

April 01, 2005

## Evaluating Single and Multiple Stressors in Watershed Risk Assessment

Jessica Keye Brooks  
*Northeastern University*

Vladimir Novotny  
*Northeastern University*

---

### Recommended Citation

Brooks, Jessica Keye and Novotny, Vladimir, "Evaluating Single and Multiple Stressors in Watershed Risk Assessment" (2005). Center for Urban Environmental Studies Publications. Paper 3.

**Center for Urban Environmental Studies  
Northeastern University, Boston, MA 02115**

---

**TECHNICAL REPORT NO. 3**

**EVALUATING SINGLE AND MULTIPLE  
STRESSORS IN WATERSHED RISK ASSESSMENT**

**Jessica Keye Brooks, MSc.**  
Research Associate

**Vladimir Novotny, Ph.D., P.E.**  
Primary Investigator

Project sponsored by the Grant No. R83-0885-010 to Northeastern University from the  
US EPA/NSF/USDA STAR Watershed Program

**Bernice Smith**  
EPA Project Officer

**Iris Goodman**  
EPA Program Director

Boston, Massachusetts  
April 2005



## **Abstract**

The Index Biological Integrity (IBI) can be used as an indicator of overall water body health and may be predicted by using artificial neural networks (ANN). Changes within a watershed, such as urbanization, can affect local water bodies by increasing chemical concentrations or degrading habitat. Risk by single stressors on all fish and benthic invertebrate species as well as key IBI metrics was evaluated using toxicity data and Maximum Species Richness. However, the risk from combined stresses may not be directly additive, as is assumed by these methods. ANN technology was used to consider the overall symbiotic effect of multiple stressors on IBI. Data for the purpose of this study was extracted from large database that combines chemical, habitat, and biological information for streams in Ohio. This database was used to train a multilayer, backpropagation network to estimate the IBI value based on water, sediment, and habitat quality. Inputs to the network included water quality parameters, such as chemical and dissolved oxygen concentrations, sediment quality, expressed by invertebrate community index, and the qualitative habitat evaluation index. Results of IBI prediction may be useful to watershed planners and decision makes evaluating new projects and restoration opportunities.

## **Acknowledgements**

The funding for this research was received through a grant from the US Environmental Protection Agency/National Science Foundation/US Department of Agriculture STAR Watershed Program. The findings and conclusions expressed in this thesis are those of the author and not of the funding agencies. The finding has been greatly appreciated.

# Table of Content

<b>Abstract.....</b>	<b>i</b>
<b>Acknowledgements.....</b>	<b>ii</b>
<b>Table of Content.....</b>	<b>iii</b>
<b>List of Figures.....</b>	<b>iv</b>
<b>List of Tables.....</b>	<b>v</b>
<b>1 Introduction.....</b>	<b>1</b>
1.1 Ecological Integrity.....	1
1.2 Ecological Risk Assessment.....	4
<b>2. Toxic Response.....</b>	<b>11</b>
2.1 Evaluating Toxicity Data.....	11
2.2 Estimation of Risk Response Function Using Toxicity Data .....	16
<b>3. Maximum Species Richness Lines.....</b>	<b>21</b>
3.1 Developing Maximum Species Richness Lines.....	21
3.2 Risk Calculation Using Maximum Species Richness Lines.....	23
<b>4. Artificial Neural Networks.....</b>	<b>27</b>
4.1 Introduction to Artificial Neural Networks.....	27
4.2 Network Data.....	29
4.3 Preliminary Network Analysis.....	31
<b>5. Conclusions and Recommendations.....</b>	<b>37</b>
5.1 Risk Calculation.....	37
5.2 Addressing Limitations of Risk Calculation.....	38
5.3 Artificial Neural Networks.....	39
5.4 Linking Biology and Watershed .....	39
5.5 Recommendations for Future Work .....	40
<b>References.....</b>	<b>43</b>

## List of Figures

Figure 1.1	Watershed, land use changes and other impacts on water quality (adapted from Novotny 2004).....	1
Figure 1.2	Increased % impervious area in a drainage basin correlates with decreased macroinvertebrate quality (Schueler 1994, Novotny 2004)...	2
Figure 1.3	Schematic of multilayer risk propagation through hierarchical model...	3
Figure 1.4	Patterns of Maximum Species Richness Lines for various stressors (Novotny 2004).....	8
Figure 2.1	Toxic response curve for cadmium evaluated at 50 mg CaCO <sub>3</sub> /l.....	15
Figure 2.2	Ecological risk from stormwater impacts (Novotny and Witte 1997)....	16
Figure 3.1	Maximum Species Richness line for lead.....	22
Figure 3.2	Calculating the risk of extinction using maximum species richness.....	23
Figure 3.3	Maximum species richness for lead: Count decreasing with increasing concentration.....	24
Figure 3.4	Maximum species richness for pH: MSRL displays a peak at pH 7-8...	24
Figure 3.5	Maximum species richness for fish vs zinc: Non-linear response.....	24
Figure 3.6	Maximum species richness for copper and zinc.....	25
Figure 3.7	Comparing cadmium MSRLs for invertebrates and fish.....	25
Figure 3.8	Comparing responses of intolerant and tolerant fish to copper.....	25
Figure 4.1	A model neuron (Demuth and Beale 1992).....	27
Figure 4.2	A model of a layered network (Demuth and Beale 1992).....	28
Figure 4.3	Examples of simulations results for Group 1 parameters.....	33
Figure 4.4	Clusters resulting from Self Organizing Map analysis of Ohio database from Virani (2005).....	34
Figure 4.5	Calibration and simulation results for network trained with Group 2....	35
Figure 5.1	Developing risk functions from Maximum Species Richness.....	38

## List of Tables

Table 2.1	Hardness Conversion Factors.....	12
Table 2.2	Data Count for Fish.....	13
Table 2.3	Data Count for Benthic Invertebrates.....	13
Table 2.4	Goodness of Fit for Distributions of Cadmium LC <sub>50</sub> s.....	15
Table 2.5	Mean and Standard Deviation of Concentrations at Site S17830.....	17
Table 2.6	Risk calculations for fish.....	17
Table 2.7	Risk calculations for benthic invertebrates.....	18
Table 3.1	Description of Data used from Ohio Database.....	22
Table 4.1	Habitat Parameters Available in Ohio Database.....	30
Table 4.2	Network Training Group 1.....	32
Table 4.3	Network Training Group 2.....	35

