Final Report

Creating a Screening Tool for Identification of the Ecological Risks of Human Activity on Watershed Quality

Steven I. Gordon, Sanjeev Arya, and Kelly Dufour
City and Regional Planning
The Ohio State University
Columbus, OH 43210

Report to the U.S. Environmental Protection Agency on Cooperative Agreement #CR826816-01-0

March, 2001

Acknowledgements

We would like to acknowledge the assistance of a number of people on this project. First, we must thank Dr. Susan Cormier and Bhagya Subramanian from USEPA in Cincinnati for their help in obtaining some of the data we used in the study. Dr. Cormier also helped to advise us on approaches to data analysis and presentation.

The biological data used in the study came from Ohio EPA. We must thank Chris Yoder, Ed Rankin, and Marc Smith for their assistance in obtaining the data and in its initial analysis. Arnold Engelmann, formerly at Ohio EPA, for help with DEM processing.

Dr. Sarada Majumder (currently at SBI) was instrumental in completion of the USEPA Star grant. Her regional model was the basis for our work.

Kang-Ping Shen, graduate student at Ohio State University, assisted with census database processing. Tracy Douglas and John Snowden, both formerly at the Center-For-Mapping at Ohio State University, for logistical support for Arc/Info on Unix and with the initial setup of our map server.

Table of Contents

List of Tables and Figures

Introduction

Recent work in the Eastern Cornbelt Plains Ecosystems of Ohio has defined a strong set of relationships among watershed corridor conditions, land use, and biological water quality as measured by the Index of Biotic Integrity (IBI) (Gordon and Majumder, 1997; Gordon and Majumder, 2000). Ohio EPA has been a pioneer in the use of indices to quantify aquatic community attributes (Ohio EPA 1987a, 1987b, 1990; Yoder and Rankin, 1995; Rankin, 1989 and 1995). The indices use the concepts of Karr's (Karr, 1981; Gammon et.al., 1976 and 1981) numerical biological criteria (biocriteria) to create measures applicable to ecoregions developed by Omernik (1987, 1988; USEPA 1988b and 1989). As part of our on-going EPA grant, we have assembled the rich database of biological quality measures available for the Eastern Cornbelt Plains Ecoregion into a Geographic Information System database for analysis. The database also includes information on the Qualitative Habitat Evaluation Index (QHEI), the Invertebrate Community Index (ICI), land cover, soils, and water quality variables.

Using this information, we have successfully created a set of empirical models that link changes in the biological diversity to changes in land use and stream condition indices. This report summarizes a project that extended this work to create a GIS-based screening tool for the World Wide Web that can be used by planning agencies, conservation organizations, government agencies, and citizens to define the nature of potential changes that could occur in their watersheds and to see the potential impacts of those changes on biological diversity in their streams.

The website we created has several components. First, we have created a set of informational materials about watersheds and the measurement of water quality using both chemical and biological measures. The information introduces the user to the concepts, definitions, and measures of environmental quality important to watersheds. Second, the website allows the user to find and explore data on 18 of the 25 major Ohio watersheds we included in our study. A web-based mapping tool lets users create their own map of the data, view data tables, and thus visualize the current state of quality in their watershed. The third part of the website allows the users to explore the potential impacts of future development scenarios in the watersheds. The forecasts are made using a revised version of the Gordon and Majumder model cited above and are presented along with information concerning the model and its derivation.

The web-based information is intended as an educational tool for planners, public officials, and citizens interested in water quality issues. The goal is to make materials relating to the complexity of watersheds and their management understandable to this audience. Accordingly, one of other project tasks was to explore other versions of the biological model that would allow these constituencies to explore the impacts of changes in their watersheds at both the "local" and "regional" scale. To create a local scale model, we needed to investigate whether there are any reasonably accurate measures of the impacts of local development decisions in terms that are understandable by the users.

This report summarizes the additional research we undertook, the creation of the website and the interactive mapping tools, and the implementation of modeling output available to the public.

Regional Scale Models

Regional scale modeling was undertaken as a part of a U.S. EPA Star Grant (Gordon et. al, 2000). Details of the data collection, analysis, and modeling are provided in that report. The study area used in that study and the current study is the Eastern Cornbelt Plains Ecoregion of Ohio. That area is shown in Figure 1 along with the boundaries of the 25 major Ohio watersheds in that region. When watersheds crossed the ecoregion boundary, the entire watershed was included in the analysis if data were available.

These watersheds were further divided into smaller units so that we could test whether watershed size made a difference in the explanatory power of the models we created. These watersheds are shown as Figure 2.

Ohio EPA was among the first to recognize that chemical standards alone are not sufficient for measuring biological integrity (Ohio EPA 1987a, 1987b). They recognized that biotic interactions, habitat structure, flow regime, and chemical contamination all interact to produce a long-term picture of community health. Thus, they derived a set of biological criteria along with chemical criteria to govern Ohio watersheds (Yoder and Rankin, 1995; Rankin and Yoder, 1991).

One of the measures that is part of the regulations is the Index of Biotic Integrity or IBI. This measure was originally developed by Karr (1981) and modified by the Ohio EPA in 1987. The index is designed to measure the aquatic vertebrate community and the surrounding conditions by using the fish species as indicators.

Along with this index, information is also gathered for Invertebrate Community Index (ICI) and Qualitative Habitat Evaluation Index (QHEI). The ICI is similar to the IBI and measures the health of the macroinvertebrate community. The QHEI creates a quantitative assessment of the physical characteristics of a stream that measures the quality of the habitat to support biological communities.

Along with these measures, our project focused on gathering related proxies for environmental stressors on the agricultural communities that allowed us to derive a set of empirical models. Our general modeling approach can be seen in Figure 3.

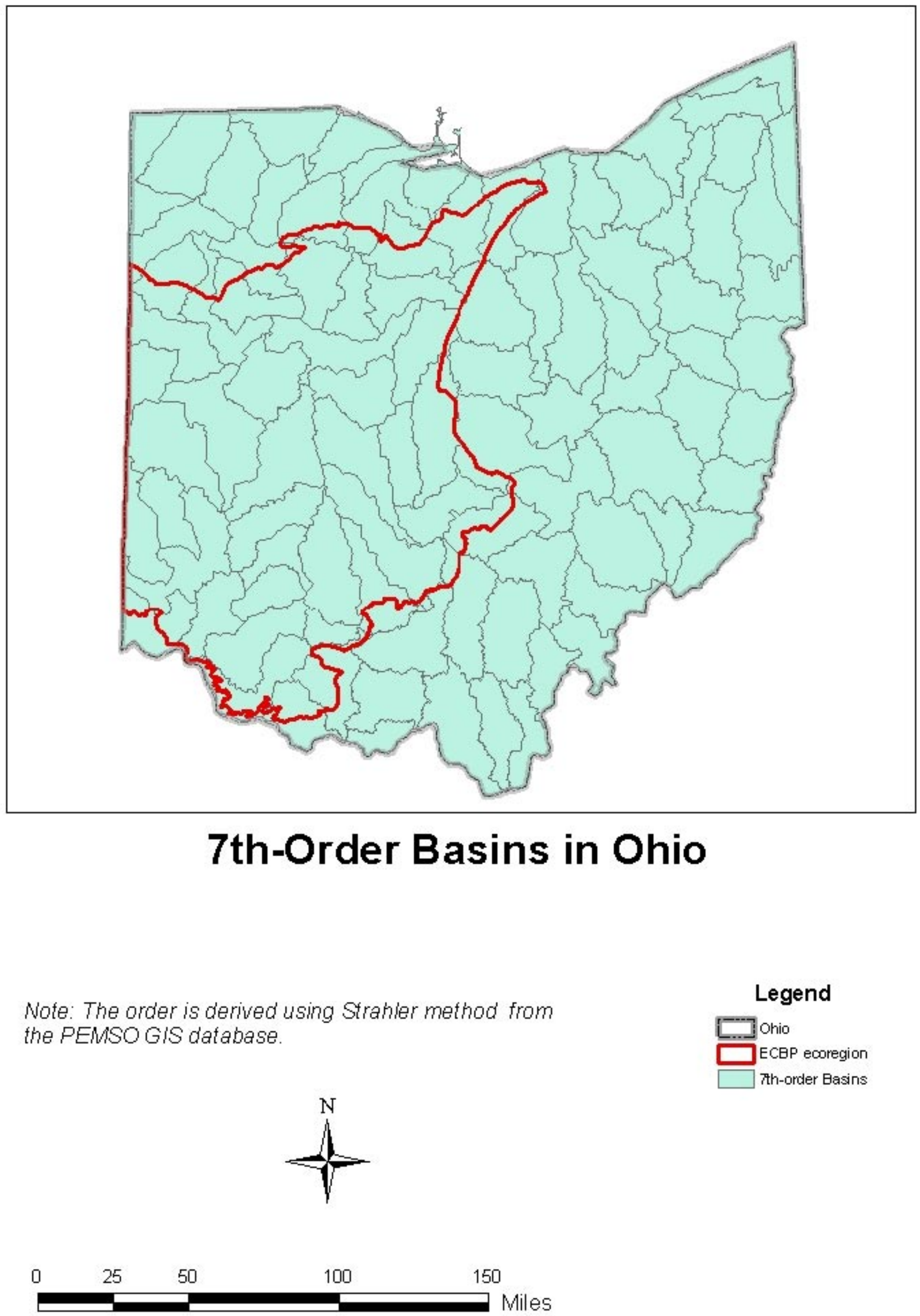

7th-Order Basins in Ohio
Note: The order is derived using Strahler method from the PEMSO GIS database.
Legend
Ohio
ECBP ecoregion
7th-order Basins
N
0
25
50
100
150
Miles

Figure 1

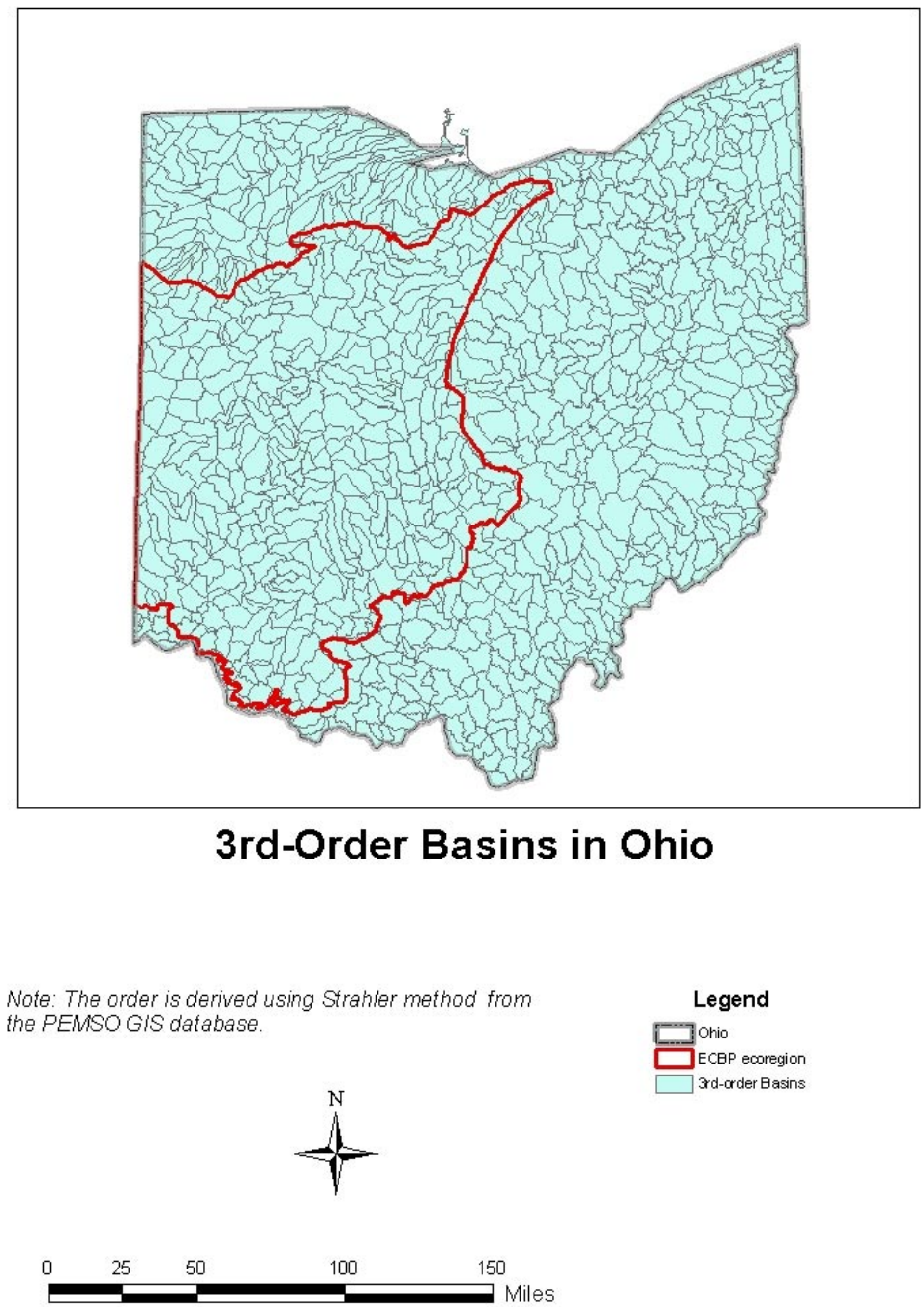

3rd-Order Basins in Ohio
Note: The order is derived using Strahler method from the PEMSO GIS database.
Legend
Ohio
ECBP ecoregion
3rd-order Basins
N
0
25
50
100
150
Miles

