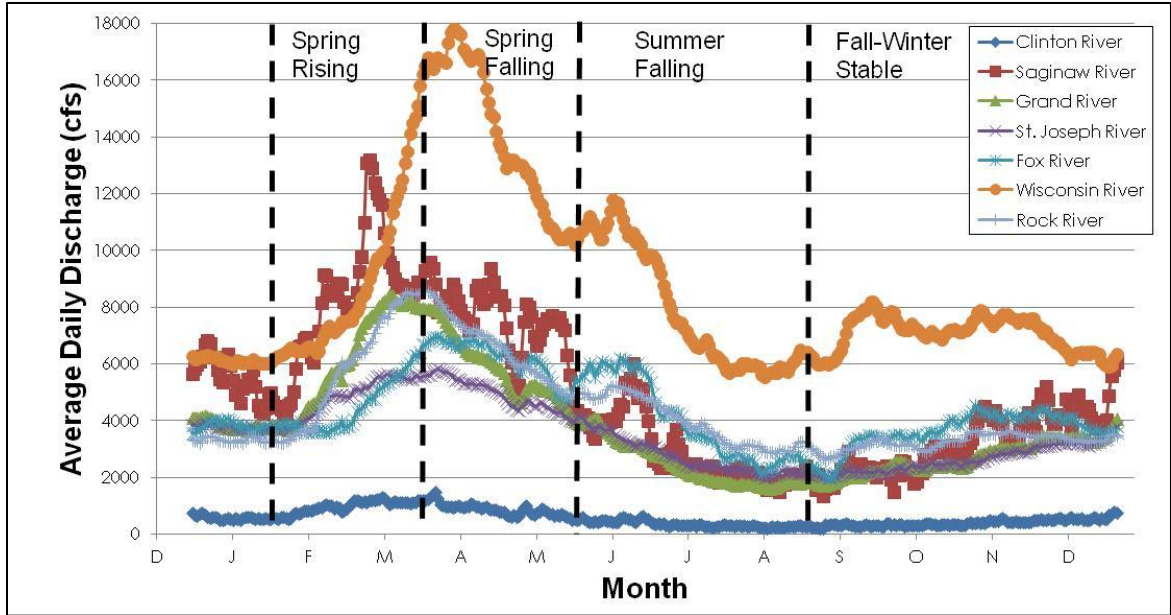


Figure 7. Average daily discharge values for seven streams from across the project area and showing the consistent annual hydrographs we used to for summarizing each SWAT water quality variable into distinct seasonal variables.

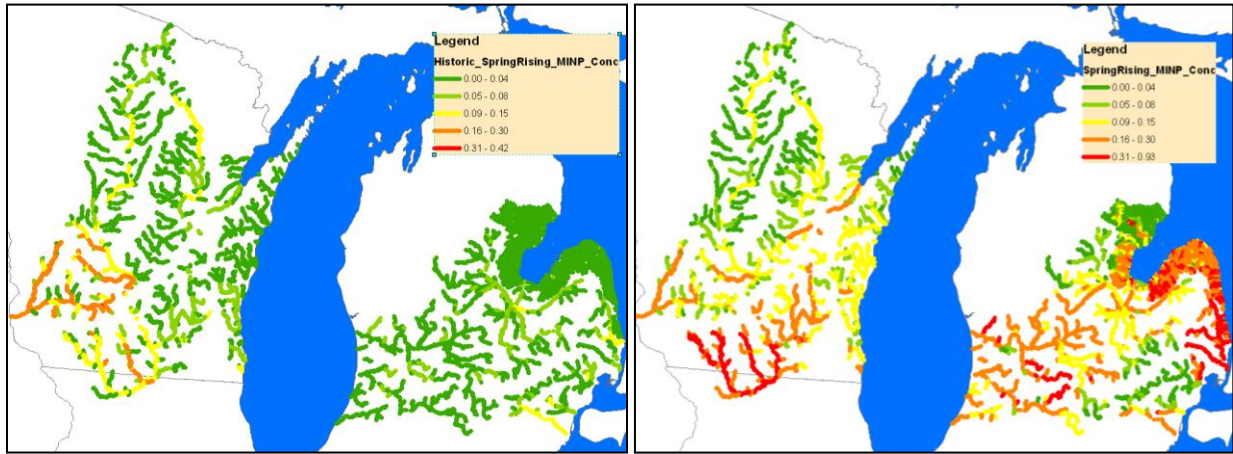


Figure 8. Maps showing predicted mineral phosphorous concentrations (mg/l) during the spring rising season based on historic (left panel) and current (right panel) land use and land cover conditions.

Natural Variables—Water quInstream habitats (including water quality and flow) and biological communities vary across landscapes both naturally, as a result of natural variation in climate and physiography, and as a result of human disturbances. In order to accurately relate fish community indicators to water quality and flow, it is important to account for variation attributable to natural variables, which have repeatedly been shown to be important in relating biological communities to water quality or anthropogenic stresses (Richards et al. 1996, Fitzpatrick et al. 2001, Wang et al. 2003). As such, we used the broad suite of natural

physiographic variables assembled as part of the USGS Great Lakes Aquatic Gap project, as well as variables assembled for the National Fish Habitat Action Plan (NFHAP; Esselman et al. 2011, Wang et al. 2011) to first relate to fish community indicators to identify dominant natural predictor variables. The 600 natural variables included measures of stream size, network position, hydrologic and thermal regime indices, surficial and bedrock geology, and natural land cover. Natural variables were quantified for five distinct spatial units; channel, local riparian, local catchment, upstream riparian, and overall watershed.

Non-Target Threat Variables—As mentioned earlier, the SWAT modeling used to develop water quality and flow predictor variables, did not fully account for all anthropogenic stresses that can significantly impact water quality, flows, physical habitat, and ultimately biological communities. For example, extremely high cattle densities can influence water quality, but cattle densities were not incorporated into the SWAT model inputs and therefore were not accounted for in the water quality predictions. Also, the twelve practices we selected are not ideally suited to addressing runoff from extremely high density cattle areas like confined animal feeding operations. Therefore, it was also important to account for variation attributable to these and other threat variables. We used the threat variables assembled for the NFHAP to relate to fish community indicators to identify dominant threat predictor variables. The 98 threat variables from this dataset included cattle density, dams, human population densities, and water withdrawals. Threat variables were quantified for both the overall watershed and local catchment.

Spatially Integrating all Response and Predictor Variables

The most difficult aspect of projects dealing with spatially-explicit data involves integrating multiple datasets that are geographically linked to different geospatial baselayers. Unfortunately, in order to integrate the full set of response and predictor variables into a single common dataset suitable for analysis we had to work with three distinct stream layers across our project area.

The NFHAP dataset was developed using the 1:100,000 scale National Hydrographic Dataset Plus (NHDPlus) as the baselayer (Esselman et al. 2011). The NHDPlus is a nationwide highly improved 1:100,000 scale hydrography datasets, which contains network of related streams, local catchments, and network catchments. The dataset contains flow direction, flow accumulation, and elevation data that can be used to study various local to network level phenomenon (<http://www.horizon-systems.com/nhdplus/>). All 1022 fish community sampling locations had already been spatially linked to the Great Lakes Aquatic GAP stream network, via a unique locational id: PU_GAPCODE. Fortunately, the Aquatic GAP stream network represents a modified version of the 1:100,000 NHD-Plus (Wang et al. 2011), which allowed us to cross-walk this modified network back to the original NHD-Plus, via the shared COMID attribute and integrate it with the NFHAP data for most stream segments. Finally, the most difficult task was spatially linking the fish community samples to the stream network used for developing the SWAT models across the region. This SWAT stream network is a much more generalized stream layer containing which was developed using the ArcSWAT tool and a 30 meter digital

elevation model (DEM) layer. We had hoped to retain all or at least most of the 1022 samples during this process. However, that would have required generating SWAT subbasins with outlets occurring at each of the 1022 sampling locations. Unfortunately, this was not technically or logistically feasible at the time and so we ended up losing nearly 70% of the fish community samples in this process. Furthermore, this process had to be done manually by visually linking sites to the appropriate subbasin to ensure that the SWAT model predictions correctly corresponded with the specific stream segment at which each fish community sample was made. As a result, we were able to successfully link only 345 of the 1022 fish community samples to the DEM derived stream network used for SWAT modeling (Figure 9).

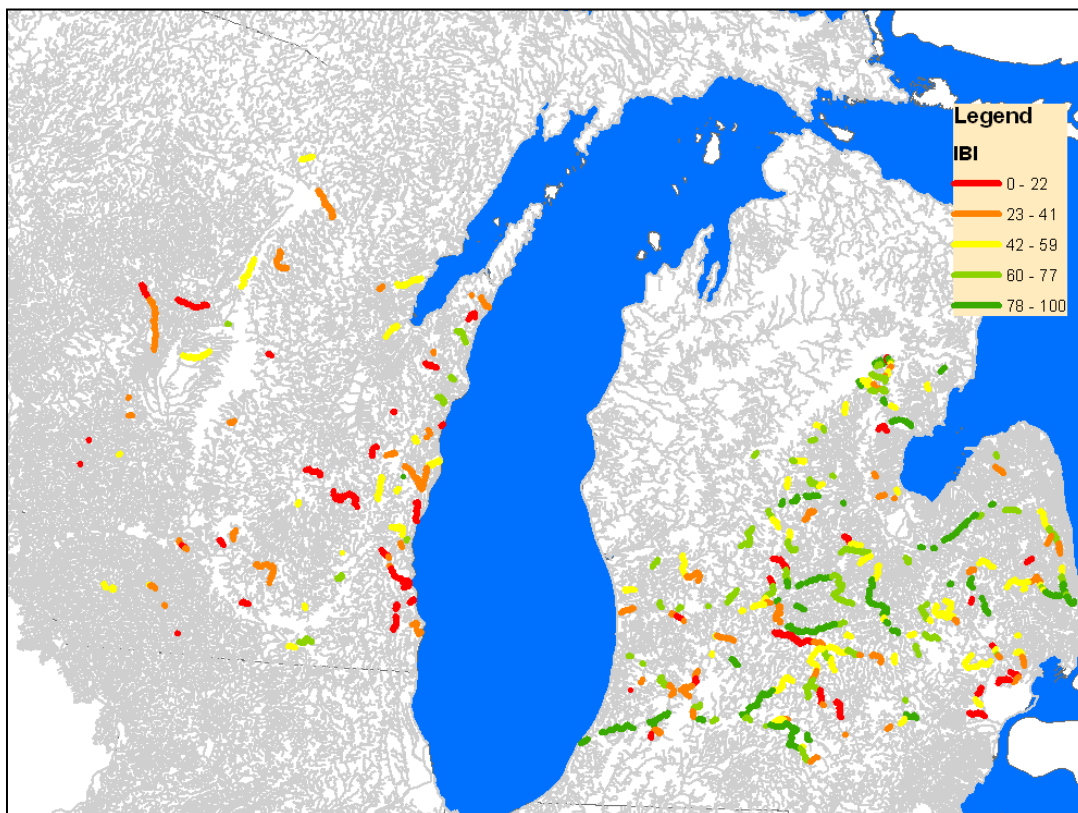


Figure 9. Map showing the location and IBI scores for the 345 fish community samples that could be spatially linked to the DEM derived stream network attributed with SWAT modeled values for instream water quality and flow.

Data Transformation and Reduction

Natural, threat, and water quality variable datasets were all analyzed for normality using skewness and kurtosis distribution tests. Variables with skewness or kurtosis values ≥ 3 were log transformed ($\log x+1$), or in the case of proportional data arcsine transformed, to attain or approximate normality. Variables with $>90\%$ zero values were deleted from analyses. Prior to performing CART modeling (see below), transformed data that remained non-normal were further transformed by placing them into bins based on distribution breaks in the data. This was

done to diminish the influence of outliers by ensuring the relatively high sample size across the range of values for each variable. Specifically, bins were created to ensure that all bins maintained at least 10% of the total data points for that variable.

Statistical Analyses

Our analyses for objective 2 focused primarily on three sets of complimentary analyses to help;

1. identify influential predictor variables and their relative degree of influence
2. identify biological thresholds and constraints for the fish community metrics, and
3. develop predictive models within a single hierarchical model or via a set of multiple models based on wedge plots

These complimentary analyses consisted of Redundancy Analysis, Classification and Regression Trees, and Simple Scatter and Wedge Plots. Redundancy Analyses and Classification and Regression Trees were used to generally evaluate which natural, threat, and water quality variables were influential across all fish community metrics, and which fish community metrics were most responsive to water quality variables. These analyses provided both a multivariate (Redundancy Analysis) and univariate (Classification and Regression Trees) assessment of predictor variables. Through these analyses, we were then able to proceed to subsequent analyses (wedge plot evaluations) with a smaller subset of variables that we knew were predictive of fish community metrics. Classification and Regression Trees were also used to attempt to predict IBI metrics, based on water quality and flow. Wedge plots were used to identify thresholds, above which a predictor variable fundamental limits a fish community metric score, regardless of other factors.

Redundancy Analyses

Redundancy Analyses (RDA) were conducted to evaluate relationships between IBI metrics and natural, threat, and water quality variables using the statistic software CANOCO (CANOCO v4.5; ter Braak and Smilauer 2002). Redundancy analysis is a direct gradient analysis that evaluates linear relationships between multiple dependent and independent variables. Natural, threat, or water quality variables that are predictive of fish community metrics were selected through a forward selection process that uses Monte Carlo permutations (999) to calculate a probability for whether a particular variable is significantly predictive. Separate RDAs were run for natural, threat, and water quality variables, and then a combined analysis to evaluate relationships between IBI variables across all predictor types. For each set of RDAs, analyses were first run with all potential variables where significant variables were selected, then rerun on the reduced set of significant variables. We elected to use RDA instead of canonical correspondence analysis (CCA)--another form of direct gradient analysis that evaluates non-linear relationships—because scatter plots of the relations between the environmental variables and IBI metrics indicated that linear responses rather than unimodal responses prevailed. This analysis helped to identify variables that consistently influence multiple IBI metrics, to evaluate how they influence IBI metrics, and in identifying which IBI metrics are more sensitive to natural, water quality, and threat variables.

For the RDA of water quality variables, we selected three dominant natural variables, drainage area, State, and Darcy (an estimate of groundwater activity based on geological features) as covariates to force into the model prior to performing the analysis; preliminary analyses without these key contextual natural variables were dominated by water quality variables that were correlated with these natural variables. For the combined RDA, only the natural and threat variables significant in the individual RDA models were included, and all land use types except urban were excluded from the analysis, because land use plays a major role in determining water quality variables in SWAT and we wanted to avoid water quality variables getting “masked” as predictors due to selection of land use variables. The SWAT modeling did not sufficiently reflect urban land use, so it was left in the analysis. All water quality variables were included in the combined analysis, in the event that additional important water quality influences might be revealed when including the context of the natural and threat variables in the analyses.

CART Analyses

Fish community metrics and the IBI were also modeled using Classification and Regression Tree (CART) analyses. These analyses were used to better understand the complex relations among the response and predictor variables and the relative strength or nesting of those relations. These analyses were also used to put SWAT variable predictors within the proper landscape/watershed context. CART analyses are nonlinear and nonparametric modeling techniques that use a recursive-partitioning algorithm to repeatedly partition the input data set into a nested series of mutually exclusive groups. Each resulting group is as homogeneous as possible with respect to the response variable (Olden and Jackson 2002). The resulting tree-shape output represents sets of decisions or rules for the classification of a particular response variable relative to a set of distinct combinations of predictor variables. These rules can then be applied to a new unclassified dataset (and corresponding GIS layer) to predict which records or, in our case, location will have a given outcome.

Nonlinear models, like CART, are gaining favor in wildlife-habitat relation modeling because the resulting nonparametric models define constraint envelopes of suitable habitat rather than correlations and thus more formally agree with niche theory (O’Connor 2002). That is, nonlinear models more accurately capture the normal distribution curve that species abundance will typically follow along an environmental gradient (ter Braak and Prentice 1988). Also, nonlinear models do not fall under the standard assumptions of linear, additive or multiplicative relationships, normally distributed errors, and uncorrelated independent variables, which are often unrealistic assumptions that are violated with correlative approaches (Olden and Jackson 2002; Huston 2002; O’Connor 2002). CART analyses, in particular, have become a popular modeling technique because they construct models with accuracy comparable to the more “sophisticated” nonlinear methods (e.g., Neural Networks; Olden and Jackson 2002), and yet are much easier to construct and interpret (Breiman et al. 1984; De’ath and Fabricus 2000).

The specific modeling algorithm we used was Exhaustive CHAID, which is a modification of CHAID developed by Biggs et al. (1991). It was developed to address some weaknesses of the CHAID method. In some instances CHAID may not find the optimal split for a variable since it

stops merging categories as soon as it finds that all remaining categories are statistically different. Exhaustive CHAID remedies this problem by continuing to merge categories of the predictor variable until only two “supercategories” are left and then examines the series of merges for the predictor and finds the set of categories that gives the strongest association with the target variable and computes an adjusted-*p* value for that association. Consequently, exhaustive CHAID can find the best split for each individual predictor and then choose which of these predictors to split on at each level in the tree by comparing the adjusted-*p* values.

Exhaustive CHAID allows the user to specify *a priori* stopping criteria related to the size of the tree (i.e., number of levels) and the minimum number of collection records that can occur in any given child node. These stopping criteria help reduce the probability of gross overfitting of the model which can be a problem with extremely large datasets containing a large number of predictor and/or response variables. We set the maximum number of levels allowable in the final tree equal to 5, which was higher than the number of levels ever achieved. We set the minimum number of collections allowable in a parent node equal to 25 and the number allowable in a child node equal to 10, for a ratio of 25:10. This ratio was selected based on results of trial runs with ratios of 25:10, 30:15 and 40:20. We set the alpha level for splitting and merging equal to 0.05 and used the Bonferoni alpha adjustment to account for the increased likelihood of a Type One error associated with multiple comparisons.

Based on the results of the RDA our CART analyses focused on just two (IBI and %Intolerant species) of the original nine fish community metrics. These two metrics consistently exhibited the strongest correlations to our all sets of predictor variables and minimal intercorrelation. Then similar to the RDAs we first ran CART independently for each set of predictor variables to identify the most informative variables within a predictor set and used this to create a subset of natural, threat, and SWAT variables. We then ran CART models for IBI and %Intolerant using this full subset of predictor variables.