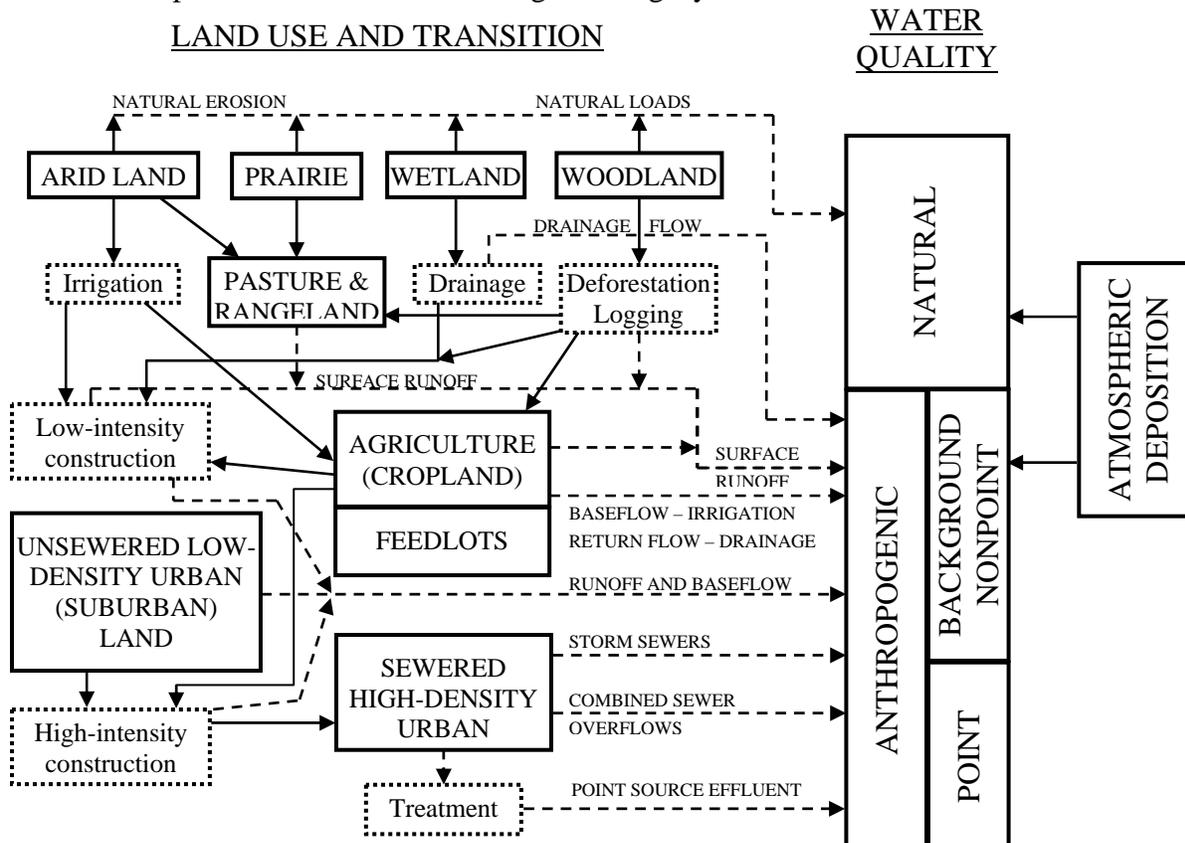


# 1 Introduction

## 1.1 Ecological Integrity

The United States' Clean Water Act of 1972 was designed to protect the ecological integrity of the Nation's waters. Ecological integrity can be defined as three parts, chemical, physical, and biological integrity. Physical integrity describes habitat conditions of the water body that would support a balanced biological community. Chemical integrity describes the chemical composition of water and sediments that would not be injurious to the aquatic biota or limit other uses of the water body such as recreation. Biological integrity describes a composition of aquatic organisms that would be balanced and would resemble that of unimpaired water bodies in the same ecoregion (Davis 1995; Karr 1995; Karr and Chu 1999). Although there are three separate measures, chemical and physical integrity are both defined with the purpose of maintaining, to some degree, biological integrity.

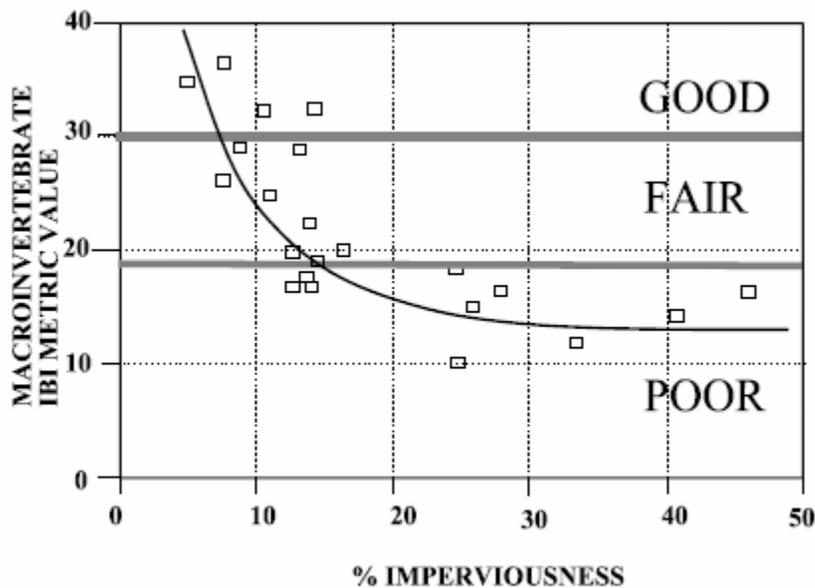
The physical and chemical conditions within a water body are dependent on the actions within the contributing watershed area. Human activities within a watershed alter water body sources and conditions. The interactions between human alterations and the changes in the water body can become quite complex, see Figure 1.1. Any alteration to a watershed presents a risk that the biological integrity will be reduced.



**Figure 1.1 Watershed, land use changes and other impacts on water quality (adapted from Novotny 2003)**

Since the inception of the Clean Water Act great reductions have been made to many point-sources of pollution primarily through the building of waste-water treatment facilities for municipal and industrial waste. Implementation of Best Management Practices (BMPs) has also been successful in reducing many non-point sources of pollution. However, ecological integrity has not yet been reached in many areas, and non-point sources are still responsible for about 50% of the remaining water quality problems (USEPA 2000). Holistically, pollution is defined as any change that causes impairment to the integrity of a system. Using this definition, inputs from human activity that change flow regime and habitat features as well as chemical quality are all considered pollution.

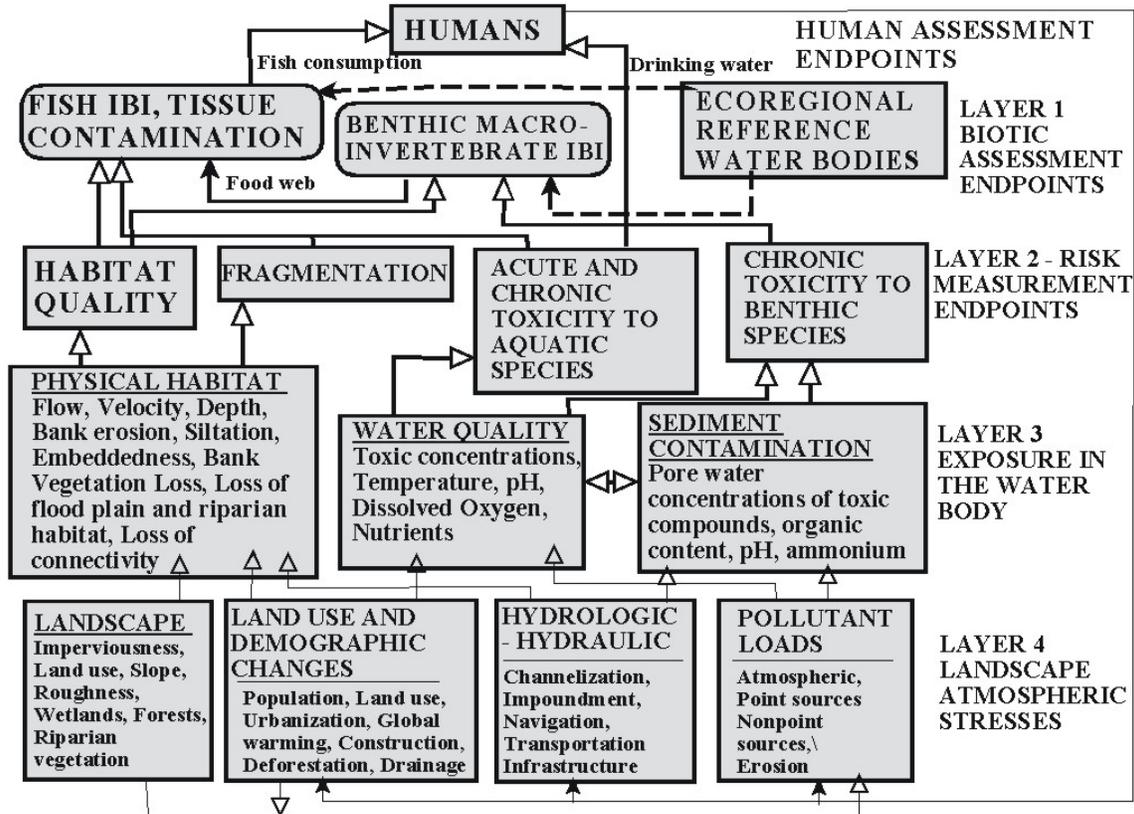
Simple models have correlated urbanization, represented by percent imperviousness, with impaired biological quality; see Figure 1.2. In this model, all the negative impacts of urbanization are grouped together as the parameter imperviousness. Imperviousness of more than 8-10% is correlated with significant biological damage. However, this does not provide enough information to develop treatment to improve quality. It is imperative that the system be modeled with more detail.