Figure 2

Regional Model of Watershed Biological Quality

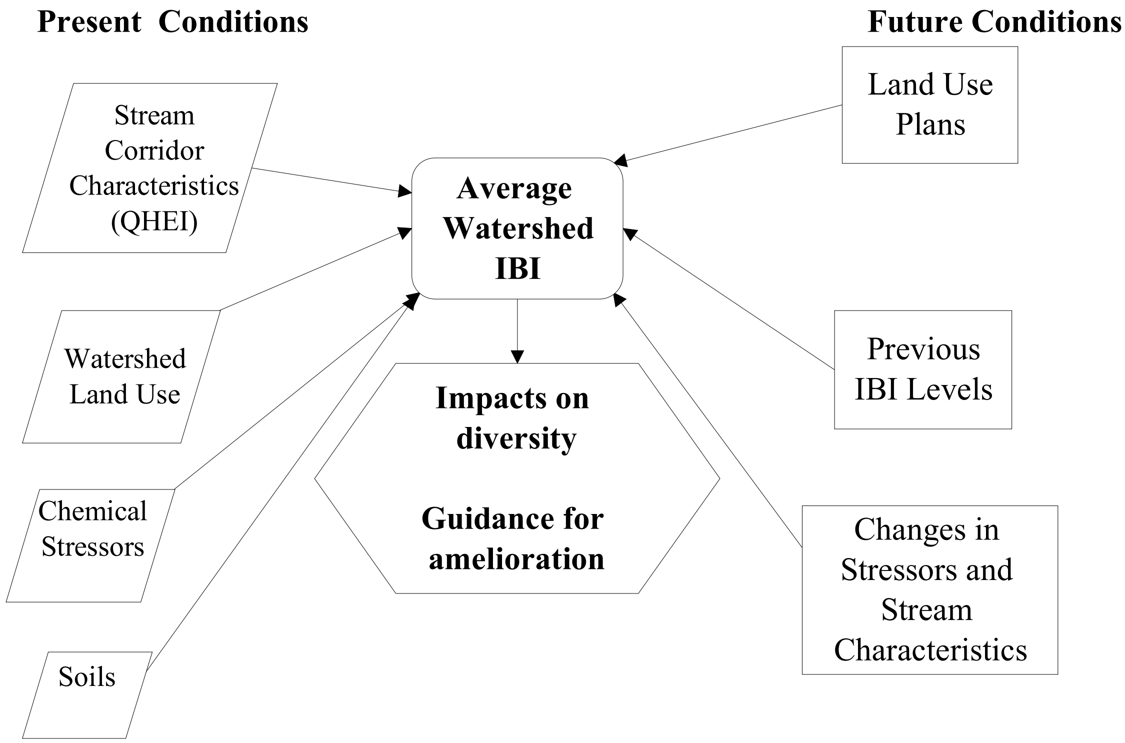


Figure 3

The watershed characteristic that is the focus of the model is the average watershed IBI. There are four groups of explanatory variables. The first set of variables relates to the characteristics of the stream habitat. Here, the components of the QHEI may be a good set of proxies for the conditions of the streambed, banks, and habitat.

Next, we need to consider the impacts of watershed land use on the stream. These of course subsume many different kinds of potential impacts: nonpoint source pollution loads, changes in runoff characteristics of the land surface, the amount of natural vegetation available to absorb or filter rainfall and pollutants before they get to the stream, etc.

Third, we need to evaluate the impacts of chemical pollutants from point sources. They need to be characterized by type, source location with respect to the locations where IBI is being measured, and both peak and average loading rates.

Finally, the soils in the watershed can be a proxy for the natural conditions that make each watershed somewhat unique. An area predominated by poorly drained, hydric soils may react much differently to land use changes than one predominated by well-drained soils.

We created statistical models that project the IBI levels based on a combination of these sets of variables. Our intention in the current project was to choose the best such model to forecast the impacts of future changes in the watersheds on the biological health of those streams. A web interface was then used to allow citizens and planners to assess relevant conditions and consider these types of consequences in making decisions concerning the amount and location of future land use changes.

Data

Table 1 lists the major watersheds. Eighteen of these 24 watersheds were used in the analysis. the table shows a summary of the number of sample points that have data on the IBI and QHEI at the same locations in the 1990-1994 timeframe for the study data. These points are shown in Figure 4. Each area of the state is not sampled each year due to cost and staff constraints, but the entire state is sampled in a roughly four year cycle. Since the measures are representative of long-term trends, this is considered to be sufficient. For our analyses, we used the data from 1989 through 1995.

The IBI is designed to measure the aquatic vertebrate community and the surrounding conditions by using fish species as indicators. Overall, there are 12 fish community variables that can be broken down into three main categories: species richness and composition, trophic composition, and fish abundance and condition. By assessing the variables within these parameters, scientists can compare a sampled site with a relatively undisturbed site with similar geographical and climatic conditions. With this rationale, the only variable would be stressors resulting from human development and disturbance. The following table lists the 12 variables measured in the IBI and their applicability depending on particular sites. Table 2 shows the components of the index.

Each of the variables can earn a score based on the comparison of a relatively undisturbed site:

5-closely approximates undisturbed site

3-somewhat approximates the undisturbed site

1-does not approximate the undisturbed site at all.

The maximum score possible from IBI assessments is: 12 (variables) * 5(highest score) = 60, whereas the minimum score possible is: 12 (variables) * 1(lowest score) = 12. Therefore, IBI scores can range from 12-60 depending on the amount of disturbance that has taken place at and around the sampling site.

The qualitative habitat evaluation index (QHEI) gives scientists a quantitative assessment of physical characteristics of a sampled stream similar to IBI and ICI biological data. QHEI represents a measure of instream geography. By combining evaluations of QHEI and IBI, for example, researchers can gain a well-rounded perspective of both the physical and biological conditions of a particular stream site. This comprehensive assessment is critical for evaluating disturbance and land use practices. There are six variables that comprise this index (represented in the Table 3).

Metrics for QHEI:
Substrate: This metric includes two components, substrate type and substrate quality.

- Type: The two most common types are two be scored, unless one substrate predominates (greater than 75%-80% of bottom area). Substrate types are defined as follows:
    1. **Bedrock:** solid rock forming a continuous surface
    2. **Boulder:** rounded stones over 256mm in diameter or large "slabs" more than 256mm in length
    3. **Cobble:** stones from 64-256mm in diameter
    4. **Gravel:** mixture of rounded coarse material from 2-64mm in diameter
    5. **Sand:** materials 0.06-2.0mm in diameter; gritty texture
    6. **Silt:** 0.004-0.06mm in diameter; fine material generally feels "greasy" when rubbed between fingers
    7. **Hardpan:** particles less than 0.004mm in diameter; usually clay that forms a dense, gummy surface that is difficult to penetrate
    8. **Marl:** calcium carbonate; usually greyish-white; often contains mollusk shell fragments

Table 1: Watershed names and numbers and biological sampling stations in the study area		
Watershed Number	Watershed Name	Number of congruent IBI/QHEI locations 1990-1994
34	Upper Scioto River	7
35	Scioto River and Mill / Bokes / Fulton Creek	10
36	Upper Olentangy River	24
37	Lower Olentangy River	27
38	Big Walnut Creek	24
39	Big Darby Creek	63
40	Walnut Creek	20
50	Upper Little Miami River	37
53	East Fork Little Miami River	25
54	Lower Little Miami River	37
55	Upper Great Miami River	37
56	Great Miami River and Loramie Creek	32
57	Stillwater River	7
58	Mad River	55
59	Twin Creek	4
60	Middle Great Miami River	2
61	Fourmile Creek	19
62	Lower Great Miami / Whitewater River	31
63	Wabash River	1
66	Blanchard River	9
67	Lower Auglaize River	9
68	Ottawa River	15
70	Upper Auglaize River	21
82	Lower Sandusky River	3

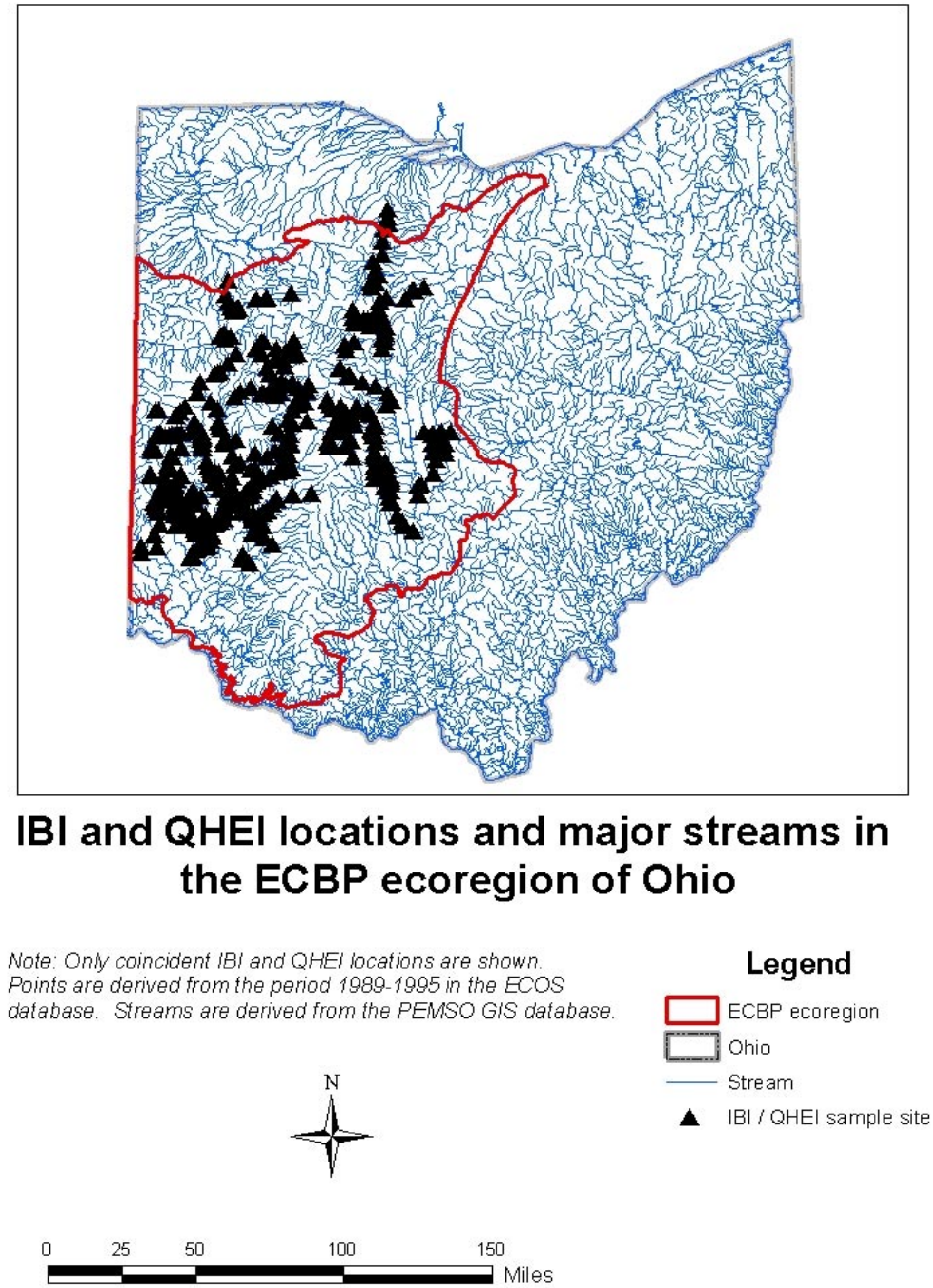

IBI and QHEI locations and major streams in the ECBP ecoregion of Ohio
Note: Only coincident IBI and QHEI locations are shown. Points are derived from the period 1989-1995 in the ECOS database. Streams are derived from the PEMSO GIS database.
Legend
ECBP ecoregion
Ohio
Stream
IBI / QHEI sample site
N
0
25
50
100
150
Miles

Figure 4

Table 2: Index of Biotic Integrity Components in Ohio		
	Variable Measured	**Type of Site**
1.	Total Number of Species	H W B
2.	Number of Darter Species	H W
	Percent Round-bodied Suckers	B
3.	Number of Sunfish Species	W B
	Number of Headwater Species	H
4.	Number of Sucker Species	W B
	Number of Minnow Species	H
5.	Number of Intolerant Species	W B
	Number of Sensitive Species	H
6.	Percent of Tolerant Species	H W B
7.	Percent of Omnivorous Species	H W B
8.	Percent of Insectivorous Species	H W B
9.	Percent of Top Carnivores	W B
	Percent of Pioneering Species	H
10.	Number of Individuals	H W B
11.	Percent of Hybrids	W B
	Number of Simple Lithophilic Species	
12.	Percent of DELT Anomalies	H W B

Type of Site: H-Headwater, W-Wading, B-Boat
DELT-Deformities, eroded fins, lesions, and tumors
Source: IBI criteria as taken from Ohio EPA 1987a

Table 3: Qualitative Habitat Evaluation Components		
Metric	**Metric Component**	**Best Possible Score**
Substrate	• Type • Quality	20
Instream Cover	• Type • Amount	20
Channel Morphology	• Sinuosity • Development • Channelization • Stability	20
Riparian Zone	• Width • Quality • Bank Erosion	10
Pool Quality	• Max Depth • Current • Morphology	12
Riffle Quality	• Depth • Substrate stability • Substrate embeddedness	8
Map Gradient		10
TOTAL		100

9. **Detritus:** dead, unconsolidated organic material covering the bottom; includes sticks, wood, and other partially decayed plant material
10. **Muck:** black, fine, flocculent, completely decomposed organic matter
11. **Artificial:** substrates such as rock baskets, gabions, bricks, trash, concrete, etc., placed in stream for reasons other than habitat mitigation
12. **Sludge:** a thick layer of organic matter that is of human origin; if originates from point source, not included

- Quality:When scoring quality, **origin** refers to the parent material that the stream substrate is derived from. **Embeddedness** is the degree to which cobble, gravel, and boulder substrates are surrounded, impacted, or covered by fine materials. Substrates should be considered embedded if more than 50% of the surface of substrates are embedded (cannot be easily dislodged). This includes substrates that are concreted or "armour-plated". Scoring **Extensiveness** of the sampling area is as follows: **Extensive** is 75% of area, **Moderate** is 50%-75% of area, **Sparse** is 25%-50% of area, and **Low** is less than 25% of area.