The RDAs and CART models consistently revealed the significant influence of measures on drainage area or stream size with our fish community metrics. Since we were interested in the residual influence of other predictor variables and to simplify our analyses, we elected to stratify our CART analyses into two categories of drainage to account for this overriding influence *a priori*. To help maintain consistency with stream size classes already used within the project area, we based our initial drainage area categories on three categories that were developed for the Michigan Water Withdrawal Assessment Tool (Hamilton and Seelbach 2010); $<80 \text{ mi}^2$ = streams, $80\text{-}300 \text{ mi}^2$ = small rivers, $>300 \text{ mi}^2$ = large rivers. We assigned each stream segment and corresponding fish community sample into these three strata based on their watershed areas and then tested for differences in the fish community metrics between the three categories. We lumped the upper two categories into one category due to a lack of strong distinction between them and for a larger sample size in the resulting categories ($<80 \text{ mi}^2$ = streams, $\geq 80 \text{ mi}^2$ = rivers) (Figure 10). The rest of our CART modeling corrected for drainage area based on these two categories.

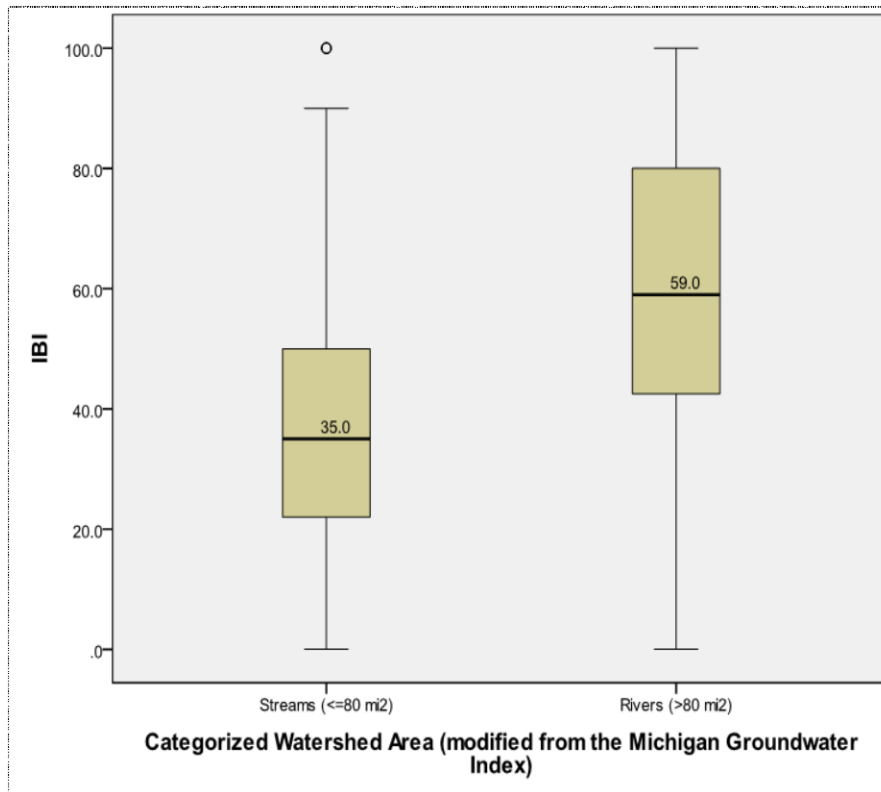


Figure 10. Box plots showing significant differences in IBI scores between streams with drainage areas less and greater than 80 mi². These two categories of Stream and River were used to apriori stratify our CART analyses for examining relationships between fish community and environmental variables.

Fish community metrics sometimes exhibited erratic patterns across the range in values for a particular environmental predictor. When this happened, we plotted the variable to the fish community metric to view where the model had split the data and to further examine patterns or anomalies in the data distribution. Sometimes the erratic patterns matched the overall distribution of the data as demonstrated by linear and loess trend lines. However, when the pattern did not match the readily observable trend across data distribution, we manually binned the predictor variable to increase the sample size of bins across the range of values. To do this the variable was binned based on the distribution shown in a histogram and the trend lines for the scatter. The newly binned variable then replaced the previously unbinned variable and the model was run again. Unfortunately such iterative data transformations were needed to account for our loss of biological data and low sample sizes which required us to use relatively low parent and child ratios (25:10). In such situations the relative influence of a handful of data points can significantly influence an otherwise visible trend. Through this process we were able to make more effective use of our limited data and generate more informative models.

Scatter and Wedge Plots

After winnowing the number of variables down through RDA and CART analyses, scatterplots with the remaining natural, threat, and water quality variables (x-axis) plotted against fish community indicators (y-axis) were examined for trends and for wedge plots. Threat, or non-target disturbances, related to row-crop agriculture (e.g. % row-crop) were specifically not examined since these landscape variables are specifically integrated into the water quality data. Wedge plots occur when a relationship between a predictor variable and response variable results in a wide scatter in the data, because the response variable is influenced by multiple factors, but along the upper limits of the predictor variable (e.g., higher urban land use) the response variable is constrained by the predictor variable so that a wedge is formed along the upper limit of the predictor variable (Brendon et al. 2008). Wedge-shaped relationships are believed to be common along aquatic gradients (Wang et al. 2003). We focused on two fish community metrics, IBI and % Intolerant, because RDAs and other preliminary analyses indicated that these two indicators were generally more responsive to threats, but also water quality variables.

We used wedge plots for natural, threat and water quality variables to identify fundamental limitations in the potential values for IBI or % Intolerant species. While wedge diagrams do not provide the specific potential for any given site, they do provide a threshold above which the response variable is limited across all sites. Upon identification of a wedge, a wedge line was drawn and the equation was generated for the slope along the wedge. Using the original data across all reaches for each natural, threat, or water quality variable with wedges, we calculated the upper maximum potential IBI or % Intolerant species for all stream reaches within the network that had values above any given threshold. These were then mapped across the network of SWAT modeled streams. Limiting natural threat, water quality and flow variable was mapped individually, and results were combined to create maps showing the upper maximum potential IBI or % Intolerant value across all variables, as well as what variable or variable type (natural, threat, water quality/quantity) was most limiting for each stream reach. An improvement capacity map was also created that represents the difference between the maximum potential IBI or % Intolerant species with natural and threat thresholds and the maximum potential based on water quality and quantity. Sites with negative maximum potential values would indicate that the upper maximum is lower based on natural or threat variables, and therefore conservation practices to improve water quantity or quality will not improve the fish community.

OBJECTIVE 2 RESULTS & DISCUSSION

Redundancy Analyses—Twenty-two natural variables were selected as significant predictors in the natural variable RDA. These explained 18.5% of the variation in IBI metrics. Natural variables represented all scales except the local riparian, with the most variables being at the overall catchment or channel scales (Figure 11). Drainage area was the most influential natural variable, as indicated by the fact that its vector in Figure x is the longest. The percent Intolerant

species metric tended to be associated with high Conifer and shrub landcover at the catchment scale and high deciduous forest at the local watershed scale, as well as high drainage scale slope, the Mixed Wood Plains Ecoregion, and channels flowing through coarse moraines. Intolerant species were negatively associated with catchment fine end-moraines and percent sand and gravel, overall riparian carbonate bedrock, and minimum July air temperature. IBI scores were similarly influenced by these variables, but also increased with drainage area. Lithophilic spawners, piscivores, insectivores and the piscivore-to-insectivore ratio were also associated with higher drainage area and somewhat negatively associated with grassland in the local watershed and channels with bedrock depths between 100 and 400 ft. Omnivores tended to have associations opposite to intolerant species and IBI, except that they were also positively associated with larger drainage area and were negatively associated with grasslands. Note that since a forward selection process was used to select variables into the model, each variable independently explains significant variation in the IBI metrics.

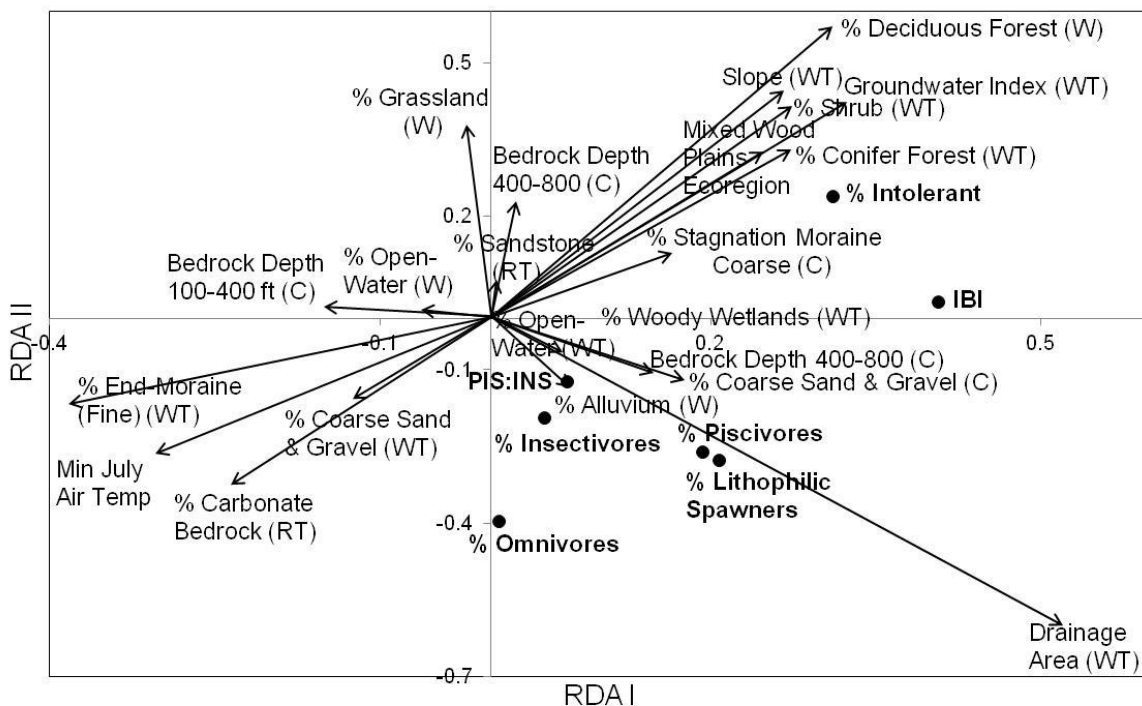


Figure 11. Redundancy Analysis plot showing the relationships between natural variables and fish Index of Biotic Integrity (IBI) scores and six individual IBI metrics. These metrics are the proportional abundance of fish species that are piscivores, insectivores, omnivores, lithophilic spawners, and intolerant of degraded water quality (% Intolerant), as well as a piscivore to insectivore ratio (PIS:INS). Natural variables were quantified at five different scales, channel (C), local riparian (R), local watershed (W), catchment riparian (RT) and catchment (WT). Vectors indicate the direction environmental factors increase in value in relation to IBI metrics.

Vectors also extend in the opposite (negative) direction but for simplicity are not shown. Smaller angles between a vector and an axis indicate higher correlation of the variable with the axis, and longer vectors indicate greater IBI metric variation accounted for. The approximate center of distribution for an IBI metric across an environmental gradient is the perpendicular intersect of a line drawn from its centroid to a vector (positive or negative).

Eleven threat variables were selected as significant predictors in the threat variable RDA. These explained 10.7% of the variation in IBI metrics. Most threat variables selected were at the catchment scale (Figure 12). The percent of the catchment in medium- and low-density urban and row-crop agriculture were the most influential threat variables, as indicated by the length of their vectors. Omnivores tended to be positively associated with each threat, while IBI, intolerant species, and to some extent piscivores tended to be negatively associated with them. The other IBI metrics demonstrated little response to threat variables.

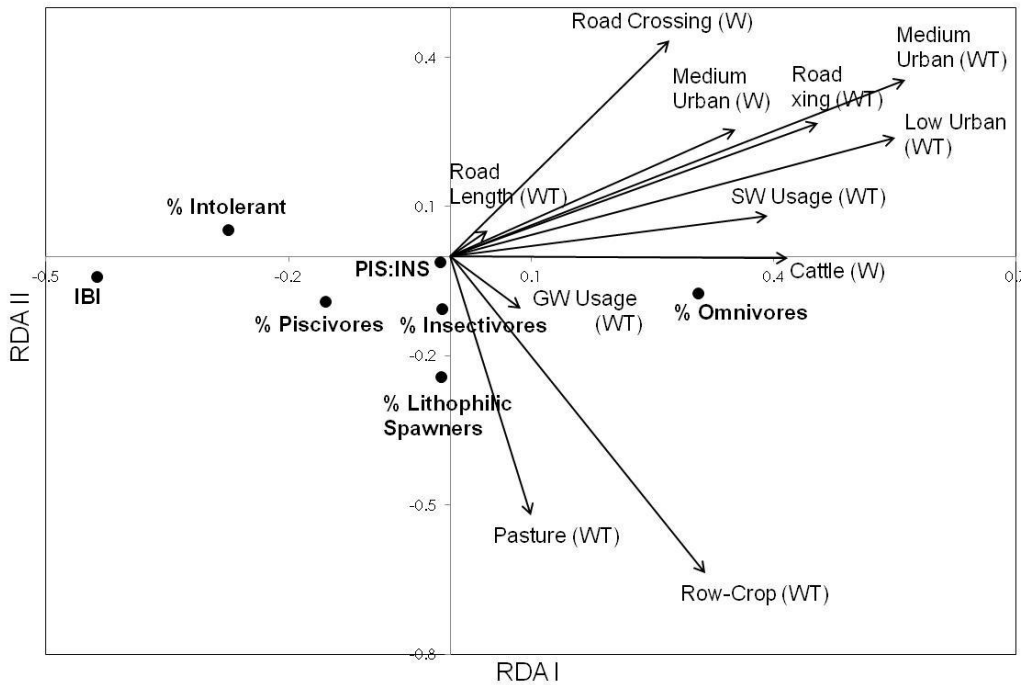


Figure 12. Redundancy Analysis plot showing the relationships between threat variables and fish Index of Biotic Integrity (IBI) scores and six individual IBI metrics (see figure 11). Threat variables were quantified at two different scales, local watershed (W) and catchment (WT).

Ten water quality (SWAT) variables were selected as significant predictors in the water quality RDA. These explained 16.1% of the variation in IBI metrics. Seasonal flow variables were the most influential water quality variables, as indicated by the length of their vectors (Figure 13). These flow variables and spring-rising nitrate (NO₃) concentrations were positively correlated with IBI, insectivores, piscivores and intolerant species. It is important to remember that this model was corrected for drainage area (it was included as a covariables), so the importance of

these flow variables is independent of stream size. As such, they likely reflect a combination of groundwater contributions and differential local climatic conditions (e.g. higher rainfall, lower evapotranspiration). Omnivores were associated with high local surface runoff and lower flows. Lithophilic spawners were positively associated with local sediment phosphorus yield, while intolerant species, piscivores, and insectivores were somewhat negatively associated with it.

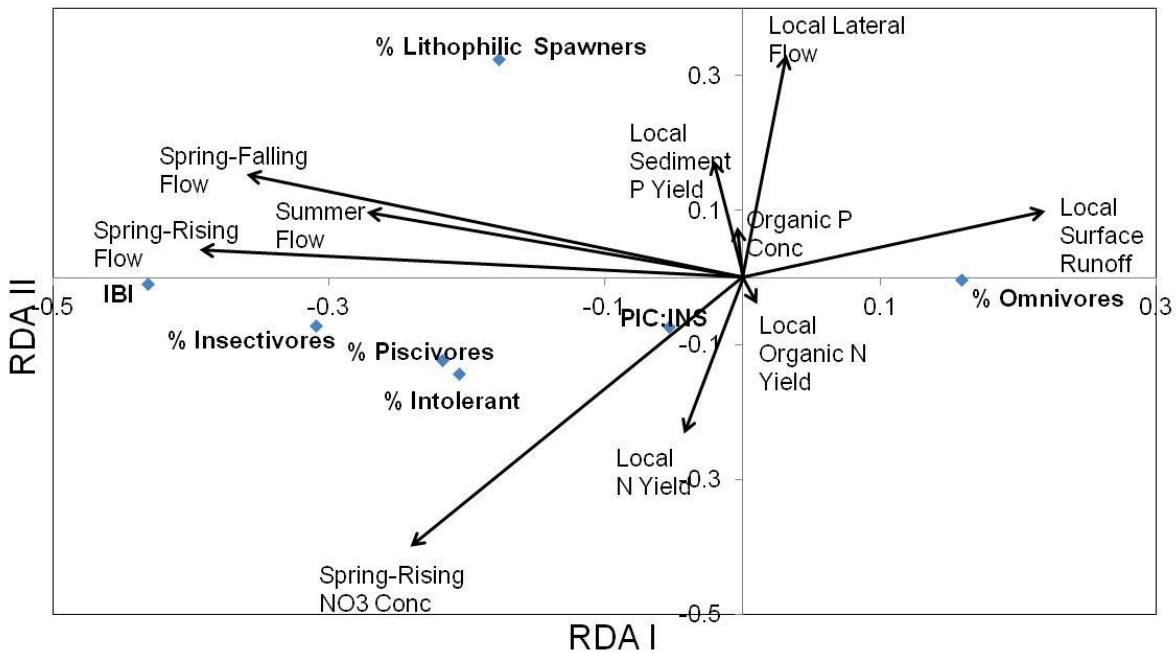


Figure 13. Redundancy Analysis plot showing the relationships between water quality (SWAT) variables and fish Index of Biotic Integrity (IBI) scores and six individual IBI metrics (see Figure 11 for details).

Eighteen variables were selected in the combined RDA, seven natural, five threat, and six water quality (SWAT) variables. These explained 20.6% of the variation in IBI metrics. Drainage area was the most influential natural variable, as indicated by the fact that it exhibits the longest vector in Figure 14. IBI and intolerant species were positively associated with open-water in the local watershed and catchment, surface water usage, woody wetlands, and drainage area, and were negatively associated with urban land use, cattle and alluvium in the local watershed, spring-rising organic phosphorus, organic nitrogen and sediment bound phosphorus runoff, and minimum July air temperature. Insectivores and piscivores were similarly associated, except that they were not as negatively correlated with the phosphorus and nitrogen variables or minimum July air temperature.