**Figure 1.2 Increased % impervious area in a drainage basin correlates with decreased macroinvertebrate quality (Schueler 1994, Novotny 2004)**

As shown in Figure 1.3, a hierarchical model can be used to link stressors to ecological integrity. This model is meant to provide assistance to watershed managers with making their decisions on methods to mitigate stream degradation and biological impairment, assess potential watershed impacts, and identify watershed restoration opportunities. The model concept is based on probabilistic risk propagation through a layered hierarchal system linking stresses with ecological endpoints. Stresses considered in the model will be focused on anthropogenic alterations of physical and chemical qualities of watershed such as impoundments, land use/cover alteration, and pollutant loads. The proposed

model was originally presented in Allen and Starr (1982) and was further developed in Novotny et al. (2001).



**Figure 1.3 Schematic of multilayer risk propagation through hierarchical model (Novotny et al. 2005)**

Layer IV of the model, Root Stressors and Land Use Effects, represents the ultimate source of stresses on the stream. As shown in Figure 1.2, the root stressors can be grouped into four major categories (Karr et al. 1986):

- Landscape morphological/riparian factors and stresses
- Land use change factors and stresses
- Pollutant inputs
- Hydrologic and hydraulic stresses

These categories describe the general types of stresses that impact stream quality. Some of the sources are related to the landscape and hydrology of the watershed, including base flow stability and flashiness, watershed slope, soil characteristics, and land use characteristics and cover. Other root stressors include those that occur directly in the stream, such as direct pollutant discharges and physical stream alterations such as impoundments and channelization. Direct pollutant discharges are relatively easy to identify and are controlled by permitting systems. Diffuse pollution remains a major problem for watershed managers because it is often difficult to identify and regulate these

discharges. Diffuse pollution abatement can reduce a significant portion of stress on watersheds.

Landscape features such as slope and soil characteristics also are important to restoration and protection efforts. Such features affect the way pollutants reach the stream and may be used to define the vulnerability of a watershed. These factors are important because they provide the context in which restoration takes place, and they provide useful information on the capacity for restoration in a stream.

The root stressors and land use effects are connected to the third model layer, Stressors. Layer III represents the actual conditions in the stream and integrates all of the negative effects of the root stressors. These effects range from hydrologic stresses, such as stream velocity, to habitat stresses, such as embeddedness and sediment type, and chemical stresses, such as nutrient and heavy metal concentrations. The stresses are primarily linked to the root stressors, but one stress also may influence another stress. For example, the slope (root stressor) of a stream will determine the velocity (stress) and the velocity will have a direct effect on other stresses such as embeddedness (Leopold et al. 1994). The interactions between stressors and stresses are very complex and interdependent.

## **1.2 Ecological Risk Assessment**

Ecological risk assessment is a process used to define the probability and magnitude of adverse effects on an environment from stressors. Risk assessment can be used to estimate the costs and benefits of changes in a watershed such as new development or restoration. A good risk assessment should be able to identify the relative importance of individual stressors as well as estimate the total stress from multiple stressors. By identifying the relative contribution of stressors and their interactions with each other, restoration alternatives can be evaluated.

There have been numerous attempts to develop a risk assessment methodology that would be useful for watershed decision makers. Such a method would need to be defensible, comprehensive, and broadly applicable. Most past aquatic ecological risk assessments have been focused on stresses from chemical toxicants. However, advances in chemical treatment and point source control regulation have greatly reduced the threat from chemicals. It is now widely agreed whether other features such as habitat quality are the limiting factors in most aquatic environments. Risk assessments should include all sources of stress from chemical, physical, and sediment qualities. Currently, a rigorous, theoretical method for evaluating the risk from multiple stressors does not exist. Even with the abundance of observational data and the development of cause and effect interpretation tools, there “are no tools that allow a confident, a priori, prediction of ecosystem response to multiple stressors” (Swanson 2005).

Current risk assessment procedures can be categorized as two fundamental approaches, identification of patterns and identification of processes. Pattern based assessments use observation evidence to correlate response patterns with stressors. Process based assessments require experimental evidence of cause and effect relationships between measured stressors and environmental responses (Swanson 2005).

### 1.2.1 Pattern Based Risk Assessments

Pattern based approach can be described as using observational, field based data to describe the current state of the system and then deducing how it got to that point. Statistical analysis is often used to examine correlations between identified biological response variables and aquatic stressors. The goal of such a statistical analysis is to determine the relative contribution of a stressor or combined stressors to a response. The response can then be statistically explained by the stressors (Faribrother and Bennett 2000). Metrics, indexes, and other integrated response variables that compare the current condition to a reference condition are commonly used to identify impaired conditions in aquatic ecosystems. This makes them popular as response variables in pattern based risk assessments as well.

The Area Degradation Value and Biological Response Signatures developed by Yoder and Rankin (1995) are two examples of pattern based risk assessments. Area Degradation Value, or ADV, is a quantitative estimate of the degree of departure from a biocriterion along a longitudinal continuum which uses the differences between the potential and actual IBI values of a stream reach to demonstrate the impact of pollution. ADV can be used to differentiate the source of pollution between point and non-point. Point sources are identified by their immediate and severe effect on ADV with recovery occurring downstream. Effects of diffuse sources tend to spread out more evenly throughout a segment. Biological Response Signatures (BSR) were developed to identify the biological response to different types of stressors, including chemical and physical changes. Elements of biosurveys, such as fish and macroinvertebrate IBIs, are combined to distinguish the impact of one type of stress over another. Both approaches work well in simple situations but break down when evaluating more complex problems. In a situation with both point and nonpoint pollution ADV has difficulty distinguishing the relative contributions and BSRs do not work well in cases with multiple stressors.

Other work has been done using IBI with more specific representations of stress. The effects of advanced waste treatment by chlorine and ammonia were successfully related to biological assessments (Karr et al. 1985). This style of evaluation may also be applicable to other contaminants. Studies have also developed links from stressors, such as geomorphic assessment (Leigh et al. 1999), features of urbanization (Younus et al. 1999) and other general types of ecological risk (Parkhurst et al. 1996b; Paulson and Parkhurst 1996), to biotic measurements.

Although pattern assessments and observational data have revealed a great deal of information about multiple stressors they have not yet provided a quantitative and probabilistic approach to risk estimation. The methods developed to this point cannot be used unaccompanied to identify the limiting factors in multiple stressor situations. As the number of stressors increases observational data becomes less and less able to distinguish the relative contributions from stressors. Metrics often represent the cumulative effect of degradation and cannot be used to identify the response from a single stressor within the group.

### 1.2.2 Process Based Risk Assessment

The second group of risk assessment protocols is one that searches for a process that governs the response to stressors. Finding a governing process requires an experimental approach (Swanson 2005). The hypotheses of such experiments are usually based on observational data, but should be applicable to situations beyond those observed. Those which generalize about cause and effect relationships are most productive and useful for prediction.

Currently, there are few experimental approaches that have developed a process for risk estimation from multiple stressors. Development in this area is important to the goal of providing a credible and defensible model for risk assessment.

The ecosystem subsidy-stress model is an example of a process based approach (Odum et al 1979). The base of this model hypothesizes that low levels of some stressors “subsidize” an ecosystem process but loading beyond a theoretical threshold will “stress” the ecosystem. Subsidies are defined to have positive affects on the ecosystem, while stresses have negative affects. The subsidies and stresses are additive, with a subsidy being a negative stress, to obtain the total stress.