  Silt cover is the extent to which the substrate is covered by silt. **Silt heavy** means that nearly all of the stream bottom is layered with a deep covering of silt. **Moderate** includes extensive coverings of silts, but with some areas of cleaner substrates. **Normal** silt cover includes areas where silt is deposited in small amounts along the stream margin or is present as a "dusting" that appears to have little functional significance. **Silt free** substrates are those that are exceptionally clean of silt.

  Instream Cover: The first half of instream cover is the *type* that is present. Any cover that is in more than five percent of the sampling area should be noted, but should not be counted if in areas of the stream that are too shallow (usually <20 cm) to make it useful. Instream cover *amount* can be categorized by: extensive, moderate, sparse, or nearly absent. **Extensive** cover is that which is present in greater than 75 percent of the sampling area. **Moderate** is about 25%-75%, **Sparse** is less than 25%, and **Nearly Absent** is when no large patch of any type exists anywhere in the sampling area.

  Channel Morphology: Relates to quality of the stream with regard to creation and stability of macrohabitat. This includes: channel sinuosity, channel development, channelization, and channel stability.
- Sinuosity: The degree to which a stream bends. **No** sinuosity means the channel is straight. **Low** sinuosity would have one or two poorly defined bends. **Moderate** has more than two outside bends, with at least one being well defined. **High** sinuosity would have more than two or three well defined outside bends with deep areas outside and shallow areas inside.
- Development: Refers to development of riffle pool complexes. **Poor** means no riffles or shallow ones with sand and fine gravel. **Fair** are poorly developed or absent riffles. **Good** implies better defined riffles with larger substrates. **Excellent** means the riffles are good and pools have a maximum depth of more than one meter and deep riffles and runs are present.
- Channelization: Refers to human-made channel modifications. **Recovered** means that the streams were channelized in the past, but have since regained most of their natural characteristics. **Recovering** means the stream was channelized, but is in the process of regaining its former , natural characteristics. **Recent** or **No Recovery** implies that the stream was recently channelized or show no significant recovery.

  Stability: How stable the channel remains. Channels with stable banks and substrates with little or no erosion are categorized as **High** stability. Artificially stable (concrete) channels also receive the **High** mark. **Moderate** scores are given to channels with stable riffle/pool and channel characteristics, but also exhibit symptoms of instability. **Low** scores go to channels with fine substrates in riffles, unstable (eroding) banks, and high bedload.

  Riparian Zone: This metric measures the quality of the riparian buffer zone of floodplain vegetation, including riparian zone width, floodplain quality, and extent of bank erosion. To score each component, one looks downstream and averages both the left and right banks.
- Width of Floodplain: This is the width of the riparian vegetation. Estimates should only be taken for forest, shrub, swamp, and old field vegetation (fairly mature successional field that has stable, woody plant growth).

- Floodplain Quality: The two most predominant floodplain quality types are to be checked. Floodplain refers to areas immediately outside of the riparian zone or greater than 100 feet from the stream, whichever is wider on each side of the stream.
- Bank Erosion: This can have one of five different scorings:
    1. **None**-streambanks are stable and not being changed by water flows or animals
    2. **Little**-streambanks are stable, but slightly changed along the transect line; less than 25% of streambank is receiving any stress, is false, broken down, or eroding
    3. **Moderate**-streambanks are receiving moderate alteration along transect line; at least 50% of streambank is in natural stable condition; 50% is false, broken down, or eroding
    4. **Heavy**-streambanks have received major alterations along transect lines; less than 50% of streambank is in stable condition; over 50% of streambank is false, broken down, or eroding

  Severe-streambanks along transect lines are severely altered; less than 25% of bank is stable condition; over 75% of bank is false, broken down, or eroding

  Pool Quality: Pool quality consists of three areas: maximum depth of pool or glide, current type, and morphology.
- Depth: This can range from a score from zero to six. A pool or glide with maximum depths of less than 20 cm are considered to have lost their function and the total metric score is zero.
- Current Type: There are seven possible categories for current type:

1. **Torrential**-extremely turbulent with fast flow and large waves; water surface very broken with no consistently connected surface
2. **Fast**-mostly non-turbulent flow with small standing waves in riffle-run areas; water surface partially broken, but some areas of consistent connectivity of surface
3. **Moderate**-detectable and visible non-turbulent flow; water surface visibly connected
4. **Slow**-water flow is perceptible, but very sluggish
5. **Eddies**-small areas of circular current usually formed in pools just downstream from riffle-run areas
6. **Interstitial**-flow only perceptible in interstitial spaces between substrate particles in riffle-run areas
7. **Intermittent**-no flow; standing pools separated by dry areas

   Morphology: This category would be checked **wide** if pools are wider than riffles, **equal** if pools and riffles are the same size, and **narrow** if riffles are wider than pools. If morphology varies throughout the site, average the types.

   Riffle Quality: If no riffles exist, a zero should be recorded. If not, riffle quality consists of three areas:

- Riffle Depth: A score from zero to four is to be chosen to describe the depth characteristics of the riffle. If the riffle is less than five cm deep, riffles are considered to have lost their function and a score of zero should be recorded.
- Substrate Stability: A score from zero to two is chosen that best describes the substrate type and stability of the riffle habitats.

  Embeddedness: This is the degree that cobble, gravel, and boulder substrates are surrounded or covered by fine material. Substrates are embedded if more than half of the surface of the substrate is embedded in the fine material (are not easily dislodged), including substrates that are cemented. *Extensiveness* of the embeddedness in the area sampled is also recorded: **extensive** is 75% of stream area; **moderate** is 50%-75% of area; **sparse** is 25%-50% of area; **low** is less than 25% of area.

  Map Gradient: Calculated from USGS 7.5 minute topographic maps by measuring elevation drop through the sampling area. First, the stream length is measured between the first contour line upstream and the first contour line downstream of sampling site and then dividing the distance by the contouring interval. A minimum distance of one mile should be used if contours are "packed" together.

In order to gain some insights into the basic biological variables, we undertook some simple statistical analyses. For the IBI, the mean value for the entire ecoregion is 38.74 while the median is 40. The distribution within watersheds if far from even however. This is illustrated by Figure 5.

The central line in the box is the median value of the dataset. The lower end of the box is the 25th percentile of the dataset, and the upper end of the box is the 75th percentile of the dataset. These are also called “hinges”. The length of the box is the interquartile range of the dataset. The whiskers are the lines extending from the box, and they extend up to 1.5 times the interquartile range from the “hinges” in each direction. The values beyond that range are mapped individually and are termed outliers. Values beyond 3 interquartile ranges of the “hinges” are also mapped individually using a different symbol than the “outliers” and are called the extremes. The boxplot is used to visually describe the distribution and is preferred over a histogram in this case because it can provide a visual estimate of the distribution of the data, as well as the presence and position of outliers (SPSS 1996, Hartwig and Dearing 1979, pp23-25).

The boxplots show that our sample includes a wide range of watersheds with both low and high IBI values and major differences in the distribution of the IBI within the watersheds. For example, one can compare the Big Darby Creek (#39) with a high median value and relatively compact distribution to several watersheds where the median drops to the 20's or 30's - the Upper Scioto River (#34) or the Lower Great Miami (#62). The Lower Olentangy (#37) is an example of a watershed with a relatively high median but a distribution skewed toward the lower end. This is caused by the heavily urbanized Columbus area in the lower reaches of this stream. The most intensively agricultural watersheds in the northern part of the region (#66, #67, #68, #70, and #82) all have IBI distributions that are lower than the rest of the region.

A similar set of plots was made for the QHEI variable and is shown as Figure 6. These plots show a similar, wide distribution. The number of outliers is higher. These turn out to be in watersheds with the most significant urban development where a great deal of channel modification has occurred.

The final data component for our modeling was soils information. Soil properties impact stormwater runoff levels and transport of pollutants through the soil, and might indirectly affect the stream IBI level. Soil properties were extracted from the Natural Resources Conservation Service’s regional State Soil Geographic (STATSGO) Data Base.

The STATSGO is chosen because of its scale and availability. These data have been generalized to 1:250,000 scale from more detailed county soil surveys. They are used because of the availability of digital coverage over a wide area and their appropriateness for regional and multi-county modeling (USDA 1994.) Each STATSGO mapping unit might contain twenty-one components of the larger scale county soil survey mapping units. There is no visible distinction of the of spatial location of the individual large scale units, but the STATSGO attributes percentages of each STATSGO mapping units that meet certain soil properties or criterion. The soil properties can be ranked according to specific criterion and a weighted average could be calculated for each STASGO soil mapping unit or polygon.

The soil properties used in this case, are explained in Table 4 (USDA 1994).