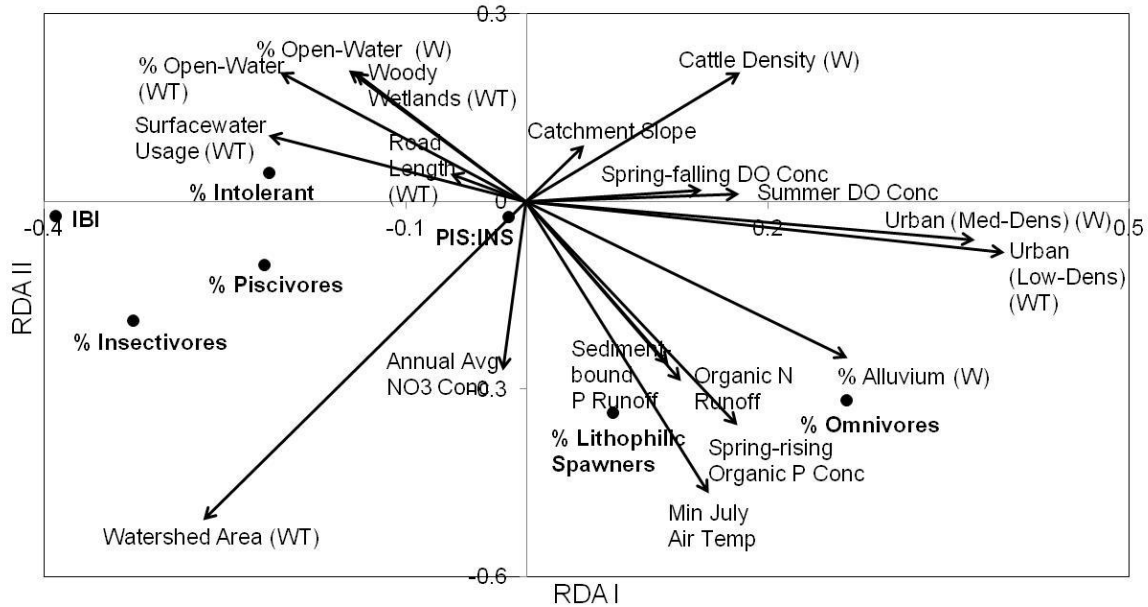


Figure 14. Redundancy Analysis plot showing the relationships between natural, threat, and water quality (SWAT) variables and fish Index of Biotic Integrity (IBI) scores and six individual IBI metrics (see Figure 11 for details).

Overall, natural variables explain more variance in IBI metrics than threat or water quality variables (Table 3). However, the model the combines natural, threat, and water quality variables provides the most thorough explanation of variation in IBI metrics. Across the analyses, IBI and intolerant species consistently demonstrated high sensitivity to (i.e. negative associations with) threats or environmental conditions we consider to be related to threats (e.g., higher nutrients, lower base flow). Similarly, omnivores consistently demonstrated positive associations with these threat or threat-related variables.

Table 3. Variance in IBI metrics explained by Natural, Threat, Water Quality, and Combined RDA models.

Environmental Variable Type	Variance in IBI Metrics Explained
Natural Variables	18.5%
Threat Variables	10.7%
Water Quality Variables*	16.1%
All Variables Combined	20.6%

* Note that the total variance here was reduced because a portion of it had already been explained by the covariables.

CART Analyses

CART model relationships between IBI and water quality (SWAT) variables are shown separately for streams (Figure 15) and rivers (Figure 16). For streams, IBI decreases with increasing runoff in the local subwatershed. Among the lowest runoff streams, IBI decreased with summer nitrate concentrations, but among the highest runoff streams IBI unexpectedly increases with spring-rising sediment concentrations. For rivers, IBI responded with a bell-shaped curve to ammonia. In subsequent tiers of the model, some relationship fit expectations (e.g. decreasing IBI with increasing spring-falling organic phosphorus, spring-rising total nitrogen and spring-falling total phosphorus), but others did not (increasing IBI with increasing spring-rising mineral phosphorus, spring-rising total phosphorus, and summer total phosphorus). Other iterations of the CART models produced similar results, where relationships between fish community metrics and water quality variables were mixed, with some relationships being quite logical and others being illogical. Frequently, illogical relationships included bins with low sample size ($n < 15$). Similar to the RDA, seasonal water quality variables were dominant predictors, with average annual variables rarely occurring in the models.

CART model relationships between IBI and natural variables (Figure 17) and threat variables (Figure 18) were much more complex and produced more logical relationships than water quality models. Watershed area was the first variable selected for both of these models. Other dominant natural variables were related to bedrock type, hydrologic soil group, groundwater index and natural land cover types. Dominant threat variables were related to cattle densities, urban land cover, and row crop land cover. The relative importance of row crop land cover was lower than anticipated though, potentially because row crop was generally predominant across the project area. It is worth noting that the more complex and logical natural and threat CART models are based on the much larger set of fish sites ($n > 1000$) than the water quality models ($n = 345$).

Ideally, we were hoping that the CART analyses would reveal the nested sets of relations where the upper levels of the trees were dominated by natural watershed features and major categories of human threats and that this initial set of strata would serve as meaningful constraints, much like ecoregional strata have been used for developing biocriteria, and then the residual variance, in fish community metrics, remaining within these upper level constraints/strata would largely be explained via relations between SWAT variables. While we saw glimmers of this idealized hierarchy of relations, it was obvious that our analyses suffered from our low sample size of 345 sites where we have SWAT variables linked to fish community samples. A simple factorial exercise illustrates why large sample sizes are needed for these types of landscape scale associative analyses. Four predictor variables, put into 3 categories of low, medium, and high, you end up with 81 distinct combinations of conditions. In order to have samples in each of those distinct combinations, which is the minimum needed to generate a mean and variance, would require 243 fish community samples. Consequently, it is easy to see that losing nearly 700 of our original 1022 fish community samples significantly hindered our ability to generate relations. Since our sample size were not sufficient to provide nested sets of relationships with

representation by natural and threat variables, as well as water quantity and flow variables, analyses to identify ecological thresholds are focused on the results from the wedge diagrams, and resulting upper maxima analyses.

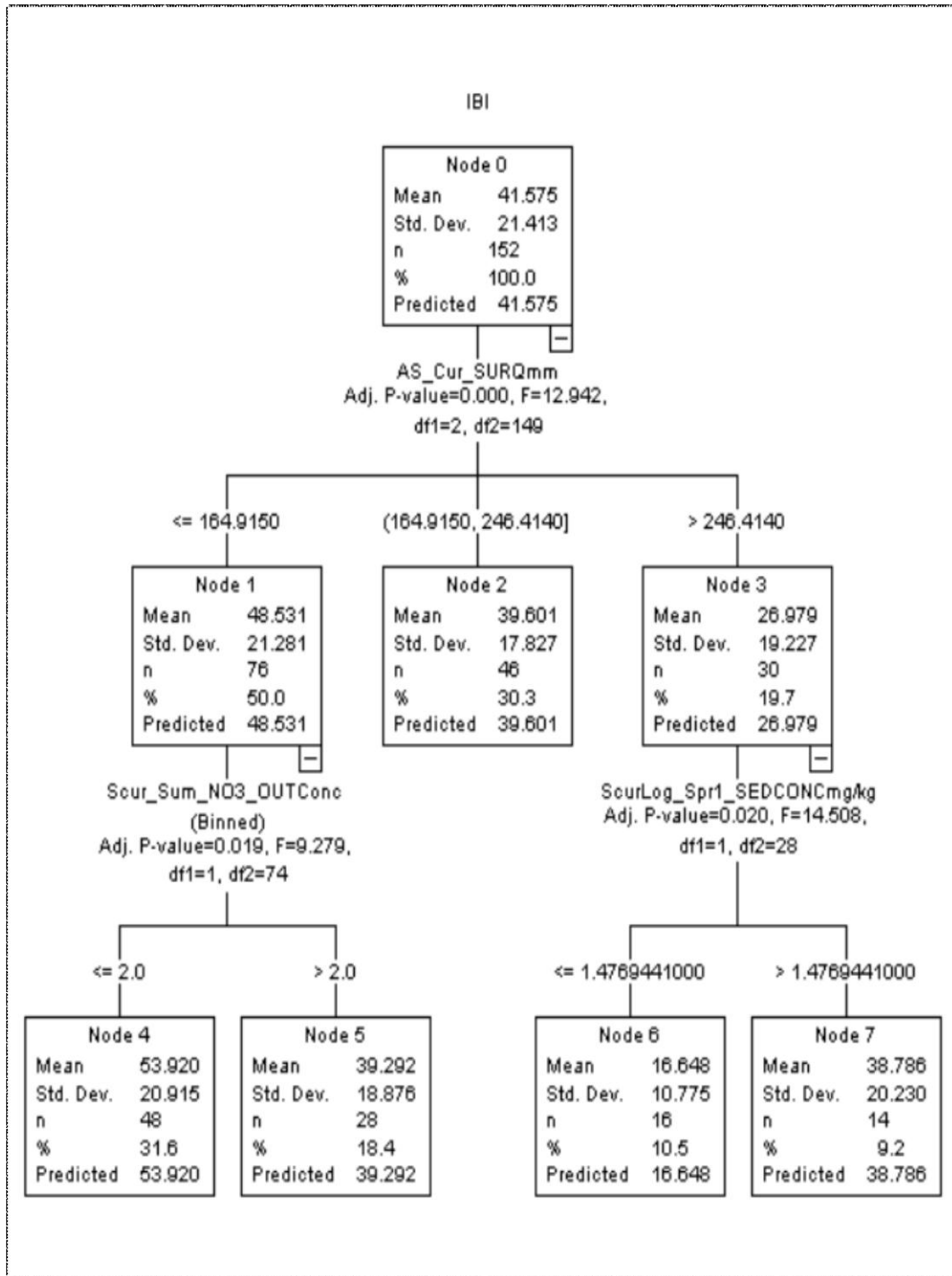


Figure 15: CART model for predicting IBI in streams using SWAT data.

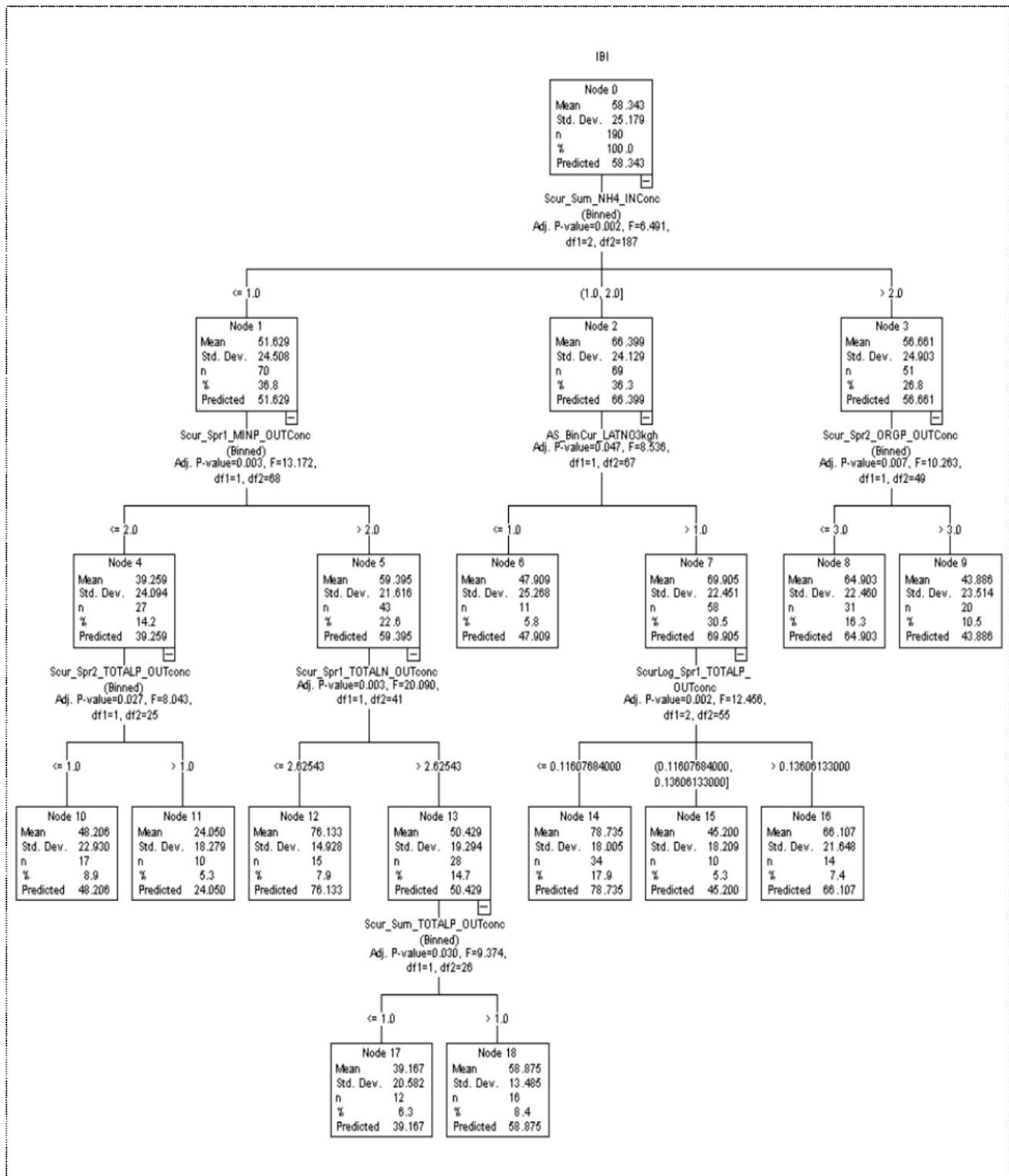


Figure 16: CART model for predicting IBI in rivers using SWAT data.

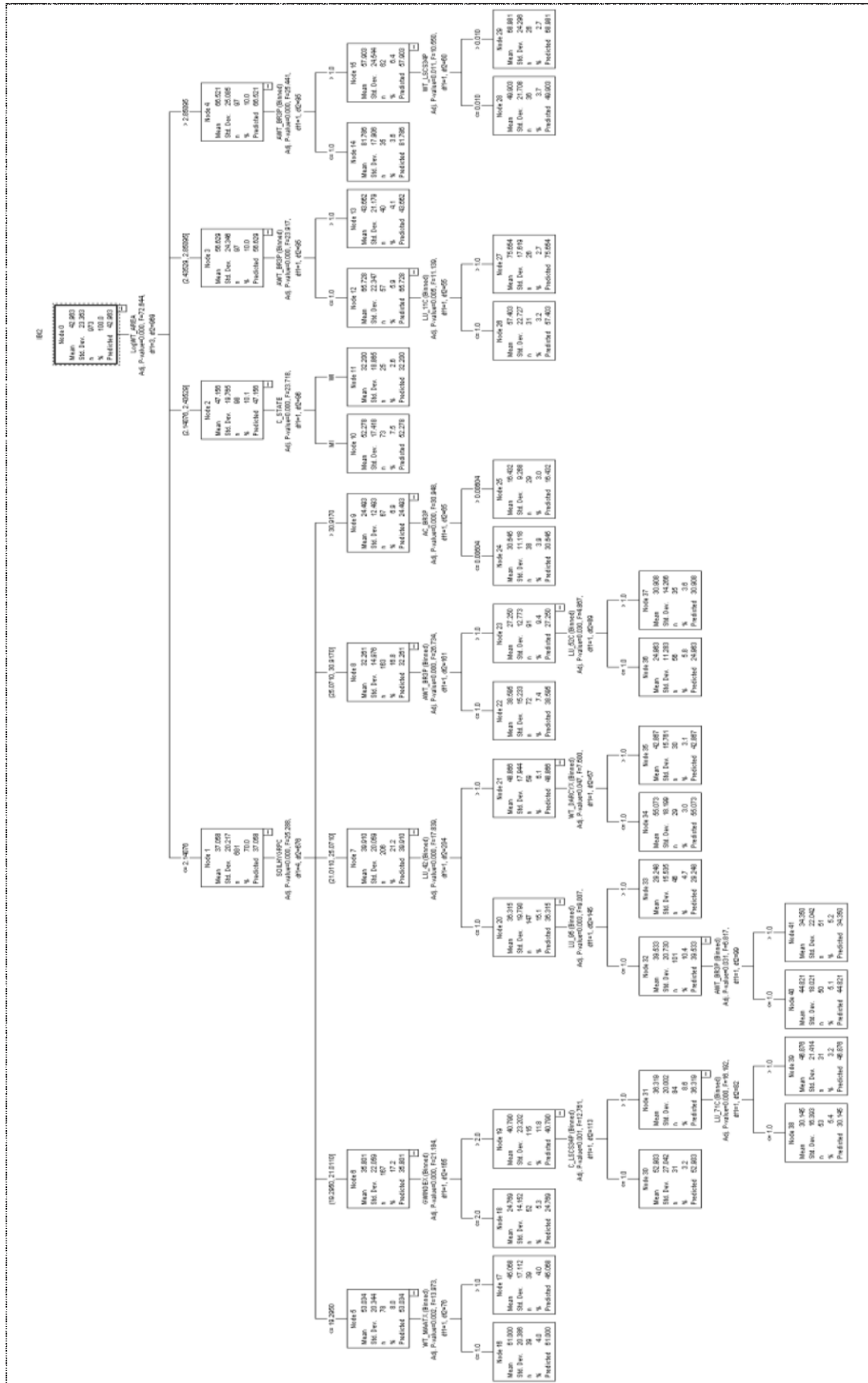


Figure 17: CART model for IBI with natural landscape variables, used to select variables for analysis.

Scatter/Wedge Plots—

Of the six natural and threat variables selected with wedges, three were selected for both IBI and percent intolerant species. Two watershed scale natural variables, fine end-moraine and size of nearest downstream lake (Figure 19), and two watershed scale threat variables, percent impervious and average cattle density (Figure 20), were selected as scatterplots that exhibited wedge relationships with IBI. Three watershed scale natural variables, size of nearest downstream lake (Figure 21a), groundwater index, and downstream Link (D-link), and two watershed scale threat variables, percent impervious (Figure 21b) and average cattle density, were selected as scatterplots exhibiting wedge relationships for percent intolerant species. For most wedges, the majority of sites fell below the threshold, so most sites do not appear to be limited by an upper maximum potential limitation specifically from the particular variable in question.