The Water Environment Research Foundation (WERF) also has a proposed risk-based framework for analyzing ecosystems, which is based on a subsidy-stress model. The model categorizes responses as one of two types, structural and functional. Structural responses are shown in the make-up of the biological community while functional responses affect the energy and nutrient cycling in the ecosystem. The base hypothesis for multiple stressors is that these two types of responses are additive (Swanson 2005). The original WERF model for multiple stressors was modified to a three tier system involving screening level risk assessment, quantitative risk assessment, and risk refinement and evaluation. Such an approach is now commonly used for ecological risk assessment. It also assumes that multiple risks are additive.

The approach of identifying a process behind cause/effect relationships in multiple stressors systems has brought the field closer to having a broadly applicable prediction system. However, there is still no approach that considers possible nonlinearities in risk. Although it is recognized that some stressors may have synergistic or interfering effects on each other, studies continue to assume that in all multiple stressors situations their responses are additive. Current approaches encounter considerable difficulty quantifying the effects of non-linear interactions if their existence is identified.

There is a serious need among watershed managers to have a reliable and accurate ecological risk assessment procedure. In the past, the largest threat to water systems could be easily identified as the major point sources of pollution from industrial and municipal waste. However, as point source control has increased, the remaining layers of watershed stressors have become prominent. Although water quality has improved a great deal with the reduction of point source pollution, it is widely recognized that continued remediation is necessary. Moreover, it is now much more difficult to identify the limiting stressor and the interactions among multiple stressors are still not entirely understood.

A defensible, probabilistic risk assessment protocol would make watershed managers able to identify possibilities for remediation and help them evaluate the most economical approach for remediation. The prediction of stressor responses can also help watershed decision makers evaluate new development projects proposed in a watershed. They may be able to guide the new development in a way that will cause the least negative impact in the watershed.

The research presented here has advanced and developed three approaches for risk assessment in an aquatic environment. Two methods are based on the response from single stressors and assume multiple stressors are additive, while the third is capable of working with non-linear relationships between multiple stressors. The risk estimation methods include estimation for IBI metrics developed from toxicity data and risk estimation to fish populations through maximum species richness lines (MSRL). Both methods have been previously developed and are applied to new situations in this research. The third is focused on developing the prediction of IBI values using artificial neural networks supervised learning techniques.

### 1.2.3 Risk Estimation from Toxic Response

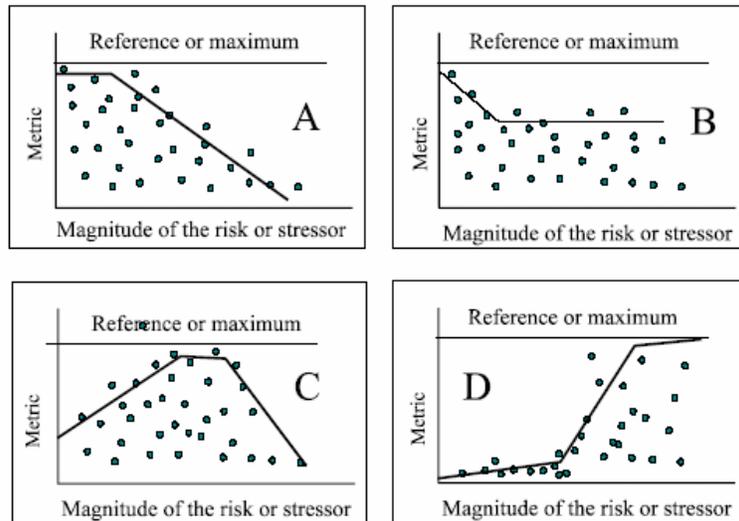
A variety of toxicity tests have been developed to evaluate the effect of stressors on organisms including fish and benthic invertebrates. Such tests usually evaluate the biological response over increasing concentrations of a stressor or over the time of exposure. The most common test and the one used by the USEPA to make recommendations on acceptable concentrations is the LC<sub>50</sub>. LC<sub>50</sub> is a test that determines the concentration which is lethal to 50% of the test species over a given time period. Acute exposure is considered to be any exposure time under 96 hours while chronic exposure is over 96 hours. Toxic response defined by LC<sub>50</sub>s can be used in conjunction with water quality distributions to estimate risk (Bartošová 2002).

Species vary in their tolerance for stressed situations. Metrics group species together based on characteristics such as trophic condition in order to describe the community. When grouped by their metrics, the individual responses can be combined to describe the response of the metric. This response can be used to evaluate the risk to fish communities from individual stressors. A stream's monitoring history can be used to determine the probability of chemical concentrations. The LC<sub>50</sub> profiles of metrics can be evaluated at the ambient concentrations to estimate the risk to the biology in a specific stream. This is a probabilistic approach to estimating risk that can be used to explain the current conditions in a stream, or predict the effect of changes in the aquatic environment.

### 1.2.4 Maximum Species Richness Lines

As suggested by their name, Maximum Species Richness Lines (MSRL) are used to define the maximum number of species that can be expected to be found in a stream. MSRLs were originally developed to account for the effect of stream order on species richness and were used for index evaluation (Fausch et al. 1984) Indexes like IBI use metrics that weigh stream features such as number of intolerant species against the number expected to be found in an undisturbed stream to calculate index values. Because of this weighting system it is important to have relatively accurate measures of expected

undisturbed value of the metric. MRSLs were originally derived from species counts at representative, undisturbed sites. They are usually a wedge shaped plot with population increasing with stream order. The actual line is the regression line at 95% of the observed population. Recently, their use has been extended to identifying other environmental factors that may limit the size of aquatic communities. Figure 1.4 shows examples of possible MSRLs patterns observed for various stressors. Use of MRSLs has also been suggested for determining threshold relationships with stream features and even for determining allowable limits for NPDES permits (Rankin and Yoder 1999).



**Figure 1.4** Patterns of Maximum Species Richness Lines for various stressors (Novotny et al. 2005)

This research, in conjunction with past research by Bartošová (2002), presents a method for using MRSLs to estimate risk from stressors. In this case, stream order remains constant and instead, a stressor of interest becomes the limiting factor for population. When population is plotted versus amount of a stressor the maximum number of expected species count, given the concentration of stressor, is shown. The risk to the population can be estimated from the difference in the maximum at the target stressor concentration and the reference maximum, the maximum when no stress is present.

As with most pattern based risk assessments, this approach is limited in the number of stressors it can involve in an analysis. However, the risk from multiple stressors can be calculated as the joint probability over all stressors of interest. It also provides a great deal of insight into the relationships between stressors and biological populations, even if it may not be useful for retroactive examinations of a stream system. That is, this approach cannot be used to identify individual stressors currently in an ecosystem. However, it can be very useful for planners attempting to predict the effect a change in watershed characteristics will have on the ecosystems as measured by fish populations. This includes the response to both degradation and remediation.

### 1.2.5 Use of Artificial Neural Networks

Artificial neural networks (ANN) have been applied to many areas of prediction and pattern recognition (Demuth and Beale 1992). Current watershed applications usually apply unsupervised learning, which can be used to identify relationships between watershed and stream stressors and stream quality measures or endpoints. This study used supervised learning to develop a prediction of biotic integrity based on measurements of stressor conditions. The model considers inputs from multiple stressors and predicts IBI values. ANN models are trained on past data and learn patterns, which give them the ability to predict future values. The output of the model may be specific to the stream on which it was trained; however, once the network structure that provides good predictions is identified, additional networks can easily be built and trained with data from other streams and be used to predict risk in those locations. Additionally, a trained network can be tested on data with a known output so the confidence of the model can be statistically defined.

The primary advantage of using ANN over previously developed risk models is that it can account for nonlinearities in the system (Haykin 1999). When using ANN, it is not necessary to assume that the effects of multiple stressors are additive. A predictive neural network provides the modeling environment needed to develop an ecological risk assessment that is probabilistic and reliable. It reduces the amount of professional judgment needed during analysis which makes its results more defensible. The output of a network can be easily described in well-understood statistic terms.