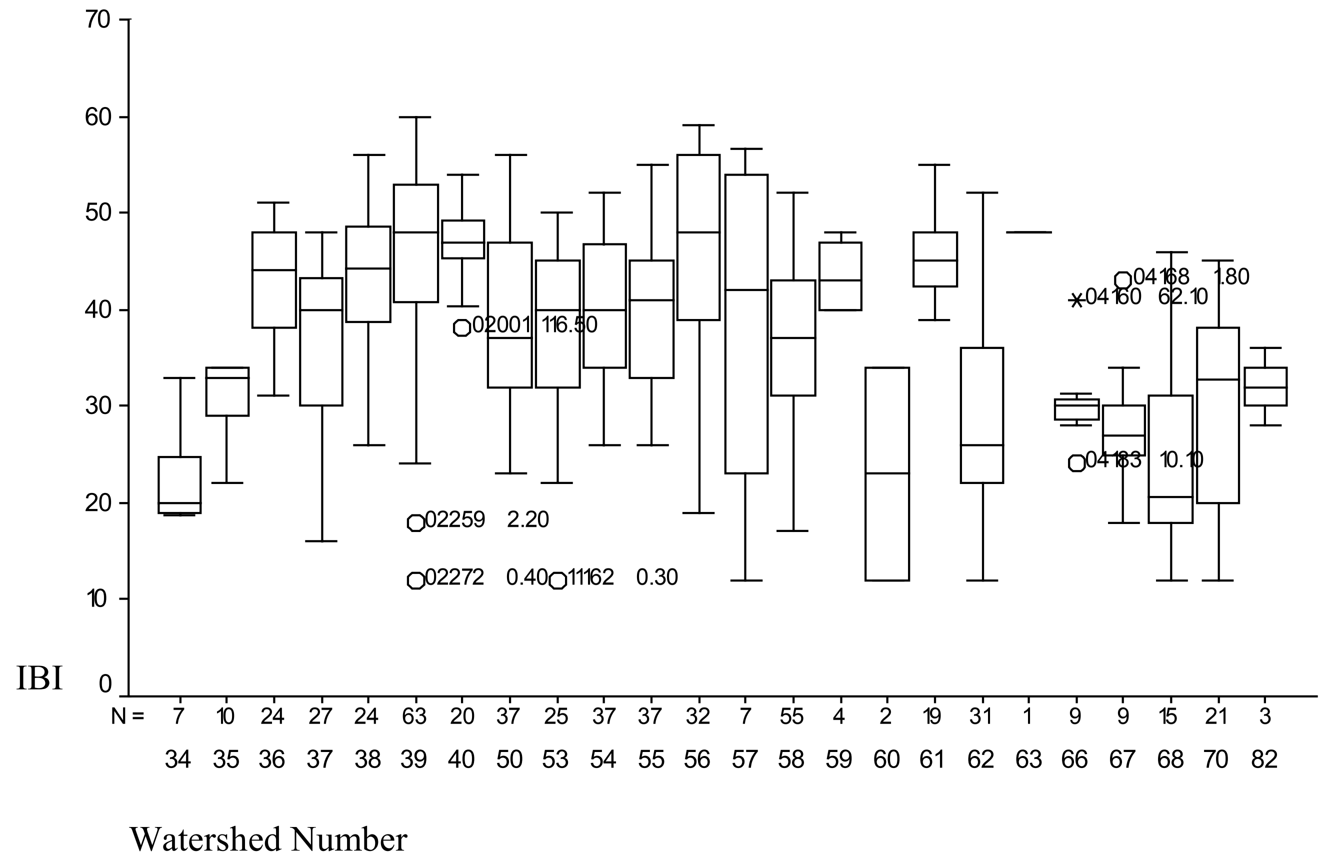


Figure 5: Boxplots of IBI Distribution by Watershed

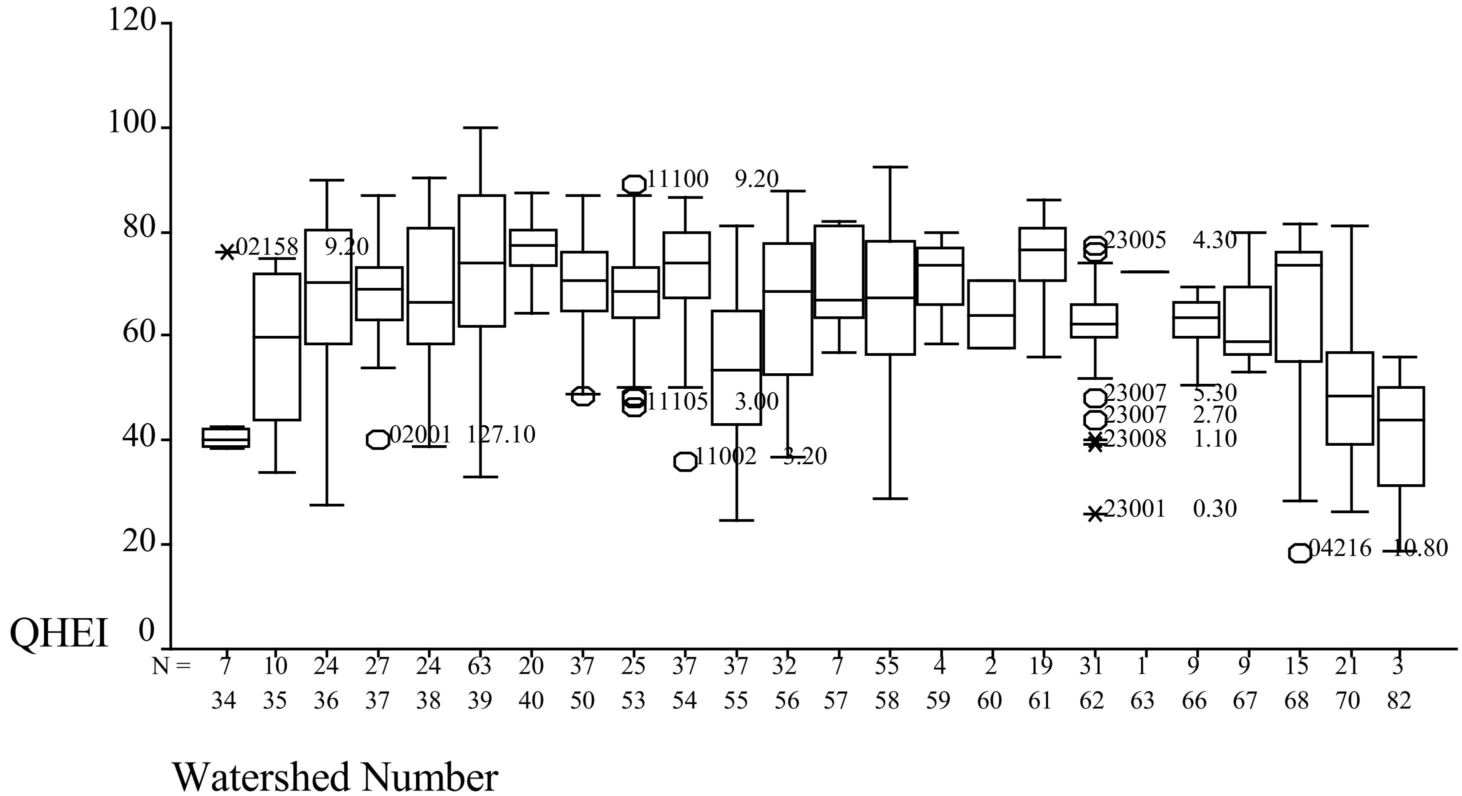


Figure 6: Boxplots of QHEI values by watershed

Table 4: Soil properties used in the analysis	
Variable Name	Soil Property
Corcon	Corrosion - Concrete : An interpretation rating of the susceptibility of concrete to corrosion when in contact with the soil
Corsteel	Corrosion - Uncoated steel : An interpretation rating of the susceptibility of uncoated steel to corrosion when in contact with the soil
Drainage	Soil Drainage class identifying the natural drainage condition of the soil and refers to the frequency and duration of periods when the soil is free of saturation. The drainage classes were rank ordered starting at 1 through 7.
Hydgrp	Soil hydrologic group such as A, B, C, D. Group A soils have low runoff rates and high infiltration while Group D at the other continuum has high runoff rates and low infiltration. The soil hydrologic group was rank ordered starting at 1 for Group A through 4 for Group D
Hydric	Hydric soil rating. Soils were binary coded with 1 for hydric soils and 0 for non hydric soils. Hydric soils are soils that are waterlogged and retain water
Primfml	Prime farmland classification - whether the any portion of the STATSGO map unit has been designated as prime farmland
Rockhard	Bedrock hardness - The degree of hardness of the underlying bedrock. Hard bedrock coded as 1, soft as 2, and others 0.
Rockdeph	The maximum value for the range in depth to bedrock
Rockdepl	The minimum value for the range in depth to bedrock
Slopeh	The maximum value for the range of slope of a soil component within a STATSGO mapunit
Slopel	The minimum value for the range of slope of a soil component within a STATSGO mapunit

Since the boundaries of the study area watersheds and the STATSGO mapping units are non contiguous, an area weighted average was calculated for each of the properties listed in Table 4, for the seventh order and third order watersheds. For each watershed, the proportion of area occupied by each unique STATSGO mapping unit is calculated. The soil properties of each mapping unit is then multiplied by the proportion of area it occupies in each watershed, and these area weighted properties are added for each unique mapping unit within the watershed to obtain the area weighted value for the whole watershed.

Because there are a number of correlated variables measuring soil properties, a principal components analysis was performed on the soils data. Separate components are extracted from the data summarized at the level of the seventh order watersheds, and at the level of the third order watersheds. The first two components are retained according to the "Scree" criterion (Hair et. al. 1998, pp106). The component loadings were rotated orthogonally for ease of interpretation, and Table 5 shows the rotated component loadings for each soils variable.

Table 5: Rotated Principal component loadings for soil variables using data summarized at the level of seventh and third order watersheds

Soil Property	Seventh order watersheds		Third order watersheds	
	Soil Component 1	Soil Component 2	Soil Component 1	Soil Component 2
Corcon	-0.405	0.723	0.492	0.354
Corsteel	0.636	0.718	-0.418	-0.848
Drainage	0.643	0.699	-0.450	0.853
HydGrp	0.252	0.954	0.006	0.919
Hydric	-0.507	-0.787	0.376	0.804
Primfml	0.868	0.225	-0.691	0.622
Rockhard	-0.969	-0.004	0.952	-0.195
Rockdeph	0.868	0.198	-0.837	0.265
Rockdepl	0.949	0.100	-0.922	0.229
Slopeh	-0.943	-0.300	0.906	-0.382
Slopel	-0.927	-0.326	0.880	-0.405

In case of the seventh order watersheds, the first two components explain eighty nine percent of the variance in the dataset while in case of the third order watersheds, the first two components explain eighty four percent of the variance in a dataset. It can be seen that the corrosion and the drainage properties load highly on the second component, while the rock hardness, depth to bedrock, and the slope load highly on the first component. The second component is interpreted as well drained, corrosive soil. The first component is somewhat more difficult to interpret because the categories slope, and depth to bedrock contain two variables, maximum and minimum values of each, and they both load highly on the first component.

Data for all the chemicals were not available at all the locations, and for each chemical, only a few locations had high values. The majority of LEAPS sampling sites were either missing or low range values, resulting in highly left skewed distributions. Therefore logarithmic transformations were used to approximate linear relationships with IBI. Chemical data extracted from the LEAPS database included ammonia, phosphorus, and biochemical oxygen demand concentrations of milligrams per liter and fecal coliform counts per liter. Although we know that pesticides play a significant role in some watersheds, there were unfortunately very few places where such data were available.

Of the chemical measures that were available, only ammonia and phosphorus were available at enough locations to provide a feasible sample size for the regression models. It was observed that both of the transformed values for ammonia and phosphorous appeared to have the strongest relationship to the IBI values. They could be used separately in the regression equation, but could not be used together because of collinearity. Collinearity was anticipated by the high correlation between the log transformed ammonia and phosphorus concentrations (0.83, significant at alpha levels of 0.01). Collinearity among these two variables was also indicated by the condition index and the regression coefficient variance decomposition matrix (Hair, 1998). To deal with the collinearity, first the standard normal deviates (Z scores) were calculated for the log transformed chemical concentrations. Then the IBI/QHEI sampling locations were ranked according to these Z scores of the transformed sum of their linear distance weighted upstream chemical concentrations. As a final step, the rankings were added together to create a composite score of the chemical stress at each site. In the absence of further literature dealing with similar issues, it was assumed that the effects would be additive.

Regression Results

Here we report the regression results for selected models for the third order watersheds since these were the models we apply on the website. The full results are available in Gordon et. al. (2000).

The range of models tested are shown in Table 6. The model chosen for the website use is the final model in the table, shown in bold. Here, one can see that the model explains nearly 70% of the variance in regional levels of IBI. The substrate and pool have a positive impact on IBI as one would expect. Higher values for these QHEI components indicate a habitat in better condition.

The percentage of dense urban land use has the largest negative impact on the IBI. This land use variable is a proxy for the non-point source pollution coming from urban runoff. Also having a negative influence are soils with high corrosivity and the distance decay weighted scores for the LEAPS, point source pollution loads.

As expected, as stream order goes up the stream is capable of supporting a more diverse fish community. Again, this is because of the stability of the system in larger streams.

Table 6: Regression results for data in the third order watersheds				
Model Description	Independent variables	Partial regression coefficients	Standardized coefficients	Adjusted R squared/ Sample size
IBI as a function of land use, stream order, and stream habitat variables	Substrate** Riparian** Pool** Dense urban land use** Strahler Stream order**	0.642 0.652 1.175 -0.415 1.371	0.258 0.105 0.288 -0.305 0.157	0.395 463
IBI as a function of land use, point source pollutants, and stream habitat variables	Substrate** Pool** Dense urban land use** LEAPS Ammonia** (First order distance decay)	1.184 1.219 -0.401 -1.073	0.471 0.262 -0.234 -0.301	0.576 72
IBI as a function of land use, point source pollutants, and stream habitat variables	Substrate** Pool** Dense urban land use* LEAPS Fecal Coliform* (First order distance decay)	1.010 1.634 -0.337 -0.00097	0.384 0.361 -0.206 -0.225	0.517 68
IBI as a function of land use, point and non point source pollutants, and stream habitat variables	Substrate** Pool** Dense urban land use** STORET Ammonia** (First order distance decay)	1.042 1.059 -0.449 -0.881	0.454 0.255 -0.370 -0.163	0.472 193
IBI as a function of land use, point and non point source pollutants, and stream habitat variables	Substrate** Pool** Dense urban land use* STORET Fecal Coliform** (First order distance decay)	0.768 1.547 -0.304 -0.026	0.345 0.386 -0.167 -0.200	0.446 132
IBI as a function of land use, combined point source pollutants, and stream habitat variables	Substrate** Pool** Dense urban land use** Sum of ranks of LEAPS Ammonia and Fecal coliform**	1.022 1.207 -0.356 -2.689	0.406 0.259 -0.208 -0.364	0.609 72
IBI as a function of land use, combined point source pollutants, stream habitat variables, soil component, and stream order	**Substrate**** **Pool**** **Dense urban land use**** **Sum of ranks of LEAPS Ammonia and Fecal coliform**** **Well drained, corrosive soil component**** **Strahler stream order****	**0.925** **1.161** **-0.636** **-1.571** **-2.946** **5.575**	**0.368** **0.249** **-0.237** **-0.213** **-0.223** **0.234**	**0.691** **72**

**Coefficient significant at the 0.01 level (2 tailed)
*Coefficient significant at the 0.05 level (2 tailed)

Text in bold is the model used for the website

Small Area Models

Although the regional models are potentially useful as screening tools for assessments of watershed scale changes, they are not as useful to local decision-makers who are most likely to be approving local land use and zoning changes impacting much smaller areas. The variables in our model are also difficult for them to use - especially those relating to habitat changes. They would be more likely to use this type of model if it will allow them to input the proposed land use changes in their region along with more typical information about those uses. The model would then need to have some indicators of habitat change that go along with those land uses and could then give them estimates of the potential incremental changes in the biological status of the streams.

In an effort to define such models, we explored several different approaches to modeling at these "larger" scales (that is for smaller areas). We explored research questions regarding QHEI and its components that could ultimately help in our understanding of the complex relationships between local habitat changes and changes in biological diversity. We set out to find answers to some fundamental questions:

> What are the landscape or land use factors affecting the habitat quality of our streams? Can the relationship between anthropogenic and landscape factors and QHEI be reasonably explained? Along the scale-impact continuum, what is the scale at which these factors and the stream habitat interact most strongly? Along the ecotone-continuum, do riparian zones matter more than the whole watershed? Are there any anthropogenic factors in the interaction that may provide insight into our land use planning and growth management programs?