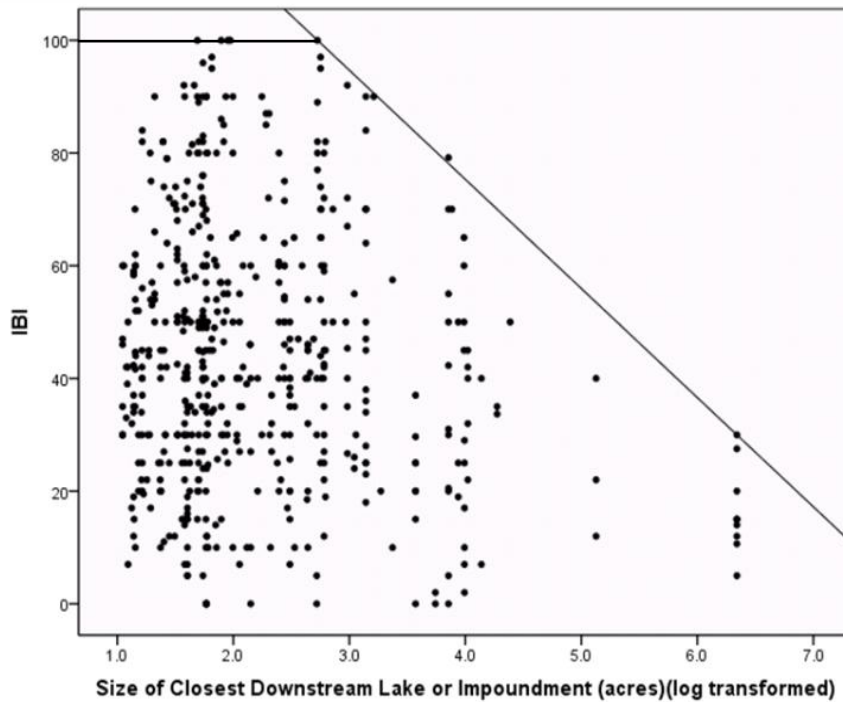
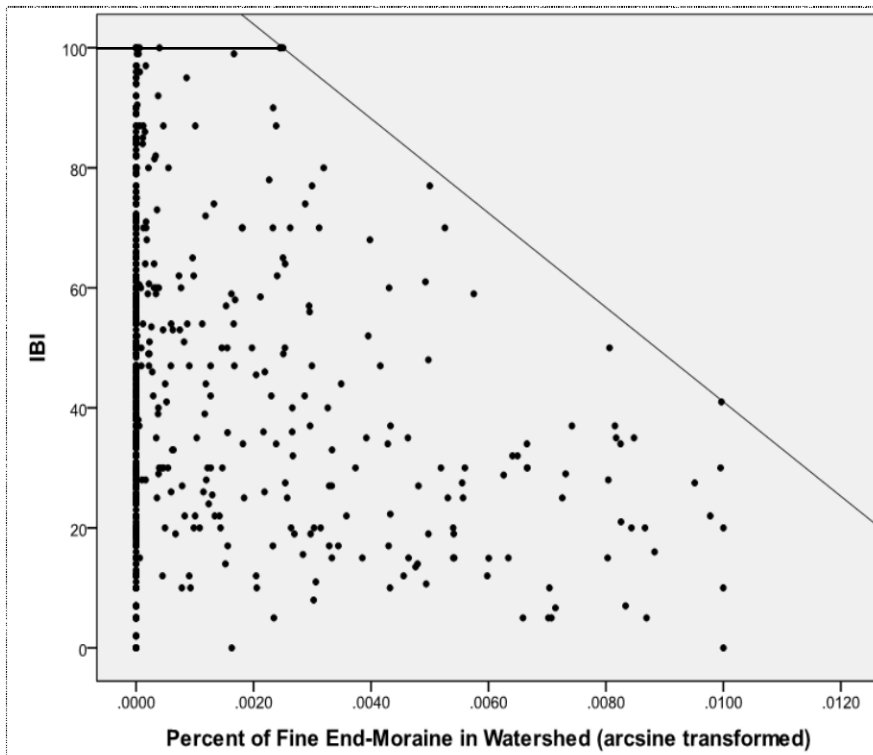


Figure 19: Index of Biotic Integrity with (a) proportion of fine end-moraine in the watershed (arcsine transformed) and (b) size of closest downstream lake or impoundment (log transformed). The wedge lines shows the upper maximum potential IBI above the threshold used to cap data.

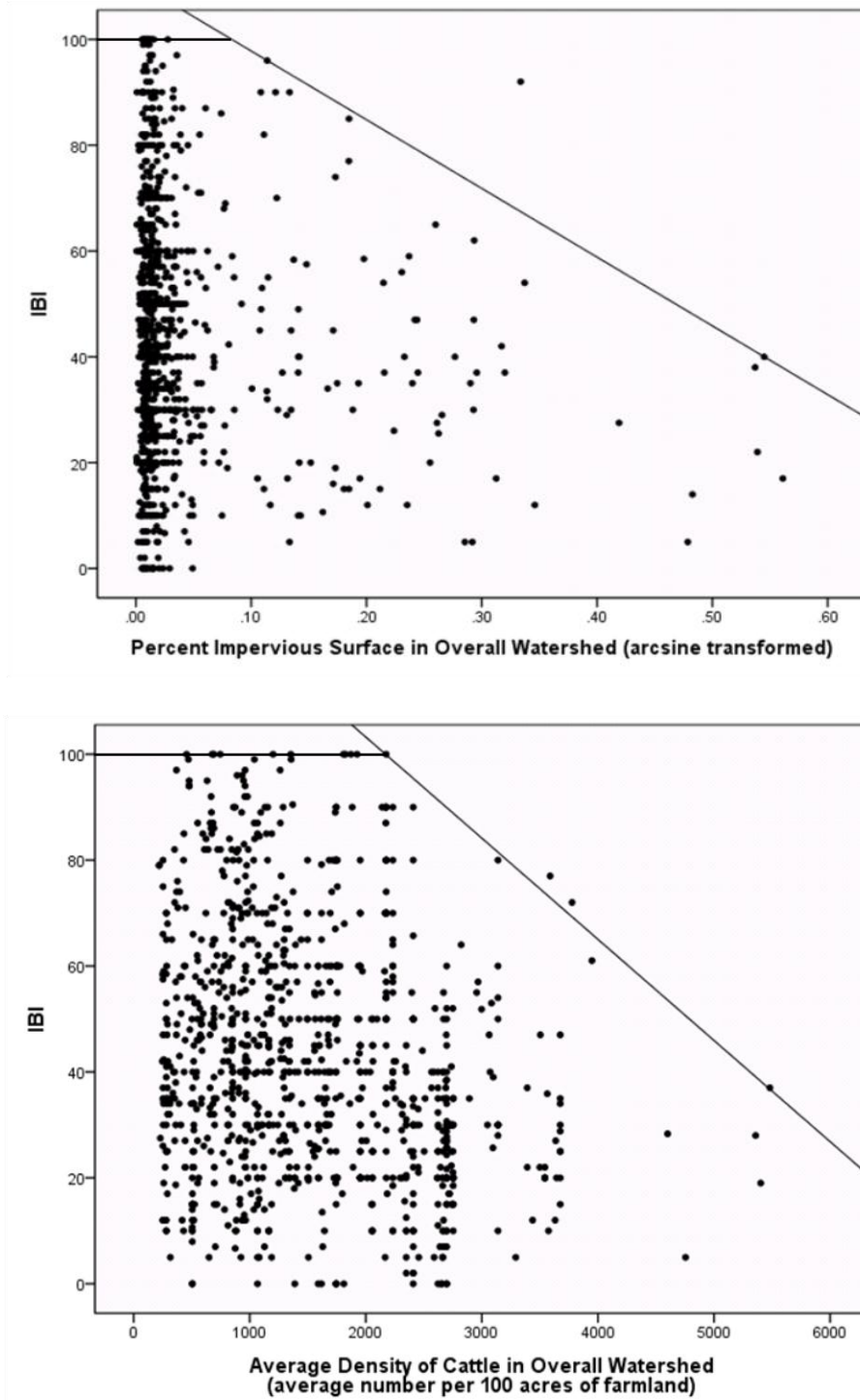


Figure 20: Index of Biotic Integrity with (a) percent impervious surface in the watershed (arcsine transformed) and (b) average cattle density in the watershed. The wedge line shows the upper maximum potential IBI above the threshold used to cap data.

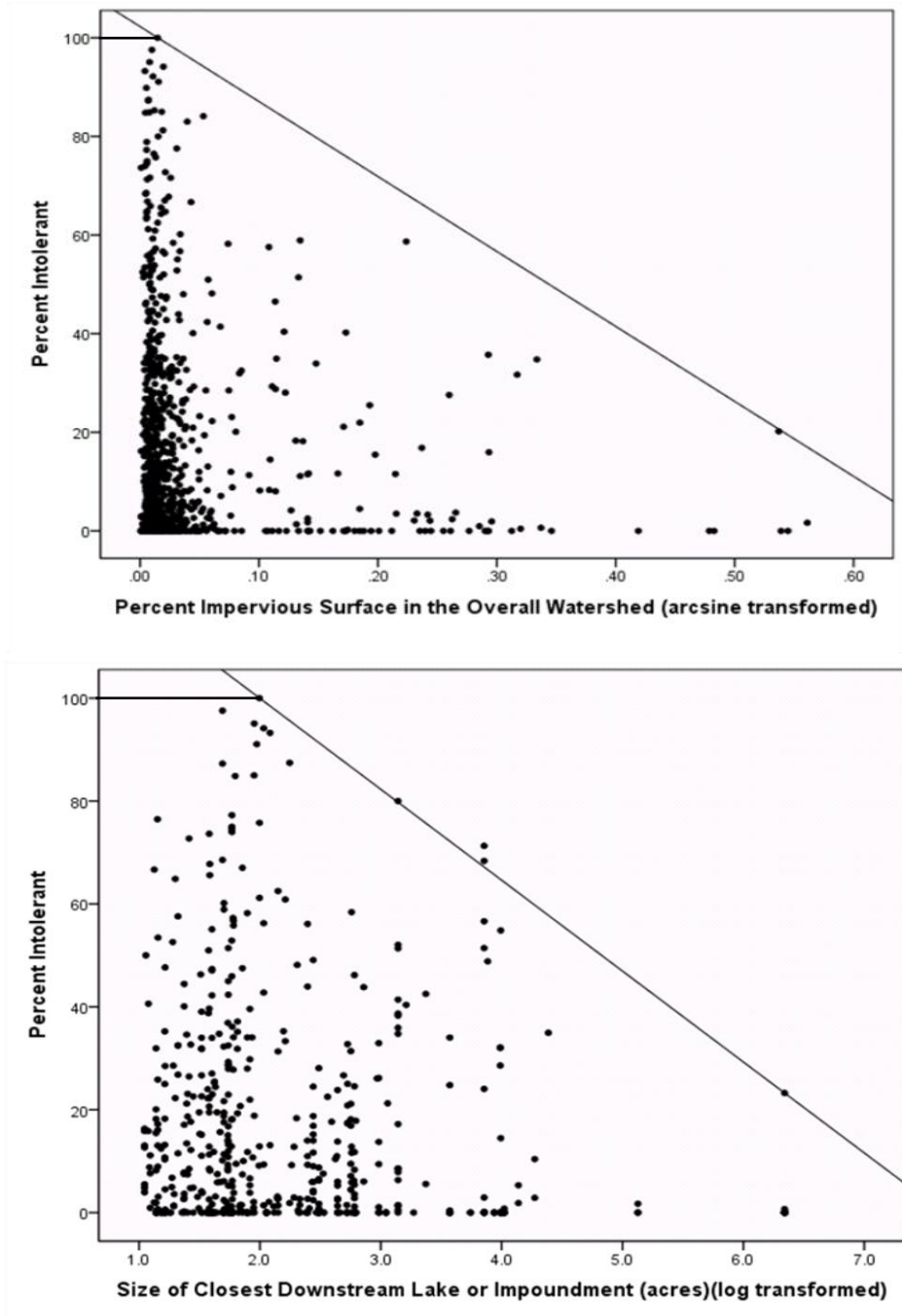


Figure 21: Percent Intolerant Species with (a) size of closest downstream lake or impoundment (log transformed) and (b) percent impervious surface in the watershed (arcsine transformed). The wedge line shows the upper maximum potential Percent Intolerant above the threshold used to cap data.

The spatial distribution and extent of stream reaches where the potential IBI or percent intolerant species would be limited by the natural and threat variables selected for capping is highly variable. For example, the natural variable percent fine end-moraine in the watershed effected fairly small areas, whereas the size of the nearest downstream lake effected large areas, especially in central Wisconsin (Figure 22). Similarly, the threat variable percent impervious surface effected potential IBI in urban areas scattered throughout the project area, but in larger concentrations around Chicago and Detroit (Figure 23), whereas the average cattle density was widely distributed as a limiting variable, but mostly in Wisconsin. Percent intolerant species were impacted by watershed groundwater index values over large areas, but like IBI is only limited by percent impervious around significant urban areas (Figure 24). The effect of impervious surfaces was slightly more widespread for IBI, but the limitations were generally more intense for percent intolerant species (Figure 23 and 24).

The size of the nearest downstream lake limited both IBI and percent intolerant species. Streams that flow into lakes tend to have more habitat generalists and fewer fluvial specialists than free flowing streams (Herbert and Gelwick 2003, Guenther and Spacie 2006). Fluvial specialists are fishes that generally reside only in flowing-water habitats (Kinsolving and Bain 1993) and tend to also be species that are more intolerant of harsh physicochemical conditions (Herbert and Gelwick 2003). Declines in fluvial species upstream from lakes result from reductions in the amount and connectivity of fluvial habitats (Winston et al. 1991, Herbert and Gelwick 2003). Increases in generalist species above lakes is due to opportunistic movement of portions of lake fish populations upstream (Herbert and Gelwick 2003). It is logical that these effects would be more pronounced upstream from larger lakes, because larger lakes would result in greater reduction and fragmentation of fluvial habitats, and would provide for larger habitat generalist source populations.

Impervious surfaces limited both IBI and percent intolerant species. Impervious surfaces have a strong influence on fish communities (Allan 2004). In a study in southeast Wisconsin—within our study area—across broad gradients of both agricultural and urban land use, impervious surfaces were the best predictor of fish community indices, including IBI (Wang et al. 2001). Impervious surfaces reduce groundwater recharge and increase surface runoff, which results in more variable stream flow and temperature regimes, and increase the amount and variety of pollutants delivered to streams (Allan 2004).

Cattle density in the watershed also influenced both IBI and percent intolerant species. Cattle can impact stream habitat and fish communities at local scales by altering bank and riparian vegetation and degrading instream habitat through trampling (Lyons et al. 2000). Cumulatively, these local impacts can impact fish communities at watershed scales under high cattle densities. However, cattle are not as frequently found to be important in shaping fish community health at watershed scales—particularly in the Midwest (Rinne 1999). The fact that this variable emerged at a watershed scale indicates that more emphasis should be placed in understanding the mechanism of these impacts.

Fine end-moraine influenced IBI scores. The importance of geological features is not surprising. Across much of the Saginaw Bay watershed, Richards et al. (1996) found that surficial geology features were very important predictors of macroinvertebrate community structure. However, the patterns exhibited between IBI scores and fine end-moraine are not entirely clear and should be explored further. Further research is needed to better understand these patterns. Downstream link, or the size of stream downstream from a given reach, influenced the percentage of intolerant species. Downstream link has been known to have a strong influence on fish communities (Osborne and Wiley 1992) resulting from adventitious movement by fish from larger streams or rivers into tributaries (Gorman 1986). Downstream link has been known to influence IBI scores (Osborne et al. 1992). Groundwater index, or the percent of flow that is derived from groundwater sources, influenced the percent intolerant species. The importance of groundwater in influencing fish assemblages is well documented within the region (Zorn et al. 2002). Groundwater would also influence IBI scores, except that IBI scores are calculated differently for cold water streams (Lyons et al. 1996), which generally are streams with high groundwater contributions.

While the majority of stream reaches were not limited by individual natural and non-target threat variables (i.e., they did not fall under the wedge), across all variables there was an upper maximum limitation for 49% of stream reaches for IBI and 58% of reaches for percent intolerant. Natural variables tended to limit potential IBI and percent intolerant species at larger spatial scales than threat variables. For non-target threats specifically, 33% of stream reaches were limited for IBI and 8% were limited for percent intolerant species. The prevalence of these “background” limitations across the study area indicates how critical it was to analyze relationships between the fish community and water quality and flow variables with these natural and non-target threat variables as a filter.

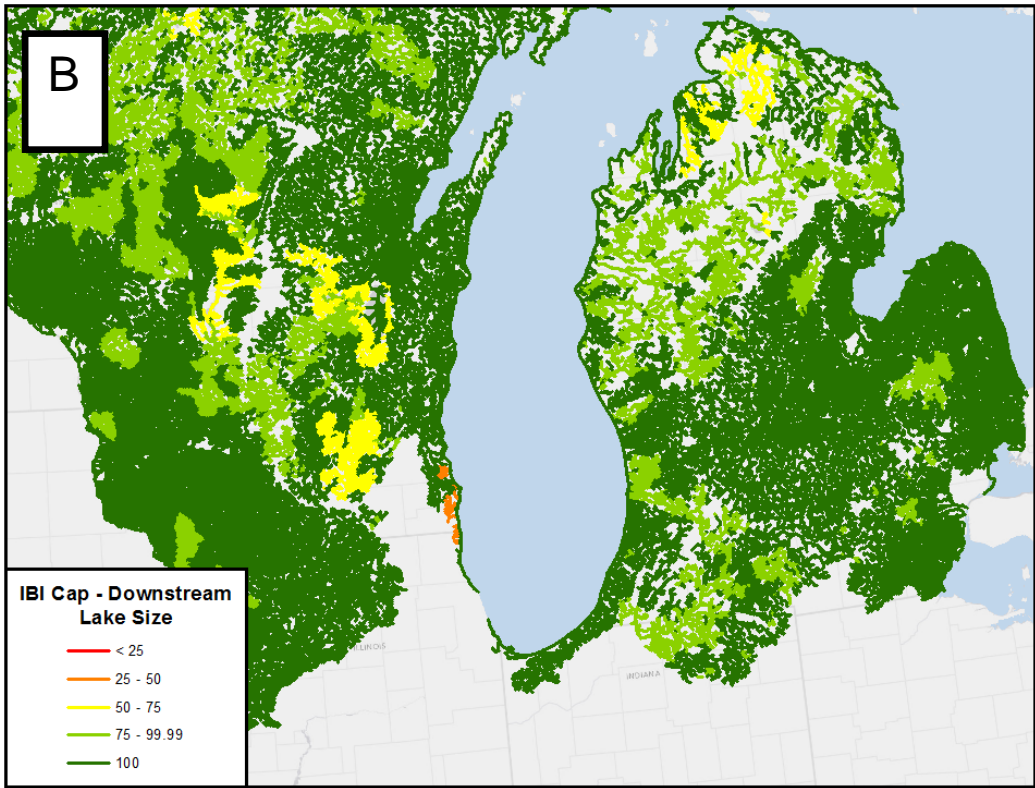
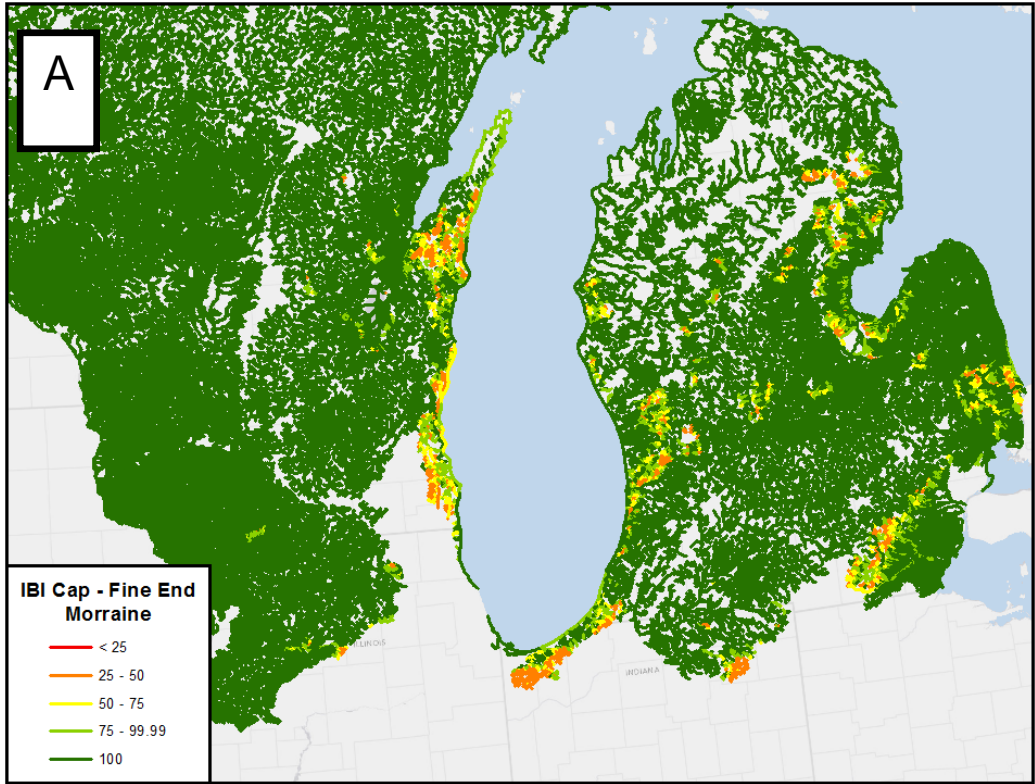


Figure 22. Maximum potential IBI scores based on wedge relationships between IBI and (A) Fine End-Moraine in the watershed and (B) size of nearest downstream lake.