## 2 Toxic Response

### 2.1 Evaluation of Toxicity Data

The median lethal concentration or dose (LC<sub>50</sub> or LD<sub>50</sub> respectively) is possibly the most important measure of toxicity. As its name implies, the endpoint of the test is the mortality of 50% of the test community in a set amount of time. Test organisms, even of the same species, do not have a uniform response to toxic influences. The mortality of 50% of species therefore represents the point at which a significant portion is affected. LC<sub>50</sub>s are generally used to quantify effects of acute exposure, exposure occurring for 96 hours or less. A second common test, the effective dose or effective concentration (ED<sub>50</sub> and EC<sub>50</sub>) is commonly used for chronic testing. The endpoint of this test is usually a physiological change such as reduced reproduction or respiratory effects. Chronic criteria are based on effects from long-term exposure instead of dose-response tests.

The conditions used for LC<sub>50</sub> testing are not entirely standardized. Tests are often performed to replicate a certain situation. This includes the chemical being tested for toxicity as well as other water conditions, particularly hardness, and the time of exposure. Tests can be normalized for hardness based on well know relationships between chemical toxicity and calcium concentrations. Tests can also be normalized for time of exposure; however, information on individual relationships between the chemical and the organism needed.

#### 2.1.1 Data Sources

Data used in this research were obtained from the ECOTOX database hosted by the U.S. EPA which is a compilation of literature results for many toxicity tests. ECOTOX is a database of toxicity information for aquatic and terrestrial biology. The database offers compiled results of many toxicity tests, including LC<sub>50</sub>s, from reliable sources. Most values listed in the database are taken from peer-review articles. The database can be queried based on the chemical, organism, and test of interest. The ECOTOX database can be queried from it's website at <http://www.epa.gov/ecotox/>.

#### 2.1.2 Data Refinement

The query returned all available results of LC<sub>50</sub> tests in the database for the six target chemicals. Because of this, the data included LC<sub>50</sub>s performed under a variety of test conditions. Data from experiments not performed in a laboratory in freshwater conditions were removed. For the metals cadmium, copper, lead, and zinc these values were then normalized to a hardness of 50 mg CaCO<sub>3</sub>/l. Aluminum and mercury results were not adjusted in this step because their toxicity is not hardness dependent. The duration of the tests ranged from 1hr up to about 2 months but only those with a 96 hour exposure were used. 96 hours is the maximum exposure to be considered acute. It was used because of both the quantity of available data and the fact that it represents the lowest concentration to have an acute effect and therefore will produce conservative estimates of risk. When multiple values were available for a species the average was taken. Species used in the model were selected based on those that might be expected to be found in the Upper Midwest Region of the United States.

### Hardness Conversion

Certain components in water can affect the toxicity of some dissolved metals. Calcium is one of the primary reactants with dissolved metals. Hardness is considered a surrogate measurement for water quality features that reduce toxicity of several metals; as hardness increases the toxicity of the dissolved metals decreases. The relationship between hardness measured as mg CaCO<sub>3</sub>/l and toxicity has been derived by laboratory experiments and is now part of all criteria involving affected metals. The data used in this research were normalized to a hardness of 50 mg CaCO<sub>3</sub>/l using Equation 2.1.

$$\frac{LC_{50}^H}{LC_{50}^T} = \left(\frac{H}{T}\right)^\alpha \quad (\text{Eq 2.1})$$

Where H is the hardness at which the test was performed, T is the target hardness value, LC<sub>50</sub><sup>H</sup> is the LC<sub>50</sub> at the test hardness, LC<sub>50</sub><sup>T</sup> is the LC<sub>50</sub> at the target hardness, and α is metal-specific conversion factor.

LC<sub>50</sub> values for cadmium, copper, lead, and zinc were normalized to a hardness of 50 mg CaCO<sub>3</sub>/l for the purpose of developing distribution curves. Aluminum and mercury are reportedly not affected by hardness and therefore were not adjusted. The hardness value of 50 mg CaCO<sub>3</sub>/l was chosen because it is commonly used in reporting by the U.S. EPA. However, when calculating the risk due to the metals affected by hardness it is necessary to adjust the toxicity based on the ambient hardness. The alpha values for the affected metals are listed in Table 2.1 (USEPA 1994).

**Table 2.1 Hardness Conversion Factors**

Metal	α
Cadmium	1.128
Copper	0.9422
Lead	1.273
Zinc	0.8473

### 2.1.3 Species Selection and Organization

The species used to develop the model should be those expected to be found in the local area. Since the work being done by this project is primarily focused on watersheds in Illinois and Ohio, species in this analysis were limited to those with habitats in the Upper Midwest Region of the United States.

Species were categorized based on general IBI metrics in order to evaluate the response of IBI metric to toxic influence. Fish metrics are usually grouped based on tolerance and trophic guild. In this study, fish were grouped as tolerant, moderately tolerant, intolerant, carnivores, insectivores, and omnivores. Benthic invertebrates are generally grouped by feeding process. In this study, they were grouped as benthic omnivores, filter/collectors, gather/collector, parasites, benthic predators, scrapers, and shredders. Some metric and chemical combinations were not analyzed because of a lack of data. Tables 2.2 and 2.3 list the number of data points available for analysis in each group for fish and benthic invertebrates respectively. Only those combinations with ten or more data points were used for risk estimation.

The species divisions used for analysis cover the basic IBI metrics. IBI metrics are locally dependant in both species and the process used. For example, the Ohio EPA groups moderately tolerant species with intolerant and has separate groups for groups of species such as darters and sunfish (Ohio EPA 1987). The data can be grouped for local IBI metrics if more precision is necessary; however, data availability may become an issue. The more specifically the data are divided, the fewer the data available for each group. This lowers the confidence in the analysis. This issue is already apparent in the data used here, particularly in the analysis of benthic invertebrates which do not have enough data to confidently estimate risk for many of the subsections. Fish metrics are more likely to have complete data sets with a smaller amount of species because species can be counted twice, once for tolerance and once for trophic guild. Double counting is accepted here because it is an artifact of the IBI analysis that is being replicated.

**Table 2.2 Data Count for Fish**

Chemical	Metrics						
	Tolerant	Mod. Tolerant	Intolerant	Carnivore	Insectivore	Omnivore	All
Aluminum	2	3	1	-	-	-	6
Cadmium	13	18	3	15	11	7	35
Copper	17	24	13	17	23	10	54
Lead	5	6	4	7	3	4	15
Mercury	7	12	2	12	5	4	21
Zinc	12	14	7	14	11	8	33

**Table 2.3 Data Count for Benthic Invertebrates**

Chemical	Metrics							
	Filter/Collector	Gather/Collector	Omnivore	Parasite	Predator	Scraper	Shredder	All
Aluminum	3	1	-	13	3	2	-	22
Cadmium	20	46	10	-	21	11	4	113
Copper	17	19	6	-	11	13	2	68
Lead	7	9	3	1	5	4	-	31
Mercury	6	23	3	2	8	5	3	50
Zinc	11	18	7	1	11	10	-	58

#### 2.1.4 Distributions

Statistical distributions can be used to describe the response of organisms to toxic influence. Although data are not available for all possible species found in the study region, it can be assumed that they will follow the general distribution as the subset presented here where sufficient data was available. Only the most sensitive species are expected to be influenced by toxic concentrations found in streams. This makes analysis and fitting at the lower end of the distribution somewhat more important than at higher concentrations. The cutoff or threshold concentration for analysis depends on the protection goals.