It may be noted here that defining watersheds of the appropriate size or scale for a particular research problem is of critical significance. Changes in QHEI are hypothesized to be site-specific and related primarily to changes only in the local riparian zone whereas overall changes in biological diversity are impacted by land use and other factors throughout the watershed (as shown in our regional models).

Two fundamental theoretical reasons guide the development of this part of the study and its study area unit. One, the area downstream of a QHEI sampling site may have negligible effect on the local stream habitat characteristics at that location. Therefore, we need to define specific watersheds for each of the biocriteria sampling sites. This approach is different from the often-used approach of pooling all sampling stations along streams in a watershed and then assigning most landscape and land use characteristics equally to all the sampling station in that watershed. The critical question then becomes at what scale to measure the characteristics of this "local" habitat - such data as land use percentage, slope values, and riparian attributes. We tested one approach to this scale question in the pilot described below and then tried to refine it in the remainder of our work.

For this study, a watershed is defined as the total naturally contributing area upstream of a certain location. This location, or pour-point or mouth, is defined using the stream location and digital elevation data to calculate the total contributing area upstream. In our regional studies, a watershed typically coincided with a biocriteria sampling site location if possible. When only a subset of the watershed is used in our study of more localized impacts we call it a local watershed, or localshed, because it delineates the area more “local” to the mouth than the area downstream or very far upstream in the watershed. It is that contributing area upstream of a location that is delimited by some criteria such as a maximum threshold area, or maximum upstream flow length. It can be thought of as a logically defined subset of the total naturally occurring area, or watershed, for a location.

Two implementation issues arise from delineating watersheds for individual bio-criteria sampling locations. One, the watersheds for many sampling sites may engulf the watersheds of the immediate upstream sampling sites. Two, due to the DEM-based algorithm used in the GIS software there may be a vast difference between the size of watersheds for spatially close bio-criteria sampling sites. Areas may range from a fraction of a square mile to hundreds of thousands of square miles. This is counter-intuitive to the idea that QHEI habitat measures are very localized and site-specific. Therefore, to deal with these two issues, two different strategies are used. These also justify the use and clarification of new terminology.

To deal with the first issue of delineation, that of watersheds within watersheds, we delimited the watersheds up to the next upstream mouth. The second issue is dealt with by generating approximately similar-size localsheds.

In the first strategy, using customized routines in Arc/Info GIS software, contributing areas were generated for each QHEI sampling site such that they extended only up to the mouth of the watershed for the next upstream sampling location. We call these **non-intersecting localsheds**. Then these localsheds were analyzed in greater detail. Since these localsheds varied in size, they were grouped into two different sizes to explore the relationship between QHEI and the land use and landscape factors at different sizes. It is hypothesized that because QHEI is a site-specific indicator of stream conditions, the model's strength will improve with smaller sizes.

Watersheds between 2,000 to 20,000 acres area were grouped as small-size watersheds and watersheds between 20,000 to 200,000 acres area were grouped as medium-size watersheds. Out of a total of 497 point-based watersheds, 147 watersheds of other sizes below and above these arbitrary thresholds were not included in this strategy. There were 225 medium-size watersheds and 125 small-size watersheds. Regression models were studied for three different groups - medium-size, small-size, and all 350 small and medium watersheds.

Riparian and watershed-level geographic data were generated for all the relevant variables for each of these watersheds and managed in a GIS. The strategy was tested with two different land use datasets. One was available in-house through prior custom categorization of a Landsat image for the regional study. The other land use dataset was obtained from the National Land Cover Data (NLCD) database. This dataset was already classified from 20 Landsat images from different time periods between 1988-94. A regression model was attempted to explain the variation in QHEI in the ECBP ecoregion.

Note that sampling strategy for the original bio-criteria data plays a critical role in our delineation of the different point-based watersheds here. The size of watersheds delineated in this manner is dependent on the sampling strategy used for selecting bio-criteria measurement sites. However, the watersheds delineated in this manner are spatially independent of each other. To balance this approach, another strategy was tested which generated watersheds of approximately equal sizes.

Based on the reasons behind defining localsheds for each sampling site, a second strategy was employed. Naturally contributing areas upstream of each QHEI sample point were clipped, at 1-mile upstream flow length from the sample point, using customized routines in Arc/Info software. This was done to generate almost equal-size localsheds distributed widely across the whole ECBP ecoregion.

The motivation for selecting the 1-mile upstream flow length was not completely arbitrary. It was found that around this distance the average area of upstream contributing areas approached the average area of the small-size watersheds of the first strategy. This could help us test whether a greater explanation of QHEI in the small-size watersheds could be validated across a wider array of ecological contexts by increasing the sample size and scope. Note that the sample size of small-size watersheds in strategy I was only 125. Using this strategy we increased the number of watersheds in the sample to 580.

Data were collected from different agencies and processed in ArcInfo GIS to conform to the same projection and coordinate system. QHEI scores were extracted from the ECOS database for the period 1989-1995, and only the latest score was used in this study for sites with repeated sampling in this period. All records in our QHEI database have location and year. Multiple samples in the same year for a given site were almost identical, so only the latest values were used.

Table 7: Data layers explored in the project

Data Layer	Source
Digital elevation model (DEM)	From USGS 1:24,000 DLG hypsography files in 7.5min quads.
Biocriteria sample points	IBI and QHEI tables in USEPA ECOS database.
Localshed	Derived from DEMs using bio-criteria sample point coverage.
Streams	From Ohio EPA PEMSO database
Streams	From USGS DLG hydrography files
Land use 1	Unsupervised classification of Landsat TM imagery of August 1994, classified into major categories such as agricultural, urban, forest. 30m resolution.
Land use 2	From NLCD project, classified into finer categories such as low and high density residential, row crops, and evergreen forests, using Landsat imagery from 1988-94. 30m resolution.
Roads	From USGS 1:24,000 DLG transportation files in 7.5min quads.
Soils	From NRCS 1:250,000 STATSGO database.
Point sources	From USEPA database of 1993 NPDES permittees.
Population	From Bureau of Census TIGER/Line95 files and block-level CENSUS database in STF1B files.
Housing units	From Bureau of Census TIGER/Line95 files and block-level CENSUS database in STF1B files.
Sinuosity	Derived programmatically from Streams coverage.
Slopes	Derived from DEMs using GIS.
Chemical parameters	From USEPA STORET database.
Stream Order	Derived from cleaned/snapped streams using Strahler algorithm.

Land use was based on two different datasets. One was derived from in-house classification for the regional IBI modeling study. Urban areas were classified using ancillary datasets such as census data to distinguish urban areas more completely. The other land use dataset was obtained from the MRLC project's NLCD archive available publicly. The classification, derived from a uniform national methodology was based on 20 Landsat scenes from the period 1988-1994. Stream order was used to identify headwater streams. Because of the imperfect matching of stream endpoints at many locations, streams with a Strahler order value of 2 were also termed as headwater, besides those with Strahler order of 1.

We hypothesized that as each of the indicators of urban development increased, the QHEI would decline. We tried a number of indicators to ascertain which might best explain changes in QHEI. Most of the indicators focused on the riparian zone since the QHEI is a rating of that area. Land use changes in that area were used as one indicator of change. Roads in the area were another indicator. We hypothesized that as the density of roads in the riparian area increased, there were increased probability that stream habitate had been altered to accommodate roads and bridges. Population and housing density in the riparian zone provided a similar indicator.

Indicators of natural conditions were also thought to be important. Sinuosity is a measure of how straight the stream segment is. We would expect that those streams that have been altered for human uses would have lower sinuosity than those left in their natural state. The problem with this measure is that the GIS data from which it is derived may be a too coarse a scale to measure the differences caused by urban development.

Slopes were also hypothesized to impact the relationships. Development on steeper slopes should produce greater impacts on bank erosion and sediment problems.

Pilot Study Results

A pilot study was performed to build a linear regression model to explain the variation in QHEI as a function of natural and anthropogenic factors and stresses. A number of preliminary models were tested. It was clear from these preliminary models that anthropogenic factors negatively impact the stream habitat, or the index for indicating the health of stream habitats – Qualitative Habitat Evaluation Index (QHEI).

GIS database is space- and resource-intensive, therefore, only two watershed regions, Big Darby Creek and Great Miami River, were selected for testing the hypothesis.[1] Big Darby Creek watershed is to the west of Columbus metropolitan area. The Great Miami River watershed covers the region around Springfield, Ohio.

Deforestation, agricultural land use and urban development may all reduce the forest cover that was significantly and positively related to QHEI. Forest cover (both riparian and non-riparian), reach sinuosity, and number of point sources in the catchment were found to be significant in the model. The best preliminary model explained more than 60% of the variation in site-specific QHEI. This model as a whole was significant at the 1% level. Riparian agriculture and riparian road density were both significantly and negatively related with QHEI.

Some degree of multicollinearity was also indicated among independent variables in the model. A study of the residuals indicated some heteroscadasticity in a few variables. However, since theoretical and intuitive reasoning did not rule out a linear relationship, it was decided not to transform many variables at this stage. Another major concern with the pilot was its lack of substantial cases in the extreme ranges of QHEI, and the small sample size of 18. These concerns were sought to be addressed by a more rigorous analysis of a greater number of cases, across the whole ECBP ecoregion, and with values fairly representing the whole range of QHEI scores.

Extended Study

The pilot was followed by a full-scale analysis of the bio-criteria sampling sites in the ECBP ecoregion using the two approaches mentioned previously. The explanatory variables for only the upstream contributing areas were used to explain the variations in QHEI. The relative significance of riparian zones vs. the variables measured at the whole localshed scale were also `studied.

Strategy I: Non-intersecting Localsheds

Models were built to explain QHEI at different locations within the ECBP ecoregion. Since QHEI is theoretically linked to site-specific indicators, considerable GIS processing was required to generate geographic data for smaller localsheds upstream of specific QHEI sample sites. Large-scale data was required for these models for factors such as roads, streams, population, land use, and slopes. Riparian data layers were generated for different strip widths, such as 30m, 90m, and 500m, to test whether the strength of ecotonal interactions varies by distance. Twenty-seven different layers were generated for each non-intersecting localshed. To give an idea of the size of the GIS database, almost 10,000 coverages were generated for small- and medium-size non-intersecting localsheds. These amounted to more than 400MB of disk space, 14 hours of processing to generate GIS data, and more than an hour to read and export the data to statistical packages, on a Windows NT4 machine with 384MB RAM.

---

[1] These regions have been defined, and named as such, by the Ohio EPA based on the watersheds for the seventh-order (Strahler) streams of Ohio. A total of 93 seventh-order watersheds have been defined for Ohio. This study comprises of regions in Ohio EPA's seventh-order watershed 39 (Big Darby) and parts of watersheds 55 and 56 (Great Miami).

A linear regression model was first built based on the GIS data extracted for all the localsheds of different sizes, and for different riparian zone widths in each localshed. The best model is described in the table here. The model explains about 24% of the variation in QHEI using independent variables such as riparian residential land use, riparian dense forest land use, riparian row crops land use, and the dummy to indicate whether the sample was on a headwater stream. It may be noted that using the stepwise selection of variables in the statistical package, riparian-scale land use variables are selected as they have a stronger relationship with the QHEI score in this ecoregion. Moreover, within the riparian variables, those measured for 30m strips across each side of the stream are more influential than those measured at 90m or 500m strip widths. This lends weight to our hypothesis that QHEI is impacted more by reach-scale riparian variables in the immediate localshed than variables measuring both upstream and downstream contexts at the basin scale.

Table 9 - Regression model results with 350 medium- and small-sized non-intersecting localsheds

Independent variables	Unstandardized coefficients		Standardized coefficients				
	b_i	Std. Error	Beta	t	Prob. > |t|	Sign OK?	Significant at 5%?
Intercept	74.014	4.525		16.355	0.000		
Low-density Resid. in 30m buffer	-0.358	0.120	-0.182	-2.981	0.003	Yes	Yes
Dense Forest in 30m buffer	0.150	0.077	0.139	1.950	0.052	Yes	Yes?
Row Crops in 30m buffer	-0.254	0.057	-0.357	-4.487	0.000	Yes	Yes
Headwater (dummy)	-4.636	1.800	-0.126	-2.575	0.010	Yes	Yes

N = 350   Prob > F = 0.000   Std.Error of Estimate: 15.95
R-Squared: 0.24   Adj. R-square: 023

All the variables have the expected sign. Low-density residential land use in riparian zones is believed to cause increased runoff due to added impervious cover in the stream's vicinity. Row crops may adversely impact stream habitat due to practices such as seasonal denuding of the land surface, and cropping practices near the river that may add to greater erosion. Headwater streams are known to be more sensitive to physical impacts because of lower volumes of water available to dilute and flush away unhealthy elements. Generally, forested land use reduces runoff, sediments, and nutrients and stabilizes stream flow and channel morphology. The presence of forests in the subwatershed may stabilize the soils in the upstream regions, reduce the uprooting of topsoil by rainwater because of canopy intervention, and deplete particle content of surface runoff through infiltration. Forest may improve the quality of a stream's overall physical habitat.