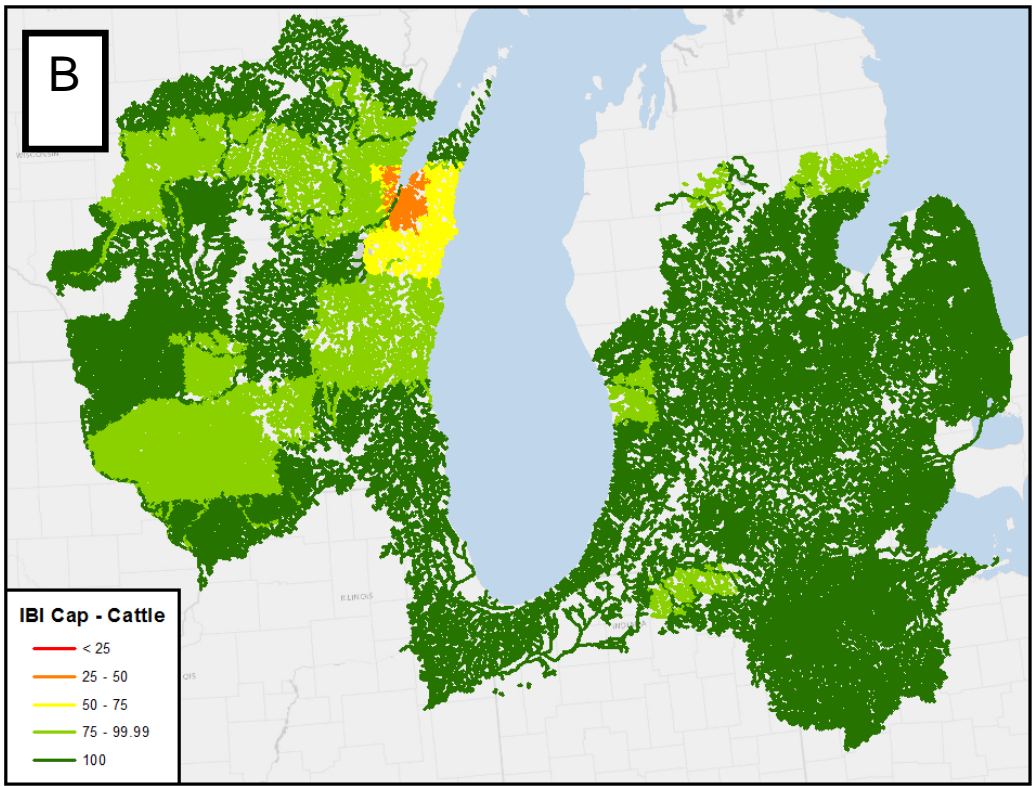
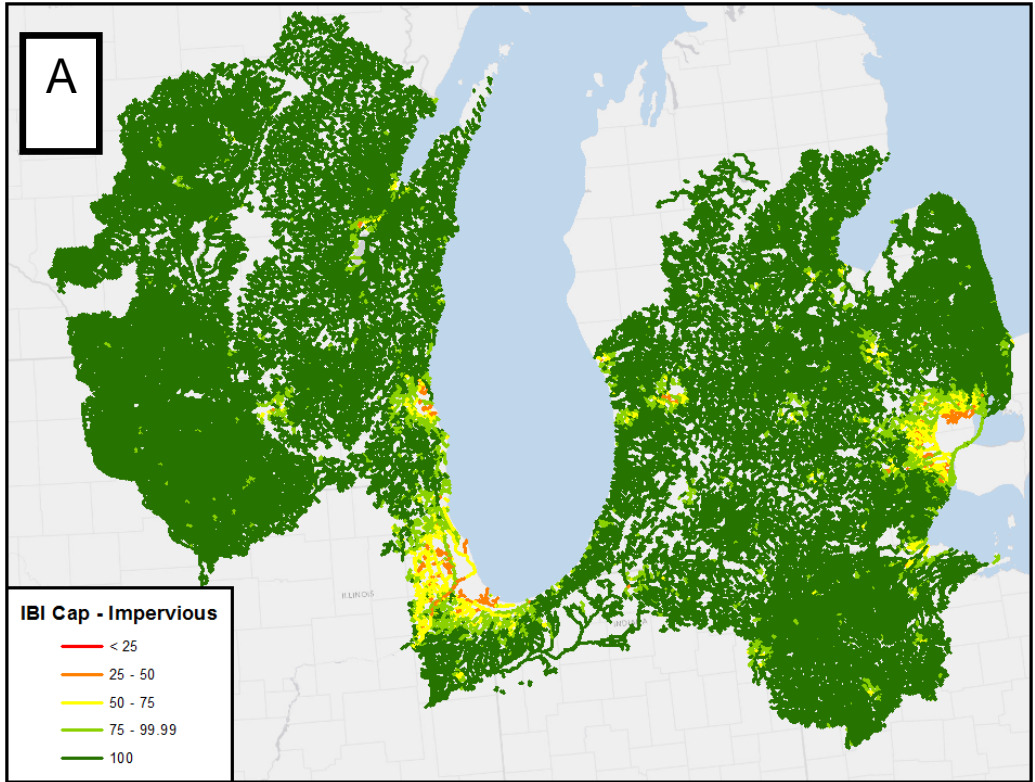


Figure 23. Maximum potential IBI scores based on wedge relationships between IBI and (A) percent impervious in the watershed and (B) average cattle density in the watershed.

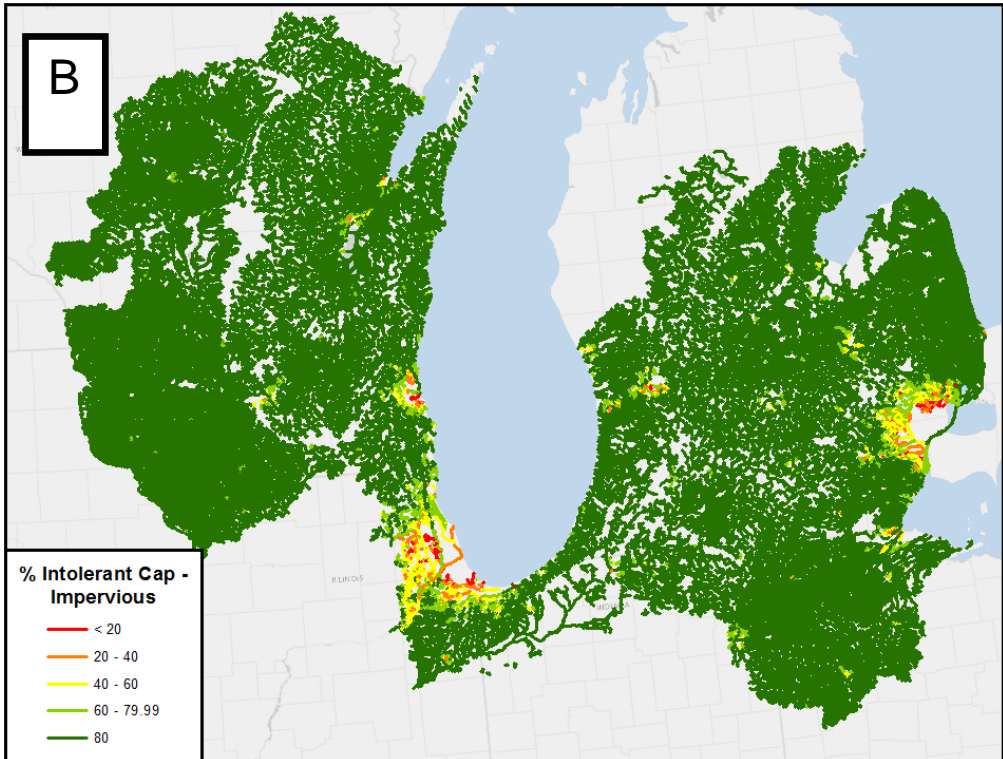
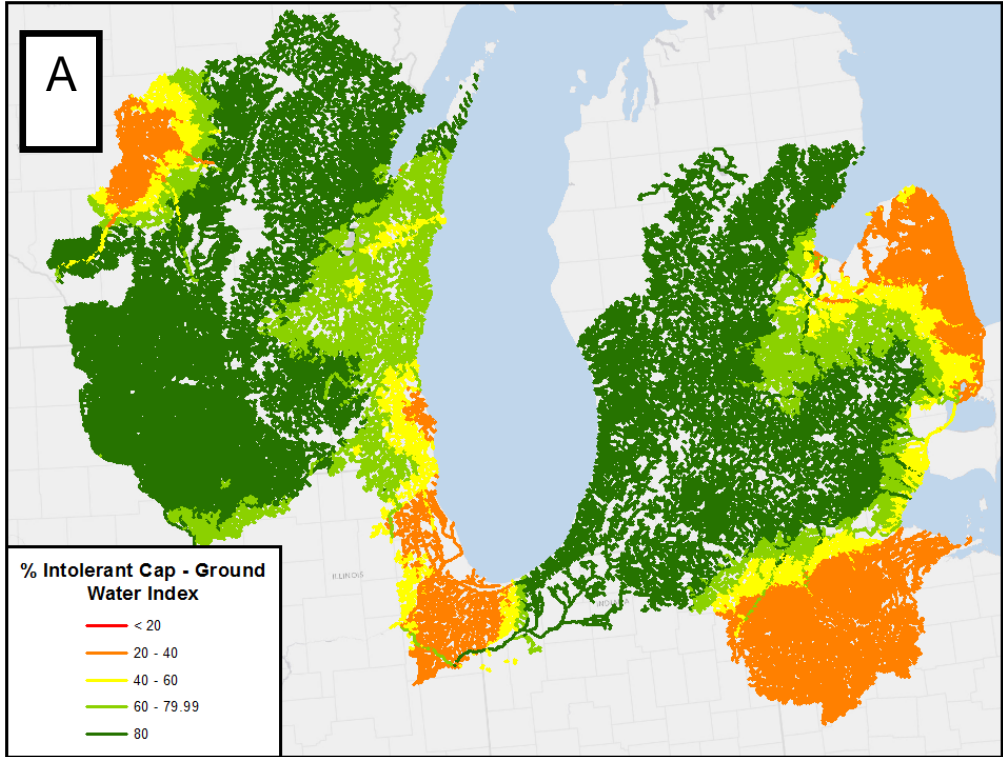


Figure 24. Maximum potential percent intolerant species based on wedge relationships between percent intolerant species and (A) watershed groundwater index value and (B) percent impervious in the watershed.

Water quality variables that exhibited threshold wedge relationships with IBI included local average annual surface runoff, local nitrate in surface runoff, summer sediment concentration, spring-rising organic phosphorus, spring-falling organic phosphorus, summer organic phosphorus, fall-winter organic phosphorus, and summer total phosphorus (see Figure 25 for examples). Water quality variables that exhibited wedge relationships with percent intolerant species include local average annual soluble phosphorus runoff, spring-falling organic phosphorus, spring-rising nitrate, summer total phosphorus, summer nitrate, summer ammonia, and fall-winter organic phosphorus (see Figure 26 for examples). Less agricultural areas in the northern portions of the study area had fewer water quality limitations, as did urban areas to some extent, where SWAT modeling was less effective in predicting water quality impacts and non-target threats dominated. However, these latter areas were largely captured in the capping for impervious surfaces.

This approach allowed us to evaluate restoration potential for water quality and flow variables constrained by limitations due to natural features and other threats. Threshold values water quality and flow variables, as well as natural and non-target threats, are shown in Table 4 for IBI and Table 5 for percent intolerant. To be clear, the wedge diagram approach is a conservative approach for threshold identification, and subsequently goal-setting, because it only identifies an upper maximum for each wedge variable and specific streams may be limited by a given variable prior to reaching that threshold, due to stream type or other local conditions. But with this conservative approach, we can be confident in the upper maximum predictions that resulted from our analyses.

Water quality or flow variables, the target variables, were generally most limiting for IBI across the agricultural dominated areas in the southern portions of the study area—especially in Michigan—and outside of urban areas with high impervious surfaces (Figure 27A). These trends were similar for percent intolerant species, except that Michigan's thumb was mostly limited by natural variables (Figure 27B).

Phosphorus variables were more frequently limiting for IBI across the study areas, except in eastern Wisconsin where nitrogen was more limiting and scattered headwater areas throughout the study areas where summer sediment concentrations or local surface runoff were most limiting (Figure 28). Spring-rising organic phosphorus was limiting at more than twice as many sites as any other water quality variable (Tables 6 and 7). Over half of stream reaches were most limited for IBI by water quality variables (Table 7), with the remaining reaches evenly divided between natural variables, non-target threats, and no variable limiting.

Limiting water quality variables were more balanced across phosphorus and nitrogen variables for percent intolerant species, and there is no clear pattern to discriminate where each tends to be limiting across the study area (Figure 29; Tables 6 and 7). Nearly half of stream reaches were most limited for percent intolerant species by water quality variables (Table 6), but most of the remaining sites (35%) were limited by natural variables and very few reaches were limited by

non-target threats (Table 7). The percentage of sites with no limiting variable was remarkably similar for IBI (16%) and percent intolerant species (15%).

Our results demonstrate the importance of considering natural and non-target threats when evaluating relationships between water quality and fish community indices/metrics. By identifying thresholds for natural and non-target threats, we were able to substantially reduce the number of reaches identified as most limiting by water quality variables from 15,564 (68%) to 11,245 (52%) for IBI and from 16,065 (75%) to 9899 (46%) for percent intolerant species. This is important because it reduces the area of focus for row-crop oriented conservation practices and ensures that the limited time and money spent implementing conservation practices will be focused in areas where it can result in improved biological communities. Still, when combining reaches most limiting for water quality and flow variables across both IBI and percent intolerant species, we see that most reaches are limited (Figure 30). But improvement capacity (Figure 31) can be used to further prioritize among reaches, by focusing on streams that can be substantially improved. In the next phase of this project, we will further prioritize by identifying locations where conservation practices can be reasonably expected to be able to result in meaningful improvements in the fish community.

Of course, reaches that are most limited by non-target threats should not be written off. Areas identified here as most limited for percent impervious surfaces or cattle should be targeted for conservation practices related to those threats.