Although choosing a distribution to fit data is statistically based, some knowledge of the underlying is useful for deciding what distributions might provide useful information. Lognormal probability distribution, which describes random variables that are limited to

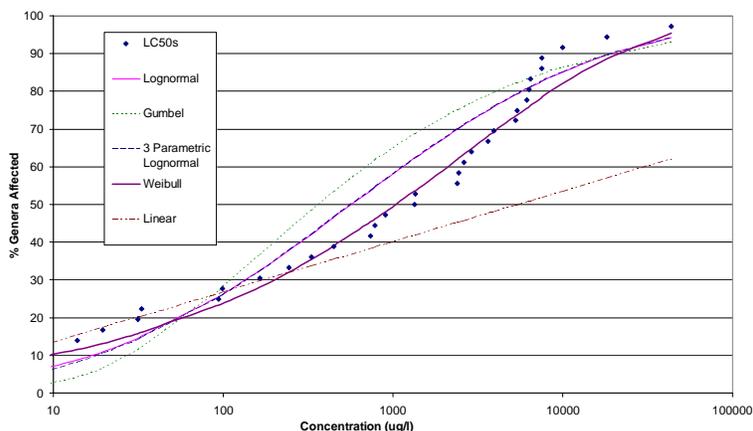
being greater than zero, is the most common distribution used in ecology as it is generally accepted as the best distribution to describe ambient chemical concentrations. It is a variant of the normal distribution, the classic bell-shaped curve. The exponential distribution, which is commonly used for describing the time between events, is characterized by a complete lack of memory. Extreme value distributions such as Gumbel, Weibull, and Pearson Type III are usually applied to modeling seasonal maxima or minima for analysis of major events such as flood size and probability.

The distribution of chemicals in the environment has been studied extensively and is generally well described; however, the distribution of toxic response is not well understood. A single distribution has not been agreed upon for modeling toxic response. Also, modeling is often limited by the number of available data points.

The choice of distribution can strongly affect the value of the calculated risk, particularly at the low concentrations expected in most streams and rivers. Several distributions were examined including the lognormal distribution, the 3-parametric normal distribution, the Gumbel distribution, and the Weibull distribution (all in logarithmic space). A linear approximation of the lower tail is also used based on Novotny and Witte's (1997) suggestion that a linear approximation of the most sensitive species best illustrates response at low concentrations, which is the region of most interest to most ecological risk assessment.

Figure 2.1 shows the toxic response curve to cadmium of all available fish data. As shown, most of the distributions fit the data reasonably well and could be used to estimate  $LC_{50}$ . Goodness of fit is estimated based on the sum of square error and square of the correlation coefficient,  $R^2$ , is shown in Table 2.4. Results from all other combinations of taxa group and stressor are shown in Appendix A. The goodness of fit was evaluated based on sum squared error and square of the correlation coefficient,  $R^2$ . Linear fit was measured just from the lower values on which the linear distribution was based. The quality of the fit for all distributions other than linear was measured from all data points. However, fit at the lower end of concentration of more interest to ecological application because they represent concentrations more likely to occur in ambient conditions.

In this example, the Weibull distribution provides the best fit for the empirical data. The extreme value distributions, Gumbel and Weibull, were most often the best fit for the test cases. Linear approximation also provided an excellent fit for most cases and could be used if only low concentrations are of concern. In this case, the linear portion of the response actually covers most of the concentration region that might be expected to occur naturally.



**Figure 2.1 Toxic response curve for cadmium evaluated at 50 mg CaCO<sub>3</sub>/l**

**Table 2.4 Goodness of Fit for Distributions of Cadmium LC<sub>50</sub>s**

Model	Sum Squared Error						
	All	Intolerant <sup>a</sup>	Mod. Tolerant	Tolerant	Carnivore	Insectivore	Omnivore
Lognormal	1740	-	652	952	372	1241	682
Gumbel	4178	-	1768	2149	1029	2226	1370
3-Param	1801	-	708	988	476	1247	735
Weibull	591	-	276	448	265	950	282
Linear	16	-	7	101	52	110	25

Model	Squared correlation coefficient, R <sup>2</sup>						
	All	Intolerant	Mod. Tolerant	Tolerant	Carnivore	Insectivore	Omnivore
Lognormal	0.937	-	0.951	0.897	0.966	0.918	0.844
Gumbel	0.848	-	0.868	0.769	0.906	0.853	0.687
3-Param	0.935	-	0.947	0.894	0.957	0.918	0.832
Weibull	0.979	-	0.979	0.952	0.976	0.937	0.936
Linear	0.999	-	0.999	0.979	0.991	0.971	0.989

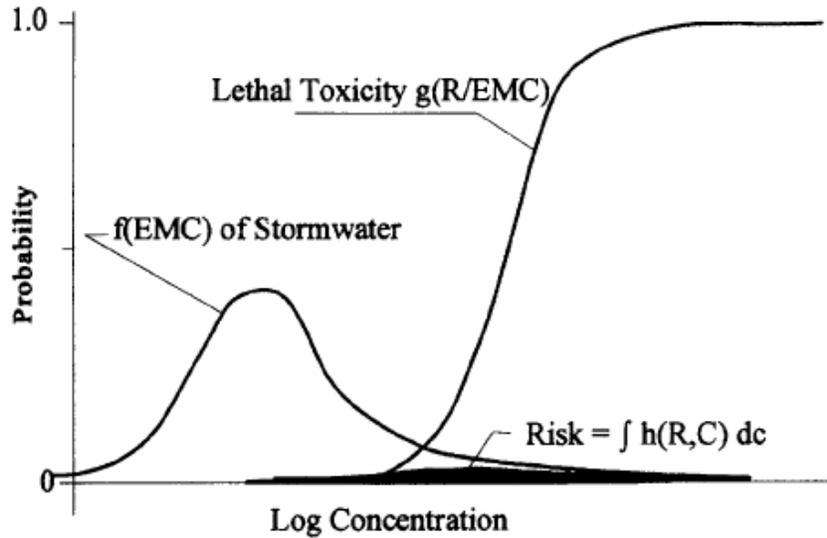
a: not enough data for analysis.

Once the LC<sub>50</sub> data is satisfactorily described by a distribution risk can be calculated. However, care should be taken in choosing the distribution as it can have a significant impact on value of risk assigned. An excellent discussion on the impacts of distribution is presented by Bartošová (2002). The calculation of risk follows the procedure outlined in the WERF methodology by Parkhurst et al. (1996a).

The risk to biota is calculated as a joint probability of 1) the probability distribution of the event mean concentration (*EMC*) of the pollutant in the water body of interest and 2) the toxic response function  $g(R/EMC)$ , which is the probability that an organisms will be adversely affected given the *EMC*. The *EMC* is adjusted for the effect of the dilution ratio (*DR*) and the water effect ration (*WER*) so that  $f(EMC) = pdf(EMC \times DR/WER)$ . The point probability then is:

$$H(R, C) = f(EMC)g(R | EMC) \quad (\text{Eq 2.2})$$

The integration of Equation 32.2 over all concentrations yields the total risk that the organisms modeled by the risk function,  $R$ , will be adversely affected. Figure 2.2 illustrates the process for the calculation of risk from stormwater events; an adaptation of the WERF methodology developed by Novotny and Witte (1997). In this case the risk function has been developed from  $LC_{50}$  values and  $R$  is the metric groupings used.



**Figure 2.2 Ecological risk from stormwater impacts (Novotny and Witte 1997)**

Bartošová (2002) developed an Excel macro that calculates risk using this method. Inputs to the macro include ambient water quality data from the study region of interest,  $LC_{50}$  values for the chemical and group of interest, and the alpha value used for hardness conversion if necessary. The data from this study was applied to the program to calculate risk to IBI metric values based on toxic response to chemicals in a sample stream site.

## 2.2 Estimation of Risk Response Function Using Toxicity Data

### 2.2.1 Risk Calculation

Risk was calculated using a macro developed by Bartošová (2002), which implements the WERF methodology for risk calculation. This method requires information about ambient water quality including the mean and standard deviation of the chemical concentration. The quality of the water in the stream is the origin of the risk to the organisms. The value of hardness is also needed to adjust toxicity. The Ohio Database was used to provide water quality information. One year of data from site S17830 was used to determine the mean and standard deviation for analysis of aluminum, cadmium, copper, lead, and zinc. Mercury data is not provided in the Ohio Database; therefore, mercury analysis for this site was not done. The water quality values for the site are shown in Table 2.5.