All the variables in this model also display a statistically significant relationship with QHEI, except dense forest land use which is almost significant at the 5% level. Riparian row crops is the land use with the strongest impact on the stream's habitat quality.

The next step was to test the hypothesis that more immediate spatial contexts impact the stream habitat more than larger regional-scale factors. We separated the 350 localsheds into two categories. The 125 localsheds with areas between 2,000 and 20,000 acres were called small-size localsheds and the rest 225 medium-size localsheds comprised of areas ranging from 20,000 to 200,000 acres.

Table 10: Regression model results with 225 medium-sized non-intersecting localsheds

Independent variables	Unstandardized coefficients		Standardized coefficients	t	Prob. > |t|	Sign OK?	Significant at 5%?
	b_i	Std. Error	Beta				
Intercept	49.606	2.045		24.258	0.000		
Dense Forest in 90m buffer	0.742	0.088	0.492	8.449	0.000	Yes	Yes

N = 225 Prob > F = 0.000 Std.Error of Estimate: 15.71
R-Squared: 0.24 Adj. R-square: 024

With only the medium-sized localsheds in the data, we attempted to build another model. The strength of this model, however, was not very different from that in the earlier model with all localsheds. About 24% of the variation in QHEI was explained by the model, the same as in the previous model with localsheds of all sizes. However, all of this explanation resulted from the only variable in the model, dense forest land use in the 90m wide strip on each side of the local streams. As expected, riparian dense forest land use was positively and significantly correlated with QHEI. Interestingly, none of the 30m riparian land uses got selected in the stepwise build of this regression model, and the standard error of the estimate improved only marginally. This led us to the next logical step of testing the relationship between stream habitat and landscape and land use factors only in the immediate and local contexts of the QHEI sample locations.

In the next group of models, only the 125 small-size localsheds were included. As expected, the models explained more variation in QHEI as the resolution of the watersheds improved when we included only the small-size localsheds.

Table 11: Regression model results with 125 small-sized non-intersecting localsheds

Independent variables	Unstandardized coefficients		Standardized coefficients	t	Prob. > |t|	Sign OK?	Significant at 5%?
	b_i	Std. Error	Beta				
Intercept	72.974	3.077		23.715	0.000		
Interaction of Row Crops in 30m buffer and Headwater	-0.204	0.075	-0.313	-2.714	0.008	Yes	Yes
Low-density Resid. in 30m buffer	-0.577	0.126	-0.330	-4.566	0.000	Yes	Yes
High Localshed Slopes (>6%)	0.336	0.140	0.184	2.406	0.018	Yes	Yes
Row Crops in 30m buffer	-0.184	0.088	-0.253	-2.102	0.038	Yes	Yes

N = 125 Prob > F = 0.000 Std.Error of Estimate: 13.910
R-Squared: 0.45 Adj. R-square: 0.43

Riparian row crops and residential land uses in the 30m buffer strip were negatively and significantly correlated with QHEI. Also, statistically speaking, the greater the area with high slopes in the localshed, the better the stream habitat quality scores, in general, for measurements made at that localshed's mouth. Interestingly, the statistically significant interaction between riparian row crops and headwater streams indicates that riparian row crops have a greater adverse effect on the habitat quality for streams with a Strahler stream order of 1 or 2, than riparian row crops in streams further downstream in their basins.

These results further improved after dealing with influential residuals. After a rigorous analysis of the residuals in the earlier model with 125 small-size non-intersecting localsheds, it was decided to test the regression with a reduced dataset of 117 localsheds. The results improved substantially, justifying the exclusion of 8 influential residuals, or about 6% of the dataset, for the sake of higher explanatory power and fit of the model. The R^2 statistic increased 27%, from 45% to 57%, and the standard error of estimate for QHEI improved by 11%, from 13.9 to 12.4. Figure 7 shows a map with the major outlier locations.

Table 12: Regression model results with 117 small-sized non-intersecting localsheds

Independent variables	Unstandardized coefficients		Standardized coefficients				
	b_i	Std. Error	Beta	t	Prob. > |t|	Sign OK?	Significant at 5%?
Intercept	74.938	2.869		26.117	0.000		
Interaction of Row Crops in 30m buffer and Headwater	-0.230	0.068	-0.354	-3.409	0.001	Yes	Yes
Low-density Resid. in 30m buffer	-0.702	0.123	-0.377	-5.708	0.000	Yes	Yes
High Localshed Slopes (>6%)	0.384	0.133	0.204	2.898	0.005	Yes	Yes
Row Crops in 30m buffer	-0.213	0.080	-0.291	-2.668	0.009	Yes	Yes

N = 117 *Prob > F = 0.000* *Std.Error of Estimate: 12.403*
R-Squared: 0.57 *Adj. R-square: 0.55*

The reduced dataset also improved the statistical significance of each of the variables in the relationship formulated in the previous model with the whole non-reduced dataset. The strength of the variable's relationship with QHEI, measured by the absolute value of unstandardized coefficients, also increased for each of the independent variables in the model, when a few influential cases were removed. Riparian row crops and residential land uses in 30m buffer are still negatively correlated with QHEI, and higher localshed slopes are correlated positively with QHEI. The presence of row crops in riparian zones of a headwater localshed is likely to have greater adverse impact on the stream's habitat quality than riparian row crops in a non-headwater localshed.

Strategy II: Similar-size Localsheds

Different layers were created and managed for each of these 580 similar-size localsheds using GIS. The average area of these localsheds was 326 acres, with almost 70% of the cases between 200 - 450 acres. Data was generated for layers such as streams, roads, population, and land use, for both riparian as well as whole localshed boundaries. Generating the GIS database, consisting of 277,000 files across 18,000 directories, took more than 20 hours on our machine.

Table 13: Regression model results with 580 similar-size localsheds

Independent variables	Unstandardized coefficients b_I	Unstandardized coefficients Std. Error	Standardized coefficients Beta	t	Prob. > |t|	Sign OK?	Significant at 5%?
Intercept	63.362	1.800		35.197	0.000		
Row Crops in 30m buffer	-0.192	0.029	-0.269	-6.732	0.000	Yes	Yes
Dense Forest in 90m buffer	0.227	0.038	0.246	6.019	0.000	Yes	Yes
Headwater (dummy)	-6.449	1.354	-0.165	-4.764	0.000	Yes	Yes
High Localshed Slopes (>6%)	0.195	0.045	0.159	4.328	0.000	Yes	Yes

N = 580 Prob > F = 0.000 Std.Error of Estimate: 13.992
R-Squared: 0.35 Adj. R-square: 0.34

The large numbers of localsheds probably induced extraneous noise into the database, therefore, a systematic study of the residuals was performed to pare down the dataset by less than 7% of the cases. The exclusion of 38 cases from the analysis improved the results and increased the explanatory power of the model. R2 improved 14% from 35% to 40%, and the standard error of the estimate improved 12% from 13.99 to 12.41.

Table 14: Regression model results with 542 similar-size small localsheds

Independent variables	Unstandardized coefficients b_i	Unstandardized coefficients Std. Error	Standardized coefficients Beta	t	Prob. > |t|	Sign OK?	Significant at 5%?
Intercept	64.645	1.650		39.183	0.000		
Row Crops in 30m buffer	-0.200	0.027	-0.291	-7.412	0.000	Yes	Yes
Dense Forest in 90m buffer	0.230	0.035	0.265	6.582	0.000	Yes	Yes
Headwater (dummy)	-6.520	1.244	-0.179	-5.240	0.000	Yes	Yes
High Localshed Slopes (>6%)	0.192	0.041	0.170	4.635	0.000	Yes	Yes

N = 542 Prob > F = 0.000 Std.Error of Estimate: 12.341
R-Squared: 0.40 Adj. R-square: 0.40

Discussion of Small Watershed Results

Our models clearly show that a variety of indicators of land use intensity in the riparian zone provide statistically significant explanation of the QHEI rating for those streams. Larger-scale (smaller area) appears to make a significant difference in the model's explanatory power. Of particular importance in this region is the presence of cropping in the riparian zone. In most of the models, this indicator provides the strongest indicator of changes in the QHEI.

Despite the changes in scale and localshed delineation scheme, the overall explanatory power of these models is not so high that we may use them for prediction with much confidence. The variables in these models are theoretically and statistically significant, so the study is useful for explanatory purposes. However, more work is needed to improve the explanatory power of the models so that they may be used for forecasting or prediction. Generally, less than 50% of the variation in QHEI is explained by the independent variables in our study. There may be several reasons for this. First, the remote sensing and other secondary indicators of riparian zone activity may still be too coarse to identify differences in habitat quality among stream reaches. Higher resolution data should be tested to ascertain whether other differences in the riparian zone could be identified.

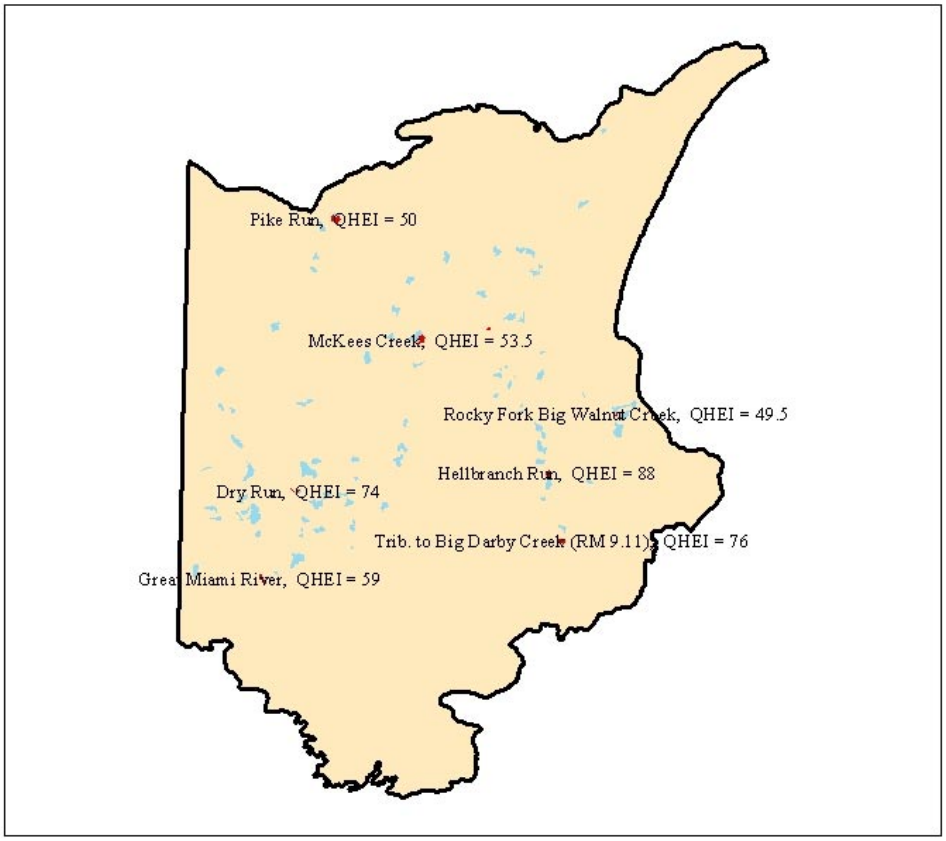
Second, the static nature of the data sources may reduce their explanatory power. The QHEI rating represents the accumulation of events in a dynamic environment over a period of several years. Our datasets represent a snapshot of conditions in that area at one or two times during that period. Major shifts in uses or management of the land over the longer period could explain the differences in habitat quality in two stream segments that might otherwise be projected to be of the same quality. Time series land cover data combined with additional sensing of the occurrence of flood events, major sedimentation episodes or other storm related phenomena might help to further explain the differences in the status of the streams.

Third, the time periods of the land use and the QHEI datasets may also be of concern in this study. The QHEI sampling sites cover the 1989-1995 time period, while the Landsat images, from which the land use dataset was derived, are from the 1988-1994 time period. It is entirely possible that there may be a critical time lag between land use alterations and their accumulated, visual, and measurable impact on the stream habitat.

Fourth, there may also be the possibility of measurement issues in the various datasets. The theory on stream habitat suggests that geomorphological attributes of streams such as sinuosity may be critical in explaining the overall habitat quality of a reach segment. However, in reality, sinuosity is a directly related to the resolution of the dataset as coarser datasets may incorrectly report different (usually lower) sinuosity values for the same stream. Data on other geomorphological variables, such as bankfull width, and depth are not available for large regional areas such as the ECBP ecoregion. The accuracy of both the land use classifications is not known. Moreover, very few studies have tried to explain stream habitat quality as the dependent variable, therefore, the validity of QHEI also has not been rigorously tested across larger regions.

Finally, there may be individual, unique circumstances that have produced unexpected habitat condition in particular watersheds. Our experiments with removing outliers from the datasets illustrate that there are a relatively small number of streams where such unusual events may be occurring. Secondary datasets such as ours cannot capture these unique circumstances.

Because the small-scale models were not able to explain more than 50% of the variance in the QHEI, we chose not to use them in our web-based modeling tools. However, these modeling experiments do indicate that it may be possible in the future to derive models that will allow a reasonably accurate prediction of the changes in habitat quality that occur as a result of human activities in the riparian zone. By investigating the use of emerging detailed remote sensing data, it may be possible to derive and implement such a model in the future.