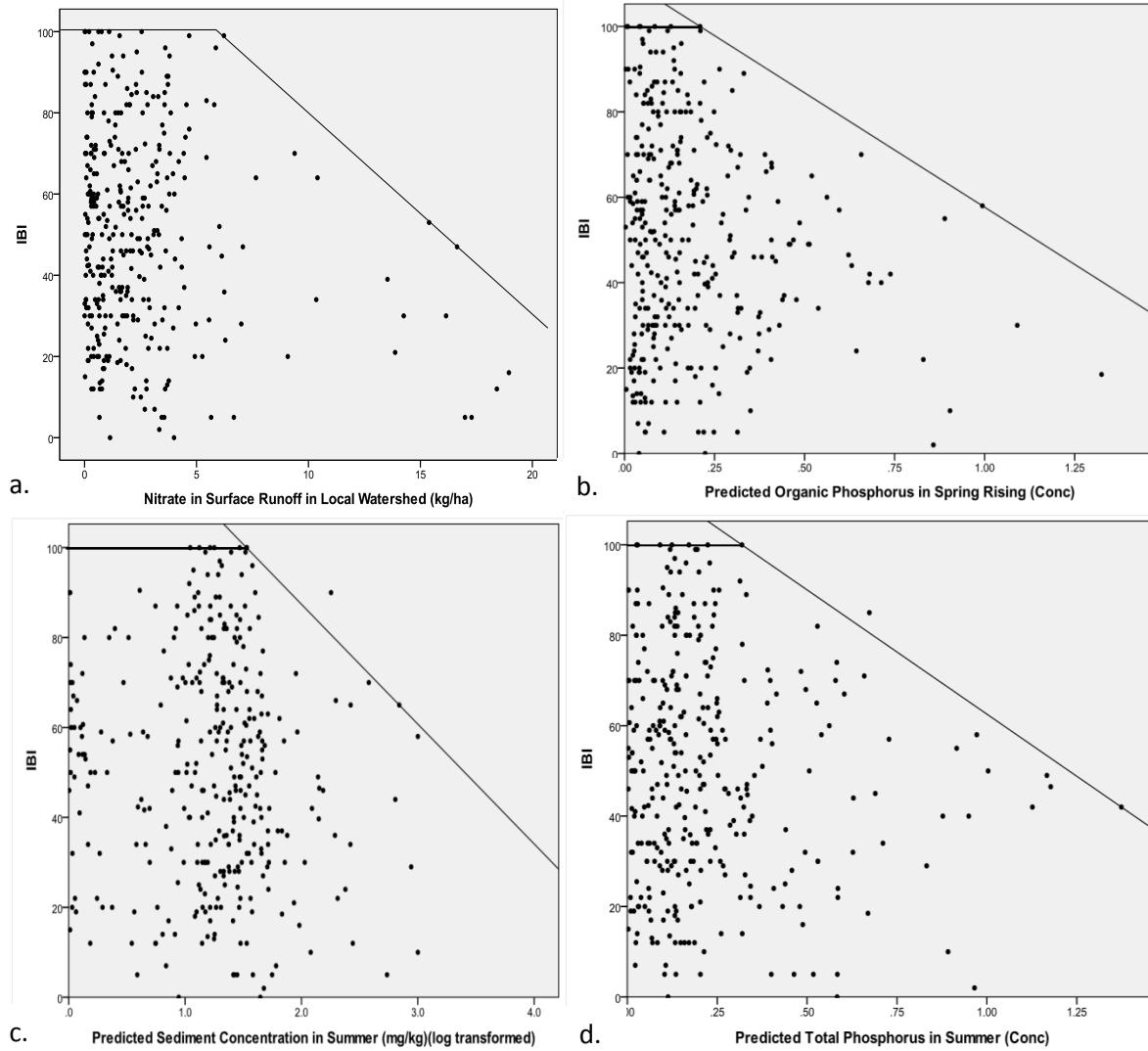


Figure 25: IBI to (a) predicted average annual surface flow (mm), (b.) predicted nitrate in surface runoff (kg/ha), (c.) predicted sediment concentration in summer (mg/kg)(log transformed), and (d.) predicted total phosphorus in summer (mg/L).

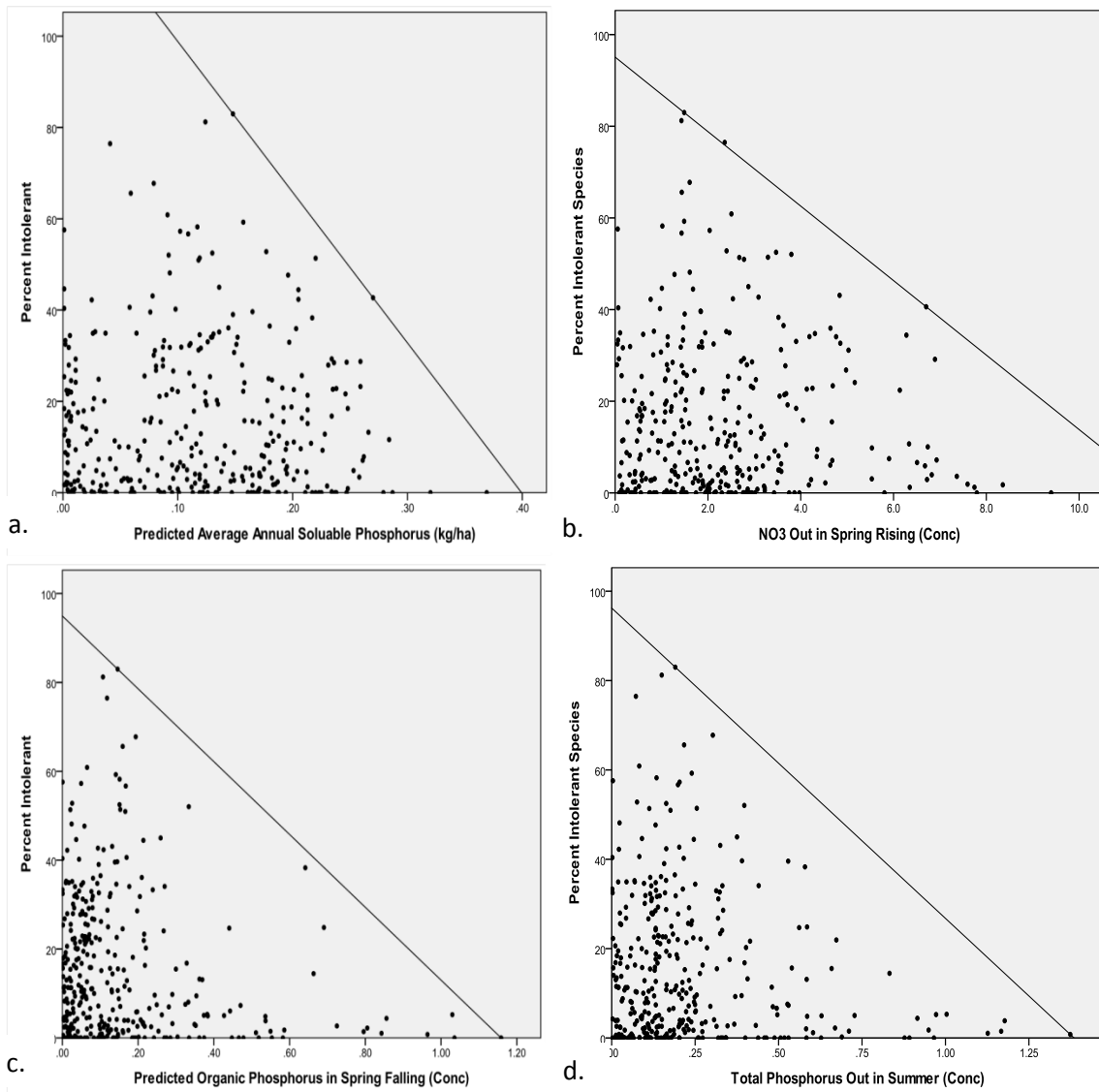


Figure 26: Wedges for percent intolerant to (a) predicted average annual soluble phosphorus (kg/ha), (b) predicted nitrate in spring rising (mg/L), (c) predicted organic phosphorus in spring falling (mg/L), and (d) predicted total phosphorus in summer(mg/L).

Table 4. IBI threshold values for natural, threat, and water quality variables. Thresholds represent the value for each variable above which IBI can no longer exceed 100, 80, 60, 40 or 20.

IBI Capping Variables	Threshold Levels				
	100	80	60	40	20
Natural					
Fine Moraine in watershed (%)	0.25%	0.50%	0.76%	1.01%	*1.27%
Downstream Lake Size (acres)	530	5726	61,747	665,744	*7,177,812
Other Threats					
Impervious in watershed (%)	8.3%	23.5%	38.1%	51.8%	*64.3%
Cattle Density on Farmland (# per 100 acre)	2169	3216	4263	5310	*6358
SWAT Variables					
Surface Runoff (kg/ha)	343	388	433	478	*523
Nitrate in Surface Runoff (kg/ha)	6.10	9.13	13.47	19.65	28.49
Summer Sediment Concentration (mg/l)	33	189	1065	6001	*33,775
Summer Total P (mg/l)	0.32	0.68	1.05	1.41	*1.77
Spring Rising Organic P (mg/l)	0.21	0.58	0.96	1.33	*1.70
Spring Falling Organic P (mg/l)	0.12	0.55	0.99	1.42	1.86
Summer Organic P (mg/l)	0.06	0.40	0.75	1.09	*1.43
Fall-Winter Organic P (mg/l)	0.12	0.41	0.78	*1.24	*1.83

*Estimates beyond data range, so values potentially inflated

Table 5. Percent Intolerant Species threshold values for natural, threat, and water quality variables. Thresholds represent the value for each variable above which the percent intolerant species can no longer exceed 80, 60, 40, or 20.

% Intolerant Capping Variables	Threshold Levels			
	80	60	40	20
Natural				
Downstream Lake (acres)	1345	18,234	247,033	*3,346,708
Downstream Link #	191	637	2122	7064
#Groundwater Index (%)	54.2%	46.5%	38.8%	31.0%
Other Threats				
Impervious in watershed (%)	14.6%	27.5%	39.8%	51.5%
Cattle Density on Farmland (# per 100 acre)	3084	3765	4445	5125
SWAT Variables				
Soluble P in Surface Runoff (kg/ha)	0.16	0.22	0.28	0.34
Spring Rising Nitrate (mg/l)	1.86	4.3	6.8	9.2
Summer Nitrate (mg/l)	1.46	8.0	14.5	21
Summer Ammonia (mg/l)	0.32	0.70	1.09	1.47
Spring Falling Organic P (mg/l)	0.18	0.43	0.67	0.91
Summer Total P (mg/l)	0.23	0.52	0.81	1.10
Fall-Winter Organic P (mg/l)	0.17	0.53	1.00	1.61

*Estimates beyond data range, so values potentially inflated

#For the Groundwater Index, the threshold represents the value below which the percent intolerant species is limited.

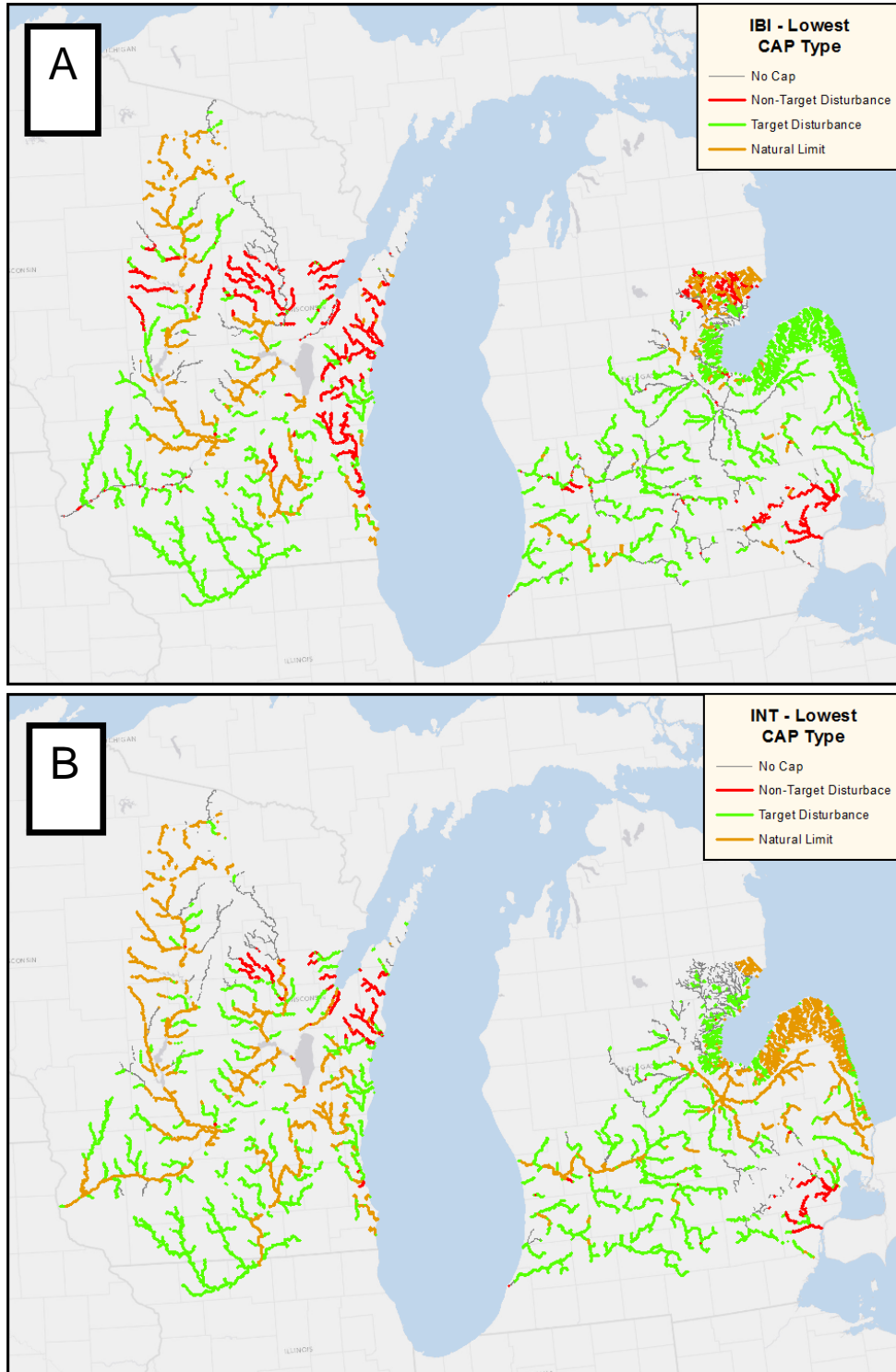


Figure 27: Lowest limiting variable group for each stream reach for (A) IBI and (B) percent intolerant species. Target disturbances are water quality and flow variables related to row crop agriculture. Non-target disturbances are anthropogenic threat variables unrelated to row-crop agriculture (e.g., impervious surfaces). Reaches with “no cap” were not limiting for any variable in our analyses.

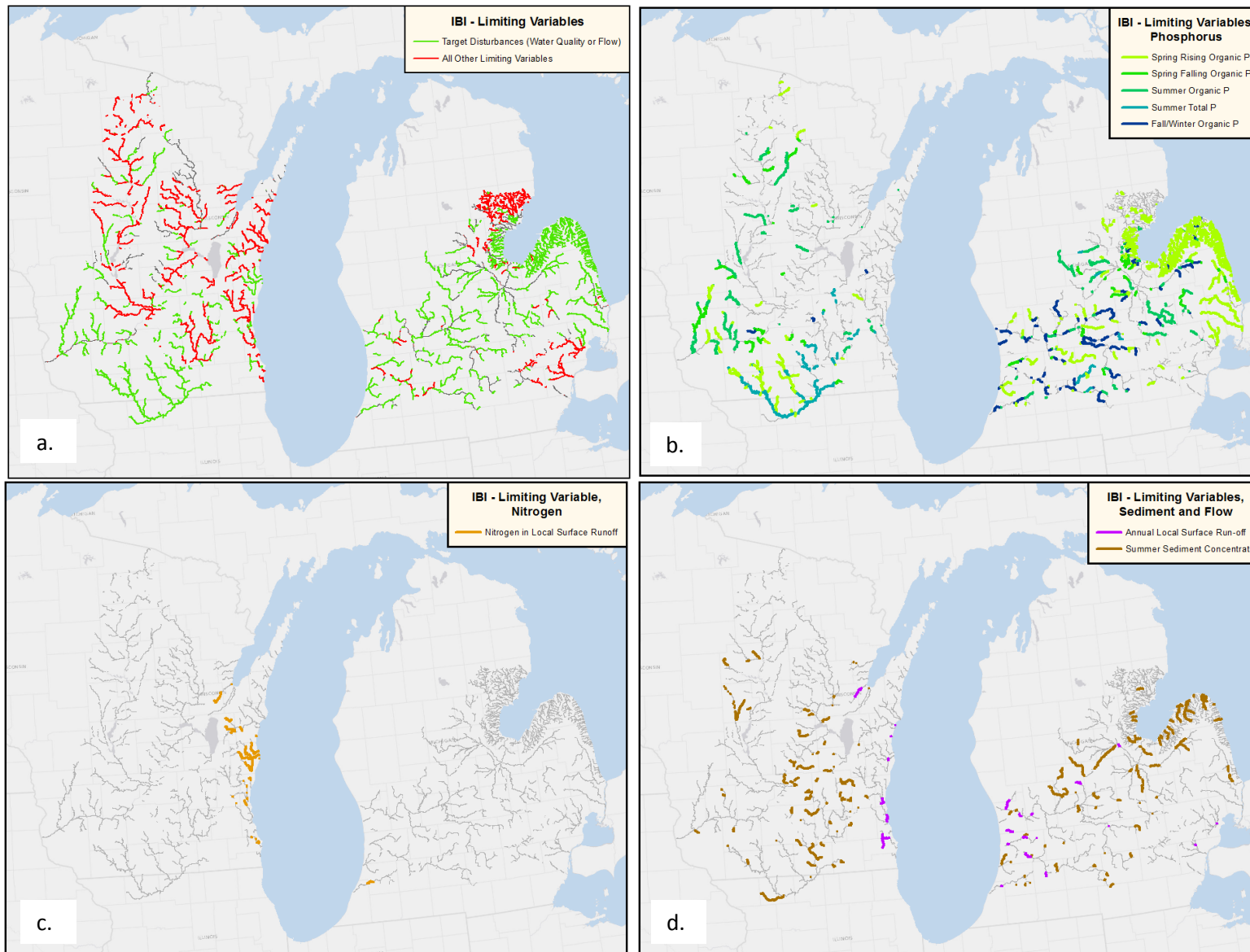


Figure 28: Stream reaches where IBI was limited by (a) any target disturbance (water quality or flow) variable, (b) various limiting phosphorus variables, (c) nitrogen in local surface runoff, and (d) local surface runoff and sediment concentration.