**Table 2.5 Mean and Standard Deviation of Concentrations at Site S17830**

Parameter <sup>a</sup>	Mean	Standard Deviation
Total Aluminum (ug/l)	1022.67	534.40
Total Cadmium (ug/l)	0.11	0.04
Total Copper (ug/l)	9.33	4.95
Total Lead (ug/l)	3.00	2.45
Total Zinc (ug/l)	16.67	14.47
Hardness (mg/l)	281.88	-

a: Dissolved concentrations would be preferable for analysis but were not available.

**Table 2.6 Risk calculations for fish**

Chemical	IBI Group	Risk Values for Fish		
		Gumbel	Lognormal	Linear
Cadmium	All	1.21E-61	2.40E-58	2.02E-68
	Intolerant	-	-	-
	Mod. Tol.	1.11E-61	2.36E-58	6.53E-88
	Tolerant	1.98E-89	1.92E-60	1.31E-195
	Carnivore	1.29E-60	3.72E-58	6.50E-88
	Insectivore	6.45E-68	3.90E-59	1.12E-132
	Omnivore	-	-	-
Copper	All	4.41E-08	1.88E-04	4.64E-06
	Intolerant	3.98E-06	1.13E-03	3.15E-05
	Mod. Tol.	5.98E-08	5.51E-04	2.07E-09
	Tolerant	7.33E-09	3.89E-04	1.63E-11
	Carnivore	6.46E-07	1.38E-03	3.17E-08
	Insectivore	1.58E-05	4.20E-03	2.55E-05
	Omnivore	1.93E-10	3.62E-05	1.21E-10
Lead	All	6.80E-46	8.39E-13	0.00E+00
	Intolerant	-	-	-
	Mod. Tol.	-	-	-
	Tolerant	-	-	-
	Carnivore	-	-	-
	Insectivore	-	-	-
	Omnivore	-	-	-
Zinc	All	1.35E-07	2.73E-04	4.99E-04
	Intolerant	-	-	-
	Mod. Tol.	7.35E-11	7.80E-06	0.00E+00
	Tolerant	4.73E-12	8.02E-06	0.00E+00
	Carnivore	5.18E-14	8.56E-07	0.00E+00
	Insectivore	1.26E-04	6.41E-03	6.73E-04
	Omnivore	-	-	-

**Table 2.7 Risk calculations for benthic invertebrates**

Chemical	Metric	Risk Values for Benthos		
		Gumbel	Lognormal	Linear
Aluminum	All	1.03E-02	2.18E-02	2.02E-02
	Filter/ Collector	-	-	-
	Gather/ Collector	-	-	-
	Omnivore	-	-	-
	Parasite	1.00E-04	4.47E-03	1.11E-03
	Predator	-	-	-
	Scraper	-	-	-
Cadmium	All	2.13E-110	6.48E-62	1.26E-98
	Filter/ Collector	3.23E-59	7.80E-58	3.46E-101
	Gather/ Collector	2.35E-113	2.13E-61	4.81E-212
	Omnivore	-	-	-
	Parasite	-	-	-
	Predator	7.75E-64	1.12E-58	7.33E-172
	Scraper	3.16E-81	4.62E-60	4.58E-183
Copper	All	8.73E-05	2.96E-03	9.13E-03
	Filter/ Collector	1.32E-02	4.38E-02	3.26E-02
	Gather/ Collector	2.14E-04	1.11E-02	4.40E-09
	Omnivore	-	-	-
	Parasite	-	-	-
	Predator	5.87E-04	1.50E-02	3.14E-03
	Scraper	1.53E-06	2.36E-04	5.15E-06
Lead	All	5.11E-08	1.06E-03	2.00E-07
	Filter/ Collector	-	-	-
	Gather/ Collector	-	-	-
	Omnivore	-	-	-
	Parasite	-	-	-
	Predator	-	-	-
	Scraper	-	-	-
Zinc	All	3.53E-07	1.99E-04	1.57E-04
	Filter/ Collector	7.99E-05	2.92E-03	3.07E-04
	Gather/ Collector	4.52E-08	2.74E-04	1.34E-08
	Omnivore	-	-	-
	Parasite	-	-	-
	Predator	4.88E-15	4.23E-06	0.00E+00
	Scraper	4.34E-09	5.19E-05	0.00E+00

The choice of distribution has a very strong affect on the value of risk assigned through this method, particularly in the lower range of concentrations. For comparison, risk was calculated using lognormal, Gumbel, and linear distributions. Tables 2.6 and 2.7 show the risk for fish and benthos respectively.

As results illustrate, differences in risk between distributions can be many orders of magnitude. Generally, the lognormal distribution results in the highest value of risk and linear in the lowest. With the linear distribution it is even possible to have effectively zero risk if distribution intercepts zero affected at a concentration higher the distribution of ambient concentration. It is very important that care is taken when choosing the distribution to use for calculating risk. It cannot be assumed that one distribution will fit all cases. Careful examination of the fit of the distributions in the range of the concentration of interest is very important for reliable results.

The risk values can be used to evaluate some relationships between the metric groups and the differences in their response to chemical stress. The groups with all data points tended to be very close to the average risk for all metrics. For fish, the risk to intolerant fish and insectivores was generally higher than average, while the risk to tolerant fish, carnivores, and insectivores was lower. In the benthic invertebrate groups, filterer/collectors tended to have the highest risk followed by predators while gatherer/collectors and scrapers had the lowest risk.

Responses are also dependent on the type of chemical creating the stress. Although direct comparisons cannot be made by these measurements because the chemicals were at different concentrations, some information can be drawn from the results. For example, the calculated risk for tolerant fish from copper was higher than that for zinc even though the copper concentration was much lower. This suggests that copper is a more toxic metal than zinc. Additional comparisons could be made by setting the influent concentrations all to the same mean and standard deviation. This would reveal more about the toxicity of the individual metals but would not describe realistic conditions.

This approach is an effective way to calculate risk for a target group of species; however, it is not ideal. The results of this method can be applied to the Upper Midwest Region of the United States, but would not be representative outside that area because the species used to develop the distribution are specific to the Upper Midwest Region. If a distribution can be fit well to the data then the resulting risk may be robust; however, research has shown that fitting a distribution can be difficult and distributions can have widely variant results. Currently the only stressors modeled are toxic chemicals. Risks from other variables such as habitat are difficult to calculate using this methodology because there are no test endpoints similar to the  $LC_{50}$ .



## **3 Maximum Species Richness Lines**

### **3.1 Developing Maximum Species Richness Lines**

Maximum Species Richness Lines (MSRLs) were examined as a second approach to estimating risk from single stressors. MSRLs can be used to determine the maximum number of species expected to live under certain conditions. Risk is defined by comparing the maximum species count in the absence of a stressor with the count at a certain concentration of stressor. They provide more flexibility than  $LC_{50}$ s because they can be applied for stressors that are not easily examined using toxicity tests. Because they rely on in situ measurements, they can be developed for any parameter that is measured at the same time as biology, including chemical and habitat quality measurements. Traditionally, MSRLs were used to relate the maximum number of species to stream order; however, they can also be adapted to express the maximum count based on other stressor concentration. Additionally, a similar process can be used to relate maximum IBI or other metric values to stressor concentration by replacing species counts with IBI and metric values.

#### **3.1.1 The Ohio Database**

Data used to develop MSRLs was extracted from a watershed monitoring database provided by Edward Rankin of the Ohio EPA (currently, Center for Applied Bioassessment and Biocriteria, Columbus, OH). The database includes measurements of water quality, habitat quality, and biotic integrity. Habitat quality is measured by the Quantitative Habitat Evaluation Index (QHEI) which was also developed by the Ohio EPA (Rankin 1995). Biotic integrity was measured as both the Index of Biotic Integrity (IBI) for fish and the Invertebrate Community Index (ICI) for benthic invertebrates. Biological data is available for the years 1981-2001; however, only data from 1991-2001 was used in this modeling.