The 8 residuals from Linear Regression for the 125 Non-intersecting Small Localsheds

Note: These influential cases were based on criteria involving Cooke's plots, Studentized residual plots, Partial regression plots, COVRATIO-residual plots, and SDFBETA plots for each residual observation. Labels show stream name and QHEI value.

Legend

- ECBP ecoregion
- Non-intersecting Small Localsheds
- Abnormal/Influential Localsheds

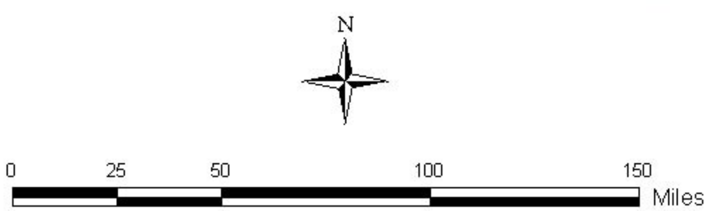


Figure 7

Deploying the Model on the Web

Using the input datasets and the results from the regional models, we produced an interactive website that allows the user to explore the datasets, to learn about the nature of watersheds and how to measure their health, and the potential impacts of future developments on watershed scale biological quality. The website resides on a server at The Ohio State University Center for Mapping. The website URL is:

http://tycho.cfm.ohio-state.edu

In this section of the report, we discuss the overall website design, the nature of the interactive mapping tool we created, and the manner in which we created watershed quality forecasts.

Website Design

The overall website was designed with the philosophy that visitors will not know very much about watersheds, environmental processes, or the terminology associated with their measurements and modeling. Accordingly, we wanted to create a site that is relatively easy to navigate, provides visual illustrations of the conditions we are explaining, and provides access to further information through a glossary of terms and links to other, related sites at U.S. EPA, Ohio EPA, and other locations.[2]

The website is divided into sections that allow users to explore a range of information. The introductory level introduces the concept of a watershed and its relationships to human activities. Links are made to illustrations of the hydrologic cycle and definitions of those basic concepts are available in the main text and the glossary.

The second section guides people through the concept of what constitutes a healthy watershed. The major characteristics of such a watershed are described and illustrated with embedded photographic examples. Thus, for example, people see that a healthy watershed is lined with vegetation that shades the surface to keep temperatures down and provide a source of food. They see that a natural stream is not straight but meanders, that it is a sequence of pools and riffles, and that it is in a dynamic balance with the waters that flow through it.

Given these illustrations of a "good" stream, the user can then review some of the measures of water quality and water pollution. Both chemical and biological measures are discussed along with illustrations of the sources of pollutants and the nature of the biotic indices such as the IBI, ICI, and QHEI. As they explore these areas, they can jump to and from the glossary getting definitions of the terms they do not understand. They can also see illustrations of streams where particular kinds of pollution problems are occurring.

The last sections of the site allow users to explore data about their watersheds. Here, they can start the interactive mapping tool and explore information about where there are watersheds in their part of Ohio and what the conditions of those watersheds are. The nature of this tool and the datasets they can explore with it are discussed in detail below.

Finally, they can explore information about selected watersheds in the ECBP where we applied the regional model of IBI to illustrate the potential impacts of three major growth and management scenarios. Live links are made to the Ohio EPA site where they can see a table of the sources of impairment in each watershed. They can then explore the future scenarios of their watersheds seeing the basis of our model and how it can be applied to gain insights into the potential future of their watershed. These scenarios and the application of the model are discussed in detail below.

---

[2] Credit must be given to Kelly Dufour who provided critical review of the basic material, created much of the site content, and undertook almost all of the website design. Without her hard work we would not have been able to complete the project.

Interactive Mapping Tool

The interactive mapping tool is intended to allow people to explore real data about their watersheds. It was built using MapObjects Internet Map Server, Visual Basic, JavaScript, and HTML. It currently runs on a dual-processor Windows NT 4 platform with 512MB RAM and Microsoft Internet Information Server 2.0 as the web server. The application can be viewed from both Internet Explorer Version 5 or higher and Netscape Version 4 or higher browsers. A server-side script also improves speed and eases machine management by automatically deleting older files once a certain threshold size is increased.

In terms of functionality, the user is able to display or hide both vector (water quality sampling sites, streams), and raster data (land use image), change the symbols (symbol shape, color, and transparency), or zoom in and out, or simply pan the digital interactive GIS data-based map. Other visualization features include map scale in different units, legend for raster map categories, and an overview or inset map in smaller scale.

The map also provides functionality for text- and spatial-attribute querying. The Identify feature generates a new pop-up window with dynamically organized fields and values based on the feature near the user's click on the map. The Hyperlink feature takes the user to another site on the internet with more detailed attributes regarding the feature near the user's click. If only partial values of some names are known (say, a county starting with "Ma…"), then the Find feature's text-based search capabilities can be utilized to zoom in to the feature if there is a single match, or to view a dynamically created pop-up table of all matching features if there are two or more matches. The Spatial Search tool lets the user perform simple spatial overlays, such as "find all counties within Great Miami River basin" or "find all NPDES sites in Franklin county", interactively. The results are both map-based (zooming in on the resulting features) and tabular (pop-up window with name and value fields for the matching features).

Table 15: Interactive Map Tool Functionality

Functionality	Implementation
Toggle data layers on/off	Click checkbox in Theme column of Legend, then click Update Map button.
Change color of a data layer	Select a color from the drop-down menu in the same row as the data layer, in the Color column of Legend, then click the Update Map button.
Change display symbol of a data layer	Select a symbol from the drop-down menu in the same row as the data layer, in the Symbol column of Legend, then click the Update Map button.
Redraw the map	Click the Update Map button.
Zoom In	Select the Zoom In radio button in the Spatial Operations section, select the zoom in factor (say, 4 times) from the drop-down menu, and click near the desired location on the map.
Zoom Out	Select the Zoom Out radio button in the Spatial Operations section, select the zoom in factor (say, 8 times) from the drop-down menu, and click near the desired location on the map.
Pan or shift the map	Select the Pan radio button in the Spatial Operations section, and click near the desired location on the map to center the map on the click position.
Identify features	Select the Identify radio button in the Spatial Operations section, select the appropriate theme (say, "Samples") from the drop-down menu, and click near the desired feature on the map. It might help to zoom in first.
Hot Link to Basins	Select the Hot Link radio button in the Spatial Operations section, and click the desired basin on the map. It might help to toggle the Basins theme on first.
Find feature	Select the Find radio button in the Spatial Operations section, type the text for the feature to search (say, "m") in the textbox, select the appropriate theme from the drop-down menu (say "Counties"), and click anywhere on the map. For this example, a table showing data about all counties starting with the letter "m" will popup.
Spatial Search	Select the Spatial Search radio button in the Spatial Operations section, select the appropriate theme from the 1st drop-down menu for the features to search (say "Cities"), select the appropriate theme from the 2nd drop-down menu for the spatial context in which to search the features (say "Counties"), and click on a feature, of the theme selected in the 2nd menu, on the map. For this example, a table showing data about all cities in the clicked county will popup and the map will zoom in and center on the first such feature.
Change scale units displayed	Select the appropriate scale units from the drop-down menu just below the map, and click the Update Map button.
See the default map	Click the "View Default Map" hyperlink.
Get help	Click the "Help" hyperlink.
Email comments	Click the "Comment" hyperlink. It might help to change the settings in your browser to open the desired email program.

Table 16: Finding solutions to simple questions using the Interactive Map Tool

Question	Solution
How can I change the layers displayed?	It is not possible to add more layers but it is possible to toggle the existing layers on and off. Their colors and symbols can also be changed.
How can I see the layer hidden below?	Select "Transparent" from the drop-down menu in the visible layer's Symbol column in Legend. This will show the layer hidden below. If "Counties" is hiding "Basins", make "Counties" transparent.
How can I find the attributes for a point, line, or polygon?	Use the Identify radio button.
How can I find more information on a basin?	Use the Hyperlink radio button.
How can I find a particular feature out of the so many displayed? Or, find a feature whose name I only know partially, say "How can I find the county whose name starts with "f"?	Use the Find radio button. Then use the Identify tool to see attributes for that feature.
How can I see results from a simple spatial overlay, such as "How many and which counties are in the Big Darby basin"?	Use the Spatial Search radio button.

Problems and future applications

The tool can be improved with more detailed design enhancements such as drag-rectangle based zoom in and zoom out functions, or scale-sensitive map layer toggling in order to make some layers visible only within certain scale ranges.

Although it is possible to display and distribute GIS-based data on the Internet, GIS-based distributed data analysis is still not a reality. It is not entirely possible to quickly integrate data from different sources over the web, in different map projections, and perform spatial analysis without caching or storing it locally. The current technology for distributed GIS-analysis suffers from serious bandwidth and networking problems because true GIS functions can only be analyzed with geometry-aware vector streaming data, not on the current standard static pixel-based images. Therefore, the best and most feasible way to perform GIS-based modeling is to let users download data, preprocess it, integrate with other data sources in a proprietary GIS, and then rely on the vendor-specific functionality or add other functionality programmatically to the local GIS server. This is also prudent if the analysis is supposed to be stored as a project locally. Storing locally analyzed projects on distant servers, like some recent Application Service Providers (ASP) enable the user to do, is not possible with GIS software applications at the current time. Using ArcExplorer, ESRI's freely distributed browser is also not a complete solution as it does not offer spatial overlay tools such as clipping, and it requires a one-time download and installation. However, local storage and analysis entails that the user be GIS-literate and have access to GIS software.

Scenarios of Watershed Development and Its Impacts

Three scenarios were created and used with the regional model to provide website users insights into the relationships between human activities on their watersheds and the impacts those activities have on the biological quality. The scenarios use historical growth rates to generate changes in population that are then tied to proportional changes in dense urban land use, a decline in stream habitat, and an increase in point source pollution loads resulting from that growth. The instream habitat decline could not be explained more than 50% in our models for QHEI. It is arbitrarily set , in relation to population growth, depending on the given scenario. The growth rate is derived by applying historical rates of population growth for the fastest growing county in the region of the watershed over the next 50 years.

The rationale for the three scenarios are described in greater detail in Table 14. Adjustments are made to the model forecasts to reflect policies that would offset the impacts of growth. The first scenario assumes increasing urbanization with no additional growth or land use regulations. In this scenario, the impacts of habitat change are reduced by 50% of the growth rate since stream habitat will be negatively affected by unregulated growth. For the scenario with some land use mitigation, although the percent of urban growth and point source pollution increases in direct proportion to the growth rate, the stream habitat improves by 40% as a result of mitigation efforts. For the growth control scenario, it is assumed that not only will there be reductions in habitat change but the amount of urban land use and point source pollution will also decline. These changes are estimated at 30% of the original growth for urban and point source pollution and 40% change in habitat from land use mitigation.

The users of the scenarios are told exactly the assumptions underlying the forecasts as well as the errors associated with the overall regional model. They are also given the model equation so that can see the nature of the relationships derived in the modeling effort.

To make the final forecasts, we applied the model to each of the subbasins in the ECBP watersheds where we had data. Since there were errors associated with the statistical model, we adjusted the starting point source indicator and habitat values so that the starting value for the predicted IBI in each subwatershed was within plus or minus 10% of the actual starting value. We then applied the forecasts for changes in the urban land use, substrate and pool, and point source indicator to create each of the scenarios.

We then created maps illustrating the impacts of growth on the watersheds over time. The resulting subwatersheds forecast IBI value was placed into the following classification:

Exceptional *	IBI of 50-60
Very Good	IBI of 42-49
Good	IBI of 34-41
Fair	IBI of 27-33
Poor	IBI of 17-26
Very Poor	IBI of < 17
Not Sampled	No data for subbasin

*IBI categories were slightly modified from the original Ohio EPA designations.

An example of the forecast maps for the Big Darby Watershed are shown in Figures 8 and 9. On the website, users can choose which major watershed to explore. When they make this choice, they are taken to the summary page at the Ohio EPA website showing the sources of impairment for that watershed. They can then explore each of the model scenarios over time or compare the end-date results of the three scenarios compared to the current status of the watershed.

Links are provided to detailed descriptions of the model assumptions, sources of errors, and scenarios. Users can also go back to the main web pages so that they can review definitions of the water quality measures or go back to the glossary as needed. In this way, people are free to explore as they wish.

Table 17: Basin Scenarios

Descriptions of Land Use Scenarios

All scenarios are based on the assumption of being able to predict the impact of population change on a river's water quality. Water quality is measured by examining physical river characteristics, such as the composition of solid materials on the river's floor, its width, current, and depth, measuring quantities and effects of chemical pollutants, and by looking at biological parameters related to the number and type of fish species in the river. For the purposes of this study, 3 scenarios have been created and are described below.

Scenario 1 - Increased Urbanization with No Regulations

Although some municipalities in the watershed have zoning and subdivision controls, they do not typically have enforceable limitations on what happens with the stormwater runoff created as a result of urbanization and whether habitat alteration along the streams takes place. Similarly, there are no overall controls on the amount of growth that is consistent among communities that share the watershed. As a result, this scenario assumes that growth will follow historic rates and proceed at the same rate with little or no control on location or stormwater runoff.