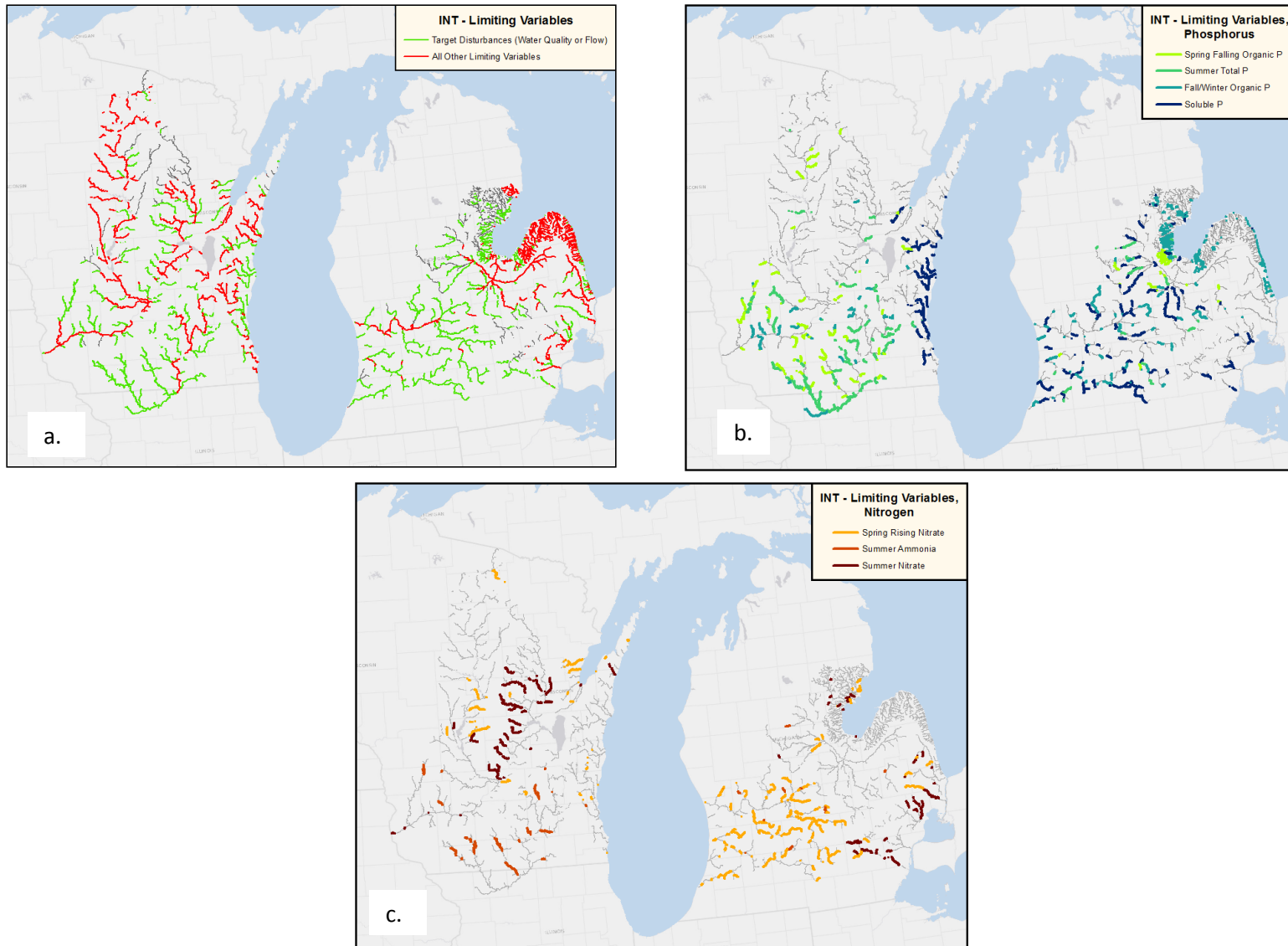


Figure 29: Stream reaches where percent intolerant species was limited by (a) any target disturbance (water quality or flow) variable, (b) various limiting phosphorus variables, and (c) various limiting nitrogen variables.

Table 6. Frequency that each wedge variable is the most limiting variable for a particular stream reach for IBI and Percent Intolerant species. Frequencies were calculated for water quality (target disturbance) variables only and across all wedge variables.

IBI Wedge Variables	Limiting Frequency		% Intolerant Wedge Variables	Limiting Frequency	
	# Reaches - Water Quality	# Reaches - All Variables		# Reaches - Water Quality	# Reaches - All Variables
Natural			Natural		
Fine Moraine	N/A	776	Downstream Lake	N/A	1622
Downstream Lake	N/A	3014	Downstream Link #	N/A	2062
Natural Subtotal	N/A	3790	Groundwater Index	N/A	3801
Other Threats			Natural Subtotal	N/A	7485
Impervious Surfaces	N/A	587	Other Threats		
Cattle Density	N/A	2506	Impervious Surfaces	N/A	283
Other Threat Subtotal	N/A	3093	Cattle Density	N/A	635
Water Quality			Other Threat Subtotal	N/A	918
Surface runoff	292	262	Water Quality		
NO3 in runoff	570	268	Soluble P in runoff	2705	1778
Summer sediment conc.	3098	1906	Spring rising NO3	3211	1619
Summer TP	1163	902	Summer NO3	1433	1264
Spring rising ORGP	4573	4333	Summer NH4	2393	965
Spring falling ORGP	1120	985	Spring falling ORGP	676	532
Summer ORGP	2618	1657	Summer TP	2481	1770
Fall-Winter ORGP	1130	932	Fall-Winter ORGP	3166	1971
Water Quality Subtotal	14,564	11,245	Water Quality Subtotal	16,065	9899
No Limiting Variable	6903	3339	No Limiting Variable	5402	3165

Table 7. Percentage of the time that each wedge variable is the most limiting variable across stream reaches for IBI and Percent Intolerant species. Percentages were calculated for water quality (target disturbance) variables only and across all wedge variables.

IBI Wedge Variables	Limiting Percentage		% Intolerant Wedge Variables	Limiting Percentage	
	% Reaches - Water Quality	% Reaches - All Variables		% Reaches - Water Quality	% Reaches - All Variables
Natural			Natural		
Fine Moraine	N/A	4%	Downstream Lake	N/A	8%
Downstream Lake	N/A	14%	Downstream Link #	N/A	10%
Natural Subtotal	N/A	18%	Groundwater Index	N/A	18%
Other Threat			Natural Subtotal	N/A	35%
Impervious Surfaces	N/A	3%	Other Threat		
Cattle Density	N/A	12%	Impervious Surfaces	N/A	1%
Other Threat Subtotal	N/A	14%	Cattle Density	N/A	3%
Water Quality			Other Threat Subtotal	N/A	4%
Surface runoff	1%	1%	Water Quality		
NO3 in runoff	3%	1%	Soluble P in runoff	13%	8%
Summer sediment conc.	14%	9%	Spring rising NO3	15%	8%
Summer TP	5%	4%	Summer NO3	7%	6%
Spring rising ORGP	21%	20%	Summer NH4	11%	4%
Spring falling ORGP	5%	5%	Spring falling ORGP	3%	2%
Summer ORGP	12%	8%	Summer TP	12%	8%
Fall-Winter ORGP	5%	4%	Fall-Winter ORGP	15%	9%
Water Quality Subtotal	68%	52%	Water Quality Subtotal	75%	46%
No Limiting Variable	32%	16%	No Limiting Variable	25%	15%

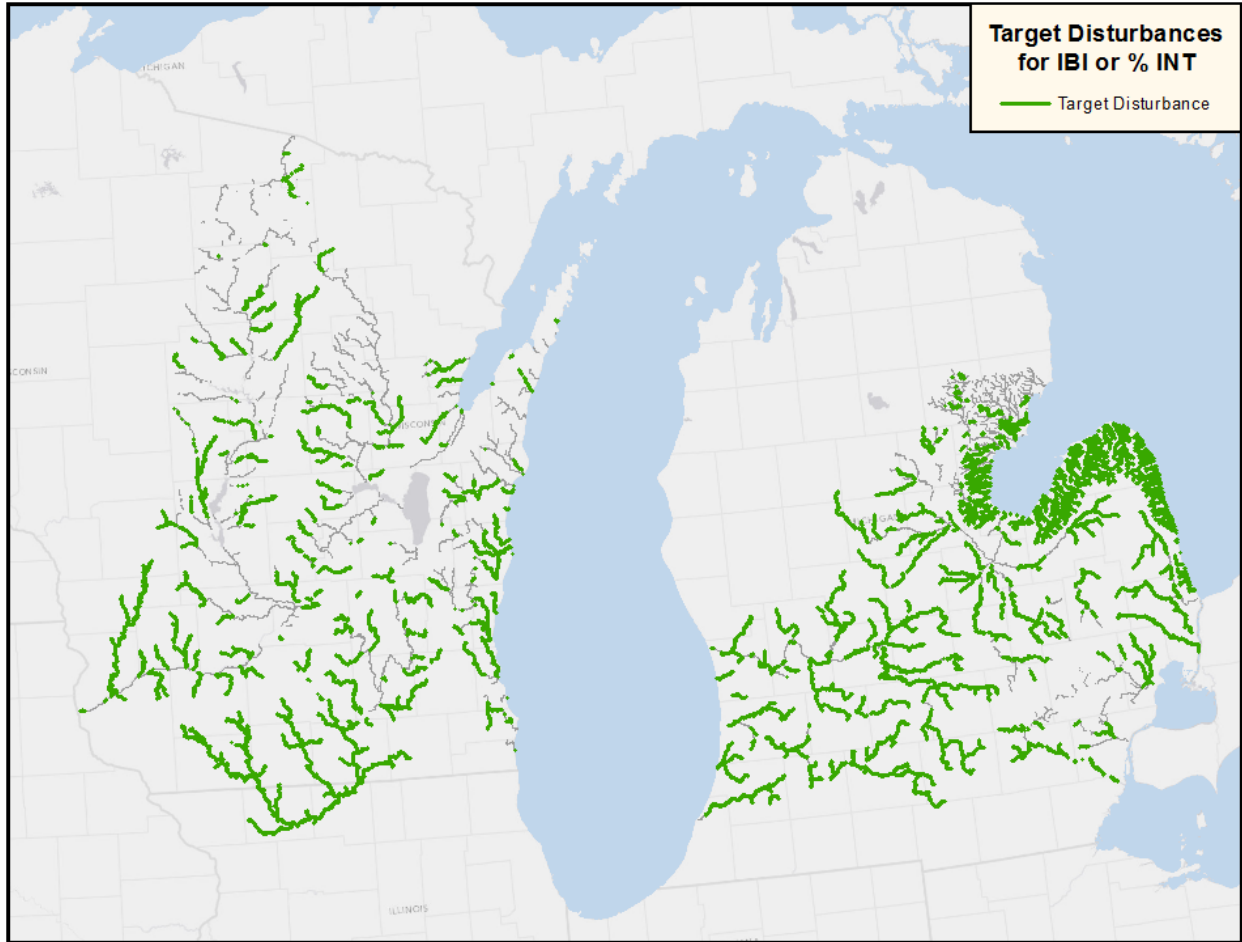


Figure 30: Stream reaches that are limited by any target disturbance (water quality or flow variable) for either IBI or percent intolerant species, or both.

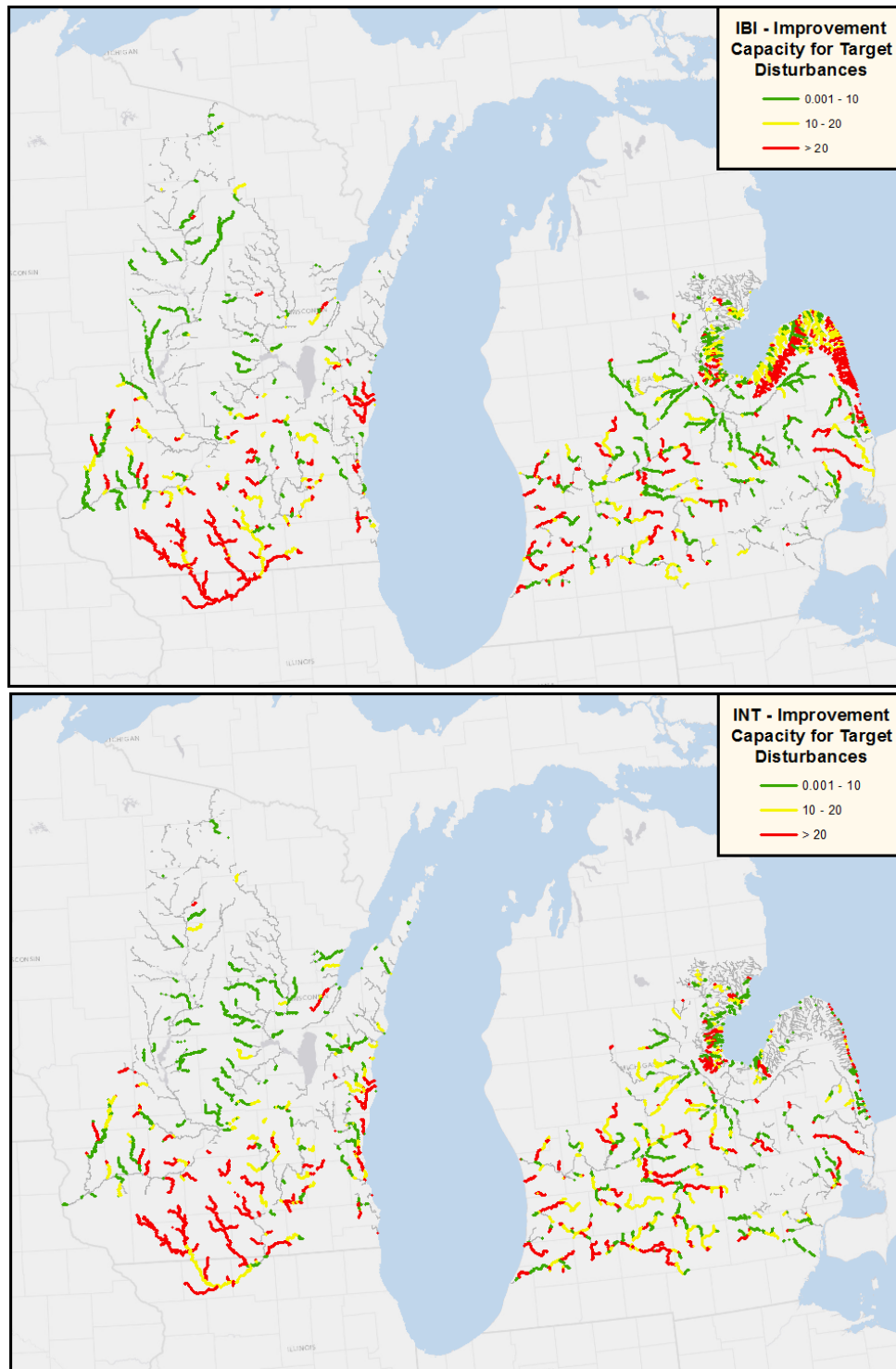


Figure 31: The improvement capacity for each stream reach for IBI or percent intolerant species. Improvement capacity is how much improvement is possible before reaching the natural limit for IBI (100) or percent intolerant species (80) or a limitation set by a wedge cap for a natural or non-target disturbance (threat) variable. Sites with no improvement capacity either had no limiting variable or were more limited by a natural variable or non-target disturbance.

In addition to scatter and wedge plots, we also developed a suite of tri-plots showing fish community metrics against both a) predicted historic water quality conditions and b) percent change from historic conditions. Preliminary examinations of some of these analyses suggest that there are a small subset of streams that should be expected to have relatively low values for IBI and percent intolerant fish species in the community even in relatively pristine conditions (Figure 32). These tri-plots also suggest that the deviation from historic conditions is possibly as much or more important than the actual current conditions, but only when placed within the proper context of the inherent potential of the site. These results are consistent with ecological theory that suggests that there is an inherent biological potential of each stream and that current biological conditions should reflect that potential (Frissel et al. 1986).

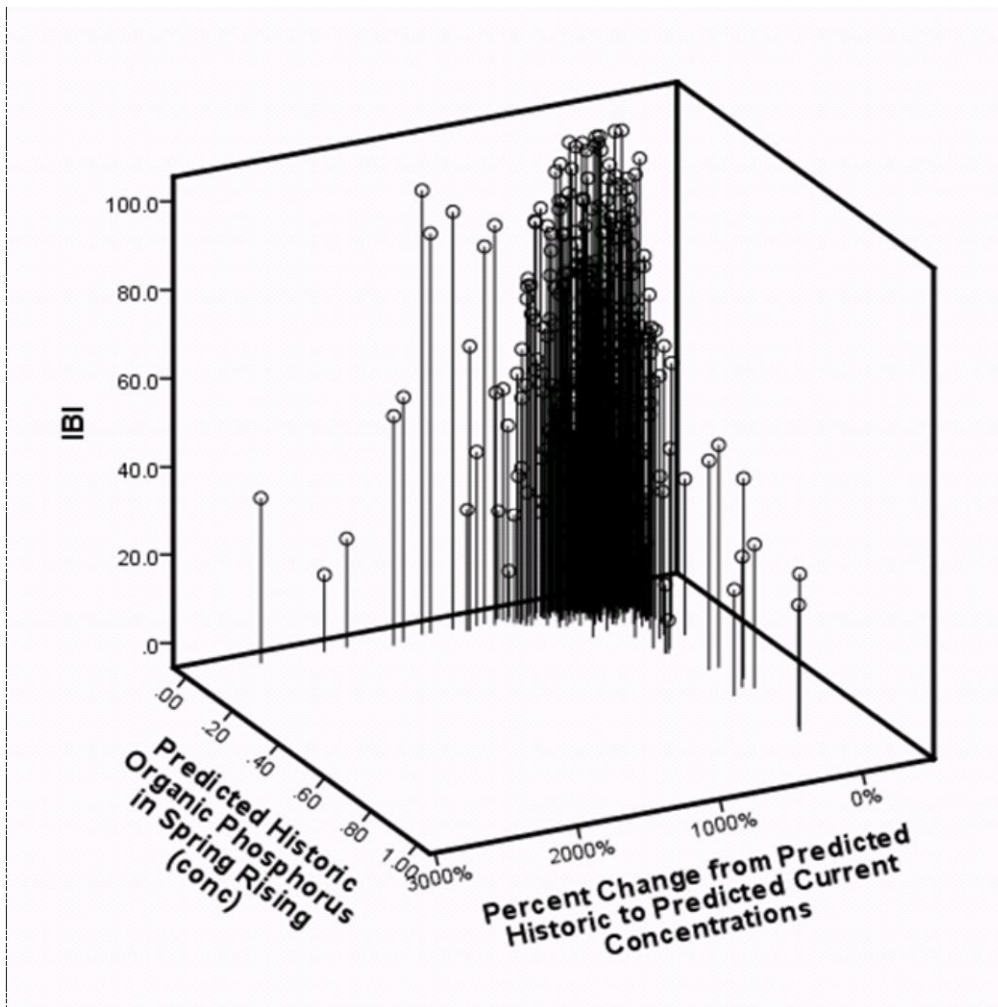


Figure 32: Tri-plot showing the relation of current IBI scores to predicted historic spring rising organic phosphorus concentrations (μl) and percent change from predicted historic to predicted current concentrations of the same parameter.

OVERALL DISCUSSION

Project Benefits

Our project successfully demonstrated that you can develop fine resolution SWAT model predictions across a large geographic area and quantitatively link the resulting water quality and flow measures to fish community indicators to generate spatially explicit predictions. Our ability to, in essence, extend the predictive capabilities of SWAT to biological endpoints and also incorporate constraints not addressed by SWAT or NRCS conservation practices allowed us to begin developing more realistic expectations to guide strategic conservation across the project area. This will help us to achieve our objectives in Phase 2 of the Great Lakes which is seeking to develop realistic goals (expectations) for fish community conditions in priority subwatersheds of the project area and working with partners to develop detailed strategies for achieving those goals.