Scenario 2 - Increased Urbanization with Land Use Mitigation Efforts

In this scenario, the negative impact of land use related activates on water quality and river characteristics are recognized. Thus, in order to minimize the harmful aspects of urban and agricultural development on water quality, some mitigation efforts and controls are implemented on both urban and agricultural endeavors in the transitional zones, those areas between the upland and aquatic environments. Some of these measures include reducing the amount of impervious surfaces when and wherever possible, constructing wetlands and stormwater retention ponds, and improving construction practices so as to minimize the amount of erosion during the clearing and construction of sites.

Scenario 3 - Implementing Growth Controls

The third scenario imposes growth restrictions and controls on all development activity in the area, with new urban developments being banned in all transitional zones, those areas between the upland and aquatic environments. Besides incorporating the mitigation and design efforts of Scenario 2, strict buffer zones are also established between all water bodies (rivers, streams, ponds, and lakes) and the urban and agricultural areas surrounding them. Land use regulations and techniques, such as conservation easements and land purchases, are also utilized to limit development in other environmentally critical and sensitive areas.

Figure 8: Big Darby Watershed Scenarios Summary Maps

Current Status | **Scenario 1 Unregulated Growth** | **Scenario 2 Mitigation incorporated** | **Scenario 3 Mitigation and Growth Control**

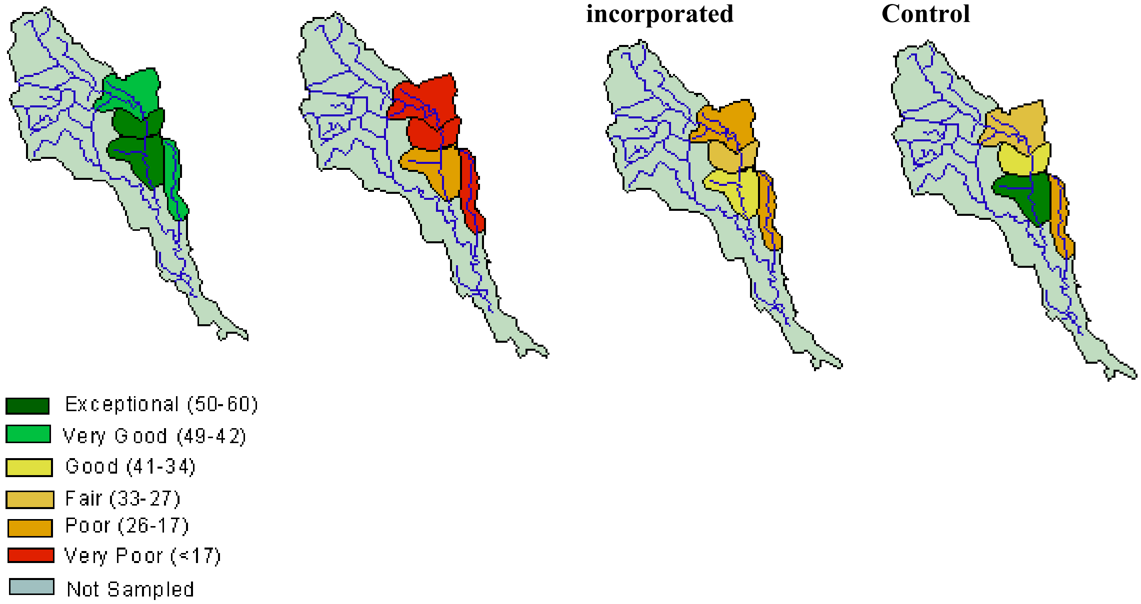
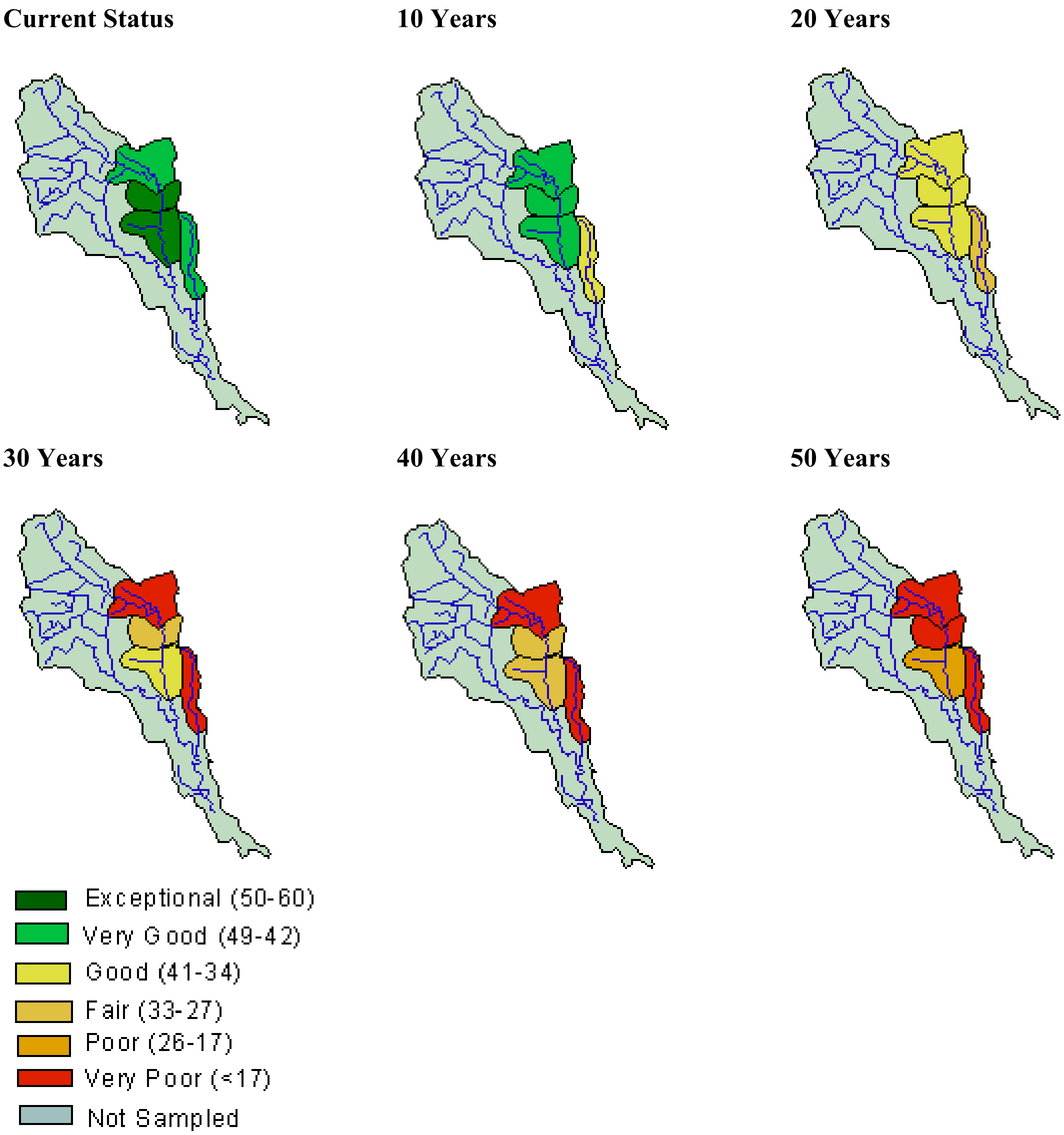


Figure 9: Big Darby Watershed Scenario 1 Maps



Summary

This project used the empirical results of regional modeling of the relationships between human stressors and watershed biological quality to derive an interactive website. The site allows users to explore basic information about watershed quality, datasets relating to watershed characteristics in the Eastern Cornbelt Plains Ecoregion of Ohio, and scenarios of the impacts of growth on watershed quality.

Detailed analysis of larger scale (smaller watershed) and localshed data provided some additional insights regarding the relationships between urban indicators and stream habitat quality. There are strong indications that development within the riparian zone has a major impact on habitat quality as would be expected. However, the empirical results, though significant, did not account for enough of the overall variance in quality to warrant the creation of an interactive modeling tool for this scale of analysis.

The regional model explaining approximately 70% of the variance in the IBI for the watersheds was used to create a series of scenarios of watershed change. The scenarios illustrate the empirical relationships found in the regional model and highlight the nature of the policy decisions that would need to be made to offset deterioration in the watershed conditions as a result of future development.

The website also allows users to explore concepts of watershed quality and how it is measured. The information can provide the basis for educating the public about watershed concepts and the relationships between human stressors and watershed biological quality. The interactive mapping tool allows people to explore information about their own watershed conditions, showing them distributions of population, land use, point source pollution loads, and biological quality in their watersheds. The site also provides live links to the related watershed exploration information at the Ohio EPA.

It is hoped that future research will allow a refinement of the models and the interactive mapping tools that will allow users to interactively explore the impacts of their decisions on watershed biological quality. In the meantime, the current site will help citizens, local planners, and public officials to understand the impacts of their activities on watershed quality.

References

Gammon, J. R. 1976. The fish populations of the middle 340 km of the Wabash River. Tech. Report No. 86. Purdue University. Water Resources Research Center, West Lafayette., IN.

Gammon, J. R., A. Spacie, J. L. Hamelink and R. L. Kaesler. 1981. Role of electrofishing in assessing environmental quality of the Wabash River. pp 307-324. In: Ecological assessments of effluent impacts on communities of indigenous aquatic organisms. ASTM STP 703, J. M. Bates and C. I. Weber (eds.) Philadelphia, PA

Gordon, Steven I., Ward, Andy D., White, Dale A., Majumder, Sarada, and Chen, Mi (2000). Integrating Planning, Forecasting, and Watershed Level Ecological Risk Assessment Techniques: A Test in the Eastern Cornbelt Plains Ecoregion. Report to the U.S. Environmental Protection Agency on Grant Number R824769. Columbus, OH: The Ohio State University, January 29, 2000.

Gordon, Steven I. and Majumder, Sarada (1997). "Ecological Risk Assessment of Aquatic Biological Diversity Using a GIS Database, Water Resources Education Training, and Practice: Opportunities for the Next Century, American Water Resources Association, Proceedings of the AWRA/UCOWR Symposium, June 29-July 3, 1997, p. 399-406

Gordon, Steven I. and Majumder, Sarada (2000). "Empirical Stressor-Response Relationships for Prospective Risk Analysis in the Eastern Cornbelt Plains Ecoregion", Environmental Toxicology and Chemistry, 19,4(2), p. 1106-1112.

Hair, Joseph F. Jr., Rolph E. Anderson, Ronald L. Tatham, et. al., 1998. *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.

Karr, J. R., L. A. Toth, and G. D. Garman. 1983. Habitat Preservation for Midwest Stream Fishes: Principles and Guidelines. United States Environmental Protection Agency, Environmental Research Laboratory Corvalis, Corvalis, Oregon. EPA-600/3-83-006 February 1983.

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

Karr, J.R. and D.R. Dudley. 1981. Ecological perspective on water quality goals. Environmental Management 5(1):55-68.

Karr, J.R. 1991. Biological Integrity: A long-neglected aspect of resource management. Ecological Applications. 1(1):66-84.

Ohio EPA 1987a. Biological criteria for the protection of aquatic life: Volume I. The role of biological data in water quality assessment. Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, OH.

Ohio EPA 1987b. Biological criteria for the protection of aquatic life: Volume II. Users manual for biological field assessment of Ohio surface waters. Division of Water

Ohio EPA 1989a. Ohio EPA manual of surveillance methods and quality assurance practices, updated edition. Division. of Environmental Services, Columbus, OH.

Ohio EPA 1989b. Addendum to biological criteria for the protection of aquatic life: Users manual for biological field assessment of Ohio surface waters. Division of Water Quality Planning and Assessment, Surface Water Section, Columbus, OH.

Ohio EPA 1989c. Biological criteria for the protection of aquatic life: Volume III. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Division of Water Quality Planning and Assessment, Columbus, OH.

Ohio EPA 1990. The use of biological criteria in the Ohio EPA surface water monitoring and assessment program. Division of Water Quality Planning & Assessment, Ecological Assessment Section, Columbus, OH.

Omernik, J.M. 1988. Ecoregions of the conterminous United States. Ann. Assoc. Amer. Geogr. 77(1): 118-125.

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

Rankin E.T., Yoder, C.O. 1991. Calculation and Uses of the Area of Degradation Value (ADV). Technical Report. Division Water Quality Planning and Assessment, Columbus, OH, USA.

Rankin, E.T. 1995. Habitat Indices in water resource quality assessments, in W.S. Davis and T. Simon (eds). Biological assessment and criteria: tools for risk-based planning and decision making. Ann Arbor, MI: CRC Press/Lewis Publishers.

U.S. Department of Agriculture (USDA), 1994. *State soil geographic (STATSGO) data base: Data use information.* Soil Conservation Service, National Soil Survey Center. Miscellaneous publication number 1492.

USEPA. 1988b. Ecoregions of the Upper Midwest Staes. EPA/600/3-88/037.

USEPA. 1989. Regionalization as a Tool for Managing Environmental Resource. EPA 1600/3-89/060.

USEPA, 1991. Quality Assurance Guidelines for Modeling Development and Application Projecta: a Policy Statement. Environmental Research Laboratory, Duluth, Minnesota, November, 1991.

Yoder, C. O. and E. T. Rankin. 1995. Biological response signatures and the area of degradation value: new tools for interpreting multi-metric data, in W.S. Davis and T. Simon (eds.). Biological assessment and criteria: tools for risk-based planning and decision making. CRC Press/Lewis Publishers, Ann Arbor.