Demonstrating the ability to predict fish community metrics from SWAT model outputs has the potential to significantly advance strategic conservation in the Great Lakes and beyond. Our results consistently demonstrated the importance of seasonal water quality and flow parameters, particularly the spring rising period, rather than average annual conditions, which are more typically available and thus used by scientists to elucidate relations of these parameters to biological endpoints. This result alone demonstrates an important benefit of SWAT, which can generate data at a variety of time steps, for advancing our understanding of the complex relations between biological endpoints and instream habitat conditions. Results like ours can also help guide conservation actions to further focus on critical periods, like early spring, to reduce runoff and associated sediment and nutrient inputs.

Another benefit of SWAT, as demonstrated by our project, is that it has the potential to be used to develop spatially comprehensive data and predictions at a fine spatial grain across a large project area and model. This ability provides benefits for both science and conservation planning. From a science perspective, the SWAT model predictions allowed us to fill gaps in water quality and flow data at locations with biological samples. In our study only a small fraction of original 1022 fish community sampling locations had existing water quality and flow data. While we were only able to link SWAT model outputs to 345 of these sites, it must be noted that most of these sites also lacked water quality data. And, the data that is available is certainly far from the consistent and comparable data we had for hundreds of parameters. However, to truly realize this benefit we must make it a priority to evaluate and improve the accuracy of hydrologic models, like SWAT, particularly as it applies to downscaling such models to finer spatial grains and making predictions beyond the gage stations used for calibration. The detailed and spatially comprehensive data provided by SWAT and the other predictors allowed us to assess and map likely fish community conditions and thresholds beyond sampled locations. Our models and maps exhibited extreme spatial heterogeneity in biological expectations under both current and historic conditions. This finding suggests that we should not

hold all streams to the same standard even within a relatively small watershed or region, which is somewhat contrary to certain methods used to establish goals for fish community endpoints in streams.

Equally important to the temporal and spatial issues described above is the fact that the SWAT model also allows you to assess past, present, and potential future conditions based on different land use, land cover and management scenarios. The demand for demonstrating the benefits of conservation, particularly to biological endpoints, has increased sharply in recent years. Monitoring program and the associated retrospective analyses are useful for addressing this demand. However, we argue that equally important to these retrospective assessments are modeling efforts that forecast the likely benefits of conservation. The ability of SWAT to forecast future instream habitat and biological conditions based on different amounts and configurations of agricultural BMPs is very appealing for conservation planning. These management scenarios provide a means of developing management alternatives needed for developing truly realistic desired conditions by allowing decision makers to simultaneously evaluate ecological benefits relative to funding needs and constraints and potentially other socioeconomic costs in terms of agricultural production, farm income, and other valued services. As stated earlier, having the ability to extend such forecasts to biological endpoints, like fish communities, provides organizations like The Nature Conservancy the ability to identify where we can make meaningful improvements in freshwater biodiversity and help secure the necessary resources and attention needed to bring about those improvements.

The SWAT modeling was focused on watershed and subwatershed scale water quality and flow relationships. Some stream reaches will be more sensitive to these water quality and flow impacts (e.g., depositional areas) and therefore may require more stringent thresholds. Other stream reaches will be more resilient. Further, the wedge approach for threshold identification is only identifying a fundamental limitation beyond which stream reaches will not attain. But many stream reaches will be affected by the limiting variable prior to reaching the threshold. Therefore, the thresholds identified here should be considered highly conservative.

Limitations and Opportunities for Improvement

Despite all of the realized and potential benefits of our project we must also be mindful of its limitations and opportunities to build upon this work and improve our ability to develop realistic expectations for biological endpoints and strategies to achieve them. Similar to previous studies (Rankin et al. 1999; Wang et al. 2007; 2008), our analyses revealed relatively good threshold relations between fish community metrics and several water quality and flow variables.

However, these preliminary RDAs and CART analyses for Phase 1 did not explain as much of the variation in fish community metrics as other efforts (Rankin et al. 1999; Wang et al. 2007; 2008; Annis et al. 2009). In fact, our analyses thus far have only explained about half (~20%) of the variance reported by these and other studies examining similar suites of predictor and response variables. These lower values could be the result of many factors related to the original

source data, transformations, or analyses that were discussed earlier in the report. However, given the potential benefits of our approach for advancing strategic conservation we again want to stress the importance of taking steps to improve the accuracy of such predictions in the future. Therefore offer several suggestions on how this might be accomplished in similar projects in the future.

1. Further downscaling of SWAT models

There is an immense number of ecological factors that collectively determine the distribution and abundance of fish and other freshwater taxa. Identifying significant relations within this realm of complexity demands an extremely large sample size for both predictor and biological response variables. This is particularly true whenever you are trying to isolate the influence of a particular subset of variables, like water quality and flow variables, as we were for Objective 2 of this project. Unfortunately, we were unable to use nearly 70% of the original 1022 fish community samples that we had compiled for this project because we were literally pushing the limits of technology for SWAT modeling at the time. We firmly believe we would have been able to explain significantly more variation in fish community metrics and develop more accurate predictive models if we had been able to use all of those 1022 samples. What prevented us from using those data was our inability to further downscale the SWAT model and generate model outputs for every single stream segment containing a fish community sample. So, we suggest every effort must be made, regionally and nationally, to develop finer resolution SWAT models.

Fortunately, in just two years since our project began, the rapid advancements in computing power combined with technical advancements in the SWAT model algorithms that have reduced computer processing and memory demands, those technical limitations that hindered our project have been eliminated (Jeff Arnold, personal communication). In fact, the CEAP Cropland Modeling team is working on the development and calibration of a national SWAT model that will provide predictions for all of the individual reaches contained within a slightly modified version of the NHD-Plus. The development of these downscaled SWAT predictions and the associated processing capabilities holds significant promise for improving the accuracy of models like ours where once again the sample size is so critical to providing the statistical power needed to collectively assess the complex array of variables that influence local biological assemblages.

2. Fill critical data gaps for certain predictor variables

We had a large number of predictor variables for our study, yet there is still significant variation in fish communities that our models could not explain. For instance, our project did not include data for drainage tiles, which occur extensively throughout much of the project area and have a significant influence on hydrology and water quality. Having and incorporating accurate geospatial on these and other critical factors for which we currently lack good data would likely help improve the SWAT models and the associated biological models.

3. Incorporate spatial statistics to account for neighborhood effects

Because organisms are mobile and utilize resources at different spatiotemporal scales, local biological assemblages may not always be a reflection of the local stream habitat (Schlosser 1991; Rabeni and Sowa 1996; Fausch et al. 2002). Local assemblages may actually be more reflective of stream habitat conditions occurring upstream or downstream, or reflect the average conditions found within a much longer stretch of stream (Schlosser 1991; Rabeni and Sowa 1996; Cooper and Mangel 1999; Fausch et al. 2002). Spatial statistics incorporates locational attributes as potential explanatory variables which can help address the influence of more complex features like patches (e.g., distinct geographic neighborhoods) on the distribution and abundance of biota (Legendre 1990; Borcard et al. 1992; Anderson and Gribble 1998).

4. Make a concerted effort to improve the accuracy of downscaled SWAT models

Obviously we should expect better relations between the fish community metrics and observed water quality and flow data than water quality and flow data based on SWAT predictions. However, because of the many potential benefits of SWAT for advancing strategic conservation we believe we must make it a priority to improve the accuracy of downscaled SWAT models and we believe there are many options for such improvements. Incorporating spatially extensive, but temporally discrete (e.g., average annual nutrient concentrations) water quality data into the SWAT model calibration process. A limitation of the SWAT modeling process used in our project, and most SWAT modeling projects, is that the model is calibrated to one or a few gage stations within the watershed. Incorporating additional calibration sites would help account for the spatial heterogeneity in water quality and flow conditions that consistently occur across large regions and are not fully accounted for by existing equations like RUSLE. Another option for improving the accuracy of downscaled SWAT models would be to follow the methods used in regional assessments by the Cropland Component of CEAP, which uses the farm survey data from Natural Resource Inventory (NRI) to better account for existing conservation practices and also APEX models to better model field scale hydrologic conditions (USDA 2011).

5. Use complimentary sets of models and water quality and flow data

All data and models have strengths and weaknesses. We have talked extensively about the strengths of SWAT, particularly its ability to be calibrated and offer predictions at a daily or any other larger time step. The results of our study, where seasonal variables consistently revealed the strongest relations to fish community metrics, clearly show the benefit of this temporally intensive calibration. SWAT was not originally designed for predictions at fine spatial scales, like we developed for our project. However, there are other models, like SPARROW, that were developed for this very purpose, yet suffer from the inability to provide detailed time step predictions (<http://water.usgs.gov/nawqa/sparrow/>). So, the strength of SWAT is the weakness of SPARROW and vice versa. We believe that integrating the strengths of these two models to produce water quality and flow predictor variables could significantly improve our ability to predict biological endpoints. Further supplementing these predictors with actual field measurements of certain water quality and flow variables could offer additional benefits.

LITERATURE CITED

- Akers, J. 1938. Drift thickness map of Lower Michigan. Lansing, MI: Michigan Geological Survey Division. 1 map (1:500,000).
- Albert, D.A., S.R. Denton and B.V. Barnes. 1986. Regional Landscape Ecosystems of Michigan. School of Natural Resources, University of Michigan 48109-115. 32 pp.
- Albert, D.A. 1994. Ecosystem Management: Saginaw Bay. Congressional Research Service of the Library of Congress, Symposium on the Federal Role in Ecosystem Management, Washington, D. C., pp. 147-151. Michigan Natural Features Inventory report number 1994-01.
- Allan, J.D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. *Annual Review of Ecology and Systematics* 35:257–84.
- Austen, D.J., Bayley, P.B., Menzel, B.W. 1994. Importance of the Guild Concept to Fisheries Research and Management. *Fisheries*. Volume 19, Issue 6, p 12-20.
- Berkman, H., C. Rabeni, T. Boyle. 1986. Biomonitoring of stream quality in agricultural areas: Fish versus invertebrates. *Earth and Environmental Science*. Volume 10, Number 3, p 413-419.
- Biggs, D., B. de Ville, E. Suen. 1991. A method of choosing multi-way partitions for classification and decision trees. *Journal of Applied Statistics*. 18, p 49-62.
- Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. *Classification and Regression Trees*. Wadsworth International, Belmont, Ca.
- Brenden, T.O., L. Wang, and Z. Shu. 2008. Quantitative identification of disturbance thresholds in support of aquatic resource management. *Environmental Management* 42:821-832.
- Comer, P.J., D.A. Albert, L.J. Scrimger, T. Leibfreid, D. Schuen, and H. Jones. 1993a. Historical wetlands of Michigan's coastal zone and southeastern lakeplain. Report to the Michigan Department of Natural Resources, Land and Water Management Division. Lansing, MI: Michigan Natural Features Inventory. 105 pp.
- Comer, P.J., D.A. Albert, T. Leibfreid, H. Wells, B.L Hart, and M.B. Austin. 1993b. Historical wetlands of the Saginaw Bay watershed. Report to the Saginaw Bay Watershed Initiative, Michigan Department of Natural Resources, Office of Policy and Program Development. Michigan Natural Features Inventory, Lansing, MI. 68 pp
- De'ath, G., and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for the analysis of complex ecological data. *Ecology* 81:3178–319.
- Denton, S. R. 1985. Ecological Climatic Regions and Tree Distributions in Michigan. Ph.D. Dissertation., University of Michigan, Ann Arbor, Michigan. 390 pp. 55
- Dorr J.A., Jr., and D.F. Eschman. 1984. *Geology of Michigan*. Univ. of Michigan Press, Ann Arbor. 476 pp.

- Eichenlaub V.L., J.R. Harman, F.V. Nurnberger, and H.J. Stolle. 1990. The Climatic Atlas of Michigan. University of Notre Dame Press, Notre Dame, Indiana. 165 pp.
- Eichenlaub, V. L. 1979. Weather and the climate of the Great Lakes region. The University of Notre Dame Press, Notre Dame, Indiana. 333 pp.
- Esselman, P., D. Infante, L. Wang, D. Wu, R. Cooper, and W. Taylor. 2011. An Index of Cumulative Disturbance to River Fish Habitats of the Conterminous United States from Landscape Anthropogenic Activities. *Ecological Restoration*. 29: 1-2, p133-151.
- Fitzpatrick, F.A., B.C. Scudder, B.N. Lenz, and D.J. Sullivan, 2001. Effects of Multi-Scale Environmental Characteristics on Agricultural Stream Biota in Eastern Wisconsin. *Journal of the American Water Resources Association (JAWRA)* 37(6):1489-150.
- Gorman, O.T. 1986. Assemblage organization of stream fishes: the effect of rivers on adventitious streams. *American Naturalist* 128:611–616.
- Guenther, C.B. and A. Spacie. 2006. Changes in Fish Assemblage Structure Upstream of Impoundments within the Upper Wabash River Basin, Indiana. *Transactions of the American Fisheries Society* 135:570-583.
- Hamilton, D. and P. Seelbach. 2010. Determining environmental Limits to Streamflow Depletion across Michigan. *The Book of the States 2010*. The Council of State Governments. p534-537.
- Herbert, M.E. and F.P. Gelwick. 2003. Spatial variation of headwater fish assemblages explained by hydrologic variability and upstream effects of impoundment. *Copeia* 2003:273-284.
- Hole, F.D, and C.E. Germain. 1994. Natural divisions of Wisconsin. Wisconsin Department of Natural Resources, Madison , WI. 1 map (1:1,000,000) and accompanying text.
- Huston, M. A. 2002. Critical issues for improving predictions. Pages 7-24 in *Predicting Species Occurrences. Issues of Accuracy and Scale*. J. M Scott, P. J. Heglund, J. B. Haufler, M. Morrison, M. G. Raphael, W. B. Wall, and F. Samson, eds. Island Press, Covelo, CA.
- Kinsolving, A.D. and M.B. Bain. 1993. Fish assemblage recovery along a riverine disturbance gradient. *Ecological Applications* 3:531–544.
- Lyons, J., L. Wang, and T.D. Simonson. 1996. Development and validation of an Index of Biotic Integrity for Coldwater Streams in Wisconsin. *North American Journal of Fisheries Management* 16:231-256.
- Lyons, J., B.M. Weigel, L.K. Paine, and D.J. Undersander. 2000. Influence of intensive rotational grazing on bank erosion, fish habitat quality, and fish communities in southwestern Wisconsin trout streams. *Journal of Soil and Water Conservation* 55: 271-276.
- Merritt, R., and K. Cummins. 1996. *An Introduction to the Aquatic Insects of North America* 3rd Edition. Kendall/Hunt Publishing Company. Dubuque, IA.

- Mitsch, W. and J. Day Jr. 2006. Restoration of wetlands in the Mississippi–Ohio–Missouri (MOM) River Basin: Experience and needed research. *Ecological Engineering* 26 (2006) p55-69.
- Morey G7.B., P.K. Sims, W.R. Cannon, M.G. Mudrey, Jr., and D. L. Southwick. 1982. Geological Map of the Lake Superior Region; Minnesota, Wisconsin, and Northern Michigan. Map s-13. Minnesota Geological Survey, University of Minnesota, St. Paul. 1 map (1:1,000,000).
- O'Connor, R. J. 2002. The conceptual basis of species distribution modeling: Time for a paradigm shift? Pages 25-34 in *Predicting Species Occurrences. Issues of Accuracy and Scale*. J. M Scott, P. J. Heglund, J. B. Haufler, M. Morrison, M. G. Raphael, W. B. Wall, and F. Samson, eds. Island Press, Covelo, CA. 56
- Olden, J.D., and D.A. Jackson. 2002. A comparison of statistical approaches for modeling fish species distributions. *Freshwater Biology* 47:1976-1995.
- Omerik, J.M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). *Annals of the Association of American Geographers* 77(1):118-125.
- Osborn, L.L., S.L. Kohler, P.B. Bayley, D.M. Day, W.A. Bertrand, M.J. Wiley, and R. Sauer. Influence of stream location in a drainage network on the Index of Biotic Integrity. *Transactions of the American Fisheries Society* 121:635-643.
- Osborne, L.L., and M.J. Wiley. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. *Canadian Journal of Fisheries and Aquatic Sciences* 49:671–681.
- Ostrom, M.E. 1981. Bedrock geology of Wisconsin. Wisconsin Geologic and Natural History Survey, University of Wisconsin, Madison, WI. 1 map
- Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, D.C. EPA 440-4-89-001.
- Rabeni, C. F., and S. P. Sowa. 1996. Integrating biological realism into habitat restoration and conservation strategies for small streams. *Canadian Journal of Fisheries and Aquatic Sciences* 53:252–25
- Richards, C., L.B. Johnson, and G.E. Host, 1996. Landscape-Scale Influences on Stream Habitats and Biota. *Canadian Journal of Fisheries and Aquatic Sciences* (Suppl. 1), pp. 295-311.
- Richter, B.D., D.P. Braun, M.A. Mendelson, and L.L. Master. Threat to imperiled freshwater fauna. *Conservation Biology* 11:1081-1093.
- Rinne, J.N. 1999. Fish and grazing relationships: the facts and some pleas. *Fisheries* 24:12-21.
- Trotta, L.C., and R.D. Cotter. 1973. Depth to bedrock in Wisconsin. Map (scale 1:1,000,000). University of Wisconsin Extension, Geological and Natural History Survey in cooperation with the U.S. Geological Survey, Madison, Wisconsin.

USDA Soil Conservation Service. 1967. Distribution of Principal Kinds of Soils: Orders, Suborders, and Great Groups.' 1 map (1:7,500,000).

Wang, Z., A.C. Chang, L. Wu, D. Crowley. 2003. Assessing the soil quality of long-term reclaimed wastewater-irrigated cropland. *Geoderma* 114: p261-278.

Wang L, J. Lyons, and P. Kanehl. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. *Environmental Management* 28:255–66.

Wendland, W.M., K.E. Kunkel, G. Conner, W.L. Decker, H. Hillaker, P. Naber-Knox, F.V. Nurnberger, J. Rogers, K. Scheeringa, and J. Zandlo. 1992. Mean 1961-1990 Temperature and precipitation over the Upper Midwest. Res. Rep. 92-01, SWS Misc. Publ. 136. Champaign, IL: Midwest Climate Center. 27 p.

Winston, M.R., C.M. Taylor, and J. Pigg. 1991. Upstream extirpation of four minnow species due to damming of a prairie stream. *Transactions of the American Fisheries Society* 120:98–105.

Zorn, T.G., P.W. Seelback, M.J. Wiley. 2002. Distributions of Stream Fishes and their Relationship to Stream Size and Hydrology in Michigan's Lower Peninsula. *Transactions of the American Fisheries Society* 131:70-85.