Biological data was extracted from the Ohio database through a query for the records for NUMSPECIES and NUMTAXA. Water quality data was extracted for cadmium, copper, mercury, lead, zinc (same chemicals as those used for risk calculation based on toxic response except aluminum which was not available), dissolved oxygen, and pH. Table 3.1 lists the parameters used and their descriptions.

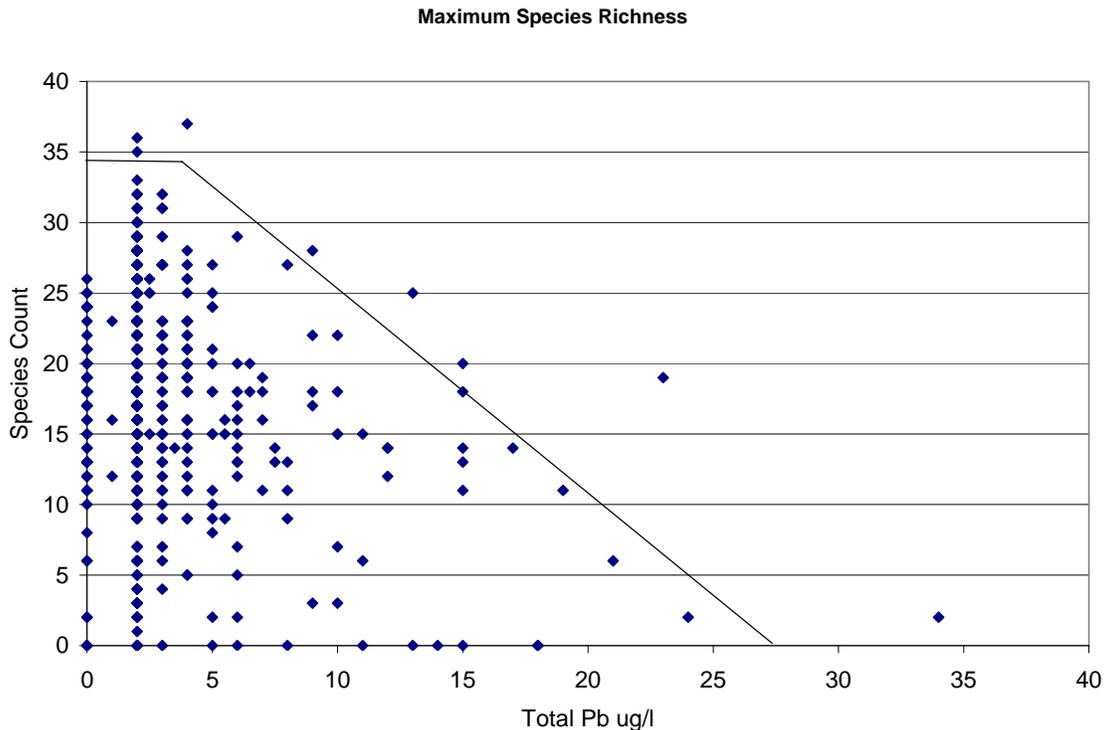
#### **3.1.2 Data Manipulation and Plotting**

Water quality sampling and biological sampling for this database did not occur at the same time; therefore, the dates for the records do not match. Records of species counts were matched with water quality data that had occurred closest in time. For example, if water quality was measured on 6/6 and 6/17 and species counts were made on 6/10 then the species counts would be matched with the water quality data from 6/6. If the two data sets occurred more than 7 days apart in time the data was not used. Biological data is understood to represent the long term affect of water quality and may be expected to remain relatively constant over a short period of time; therefore, little error is expected from this matching procedure.

**Table 3.1 Description of Data used from Ohio Database**

<b>Species Counts</b>	
FIELD NAME	DESCRIPTION
NUMTAXA	TOTAL NUMBER OF QUANTITATIVE TAXA COLLECTED FROM A HESTER DENDY SAMPLE.
NUMSPECIES	NUMBER OF SPECIES IN A SAMPLE EXCLUDING HYBRIDS AND EXOTICS ('SPECIES' USED IN IBI)
<b>Water Quality Parameters</b>	
FIELD NAME	DESCRIPTION
P1027	TOTAL CADMIUM (UG/L)
P1042	TOTAL COPPER (UG/L)
P1051	TOTAL LEAD (UG/L)
P1092	TOTAL ZINC (UG/L)
P299	DISSOLVED OXYGEN (MG/L)
P400	PH, FIELD

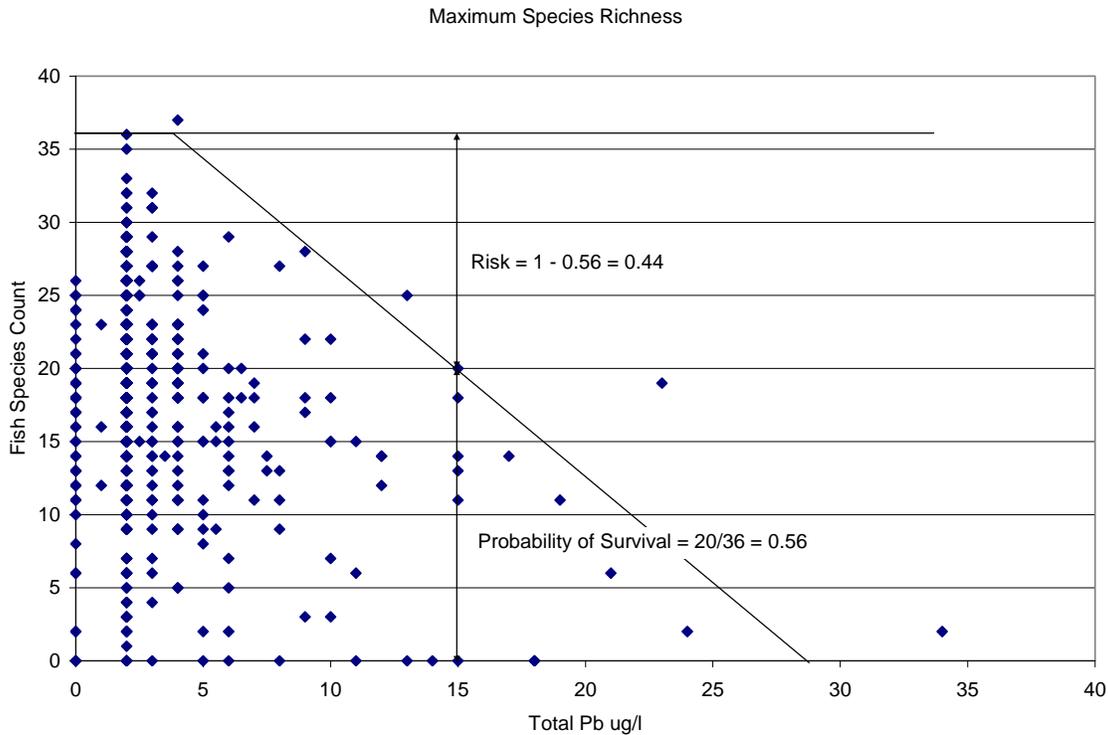
Species counts were plotted versus water quality measured concentration/value to develop the MSRL. The MSRL should be drawn to contain 95% of the data. This can be done by professional judgment or through regression analysis (Rankin 1995). As an example Figure 3.1 shows the MSRL for fish given lead concentration.



**Figure 3.1 Maximum Species Richness line for lead**

The risk of biota extinction is one minus the probability of survival which can be calculated directly from the MSR lines (Bartošová 2002). The risk to biota is represented by the difference between the maximum with no stressor present and the maximum at the

target level. The probability of survival given the habitat,  $p_S^{(fish/habitat)}$ , is calculated as the ultimate maximum over the target maximum and the risk of extinction,  $p_E^{(fish)}$ , is one minus the probability of survival. Figure 3.2 illustrates the calculation process.



**Figure 3.2 Calculating the risk of extinction using maximum species richness**

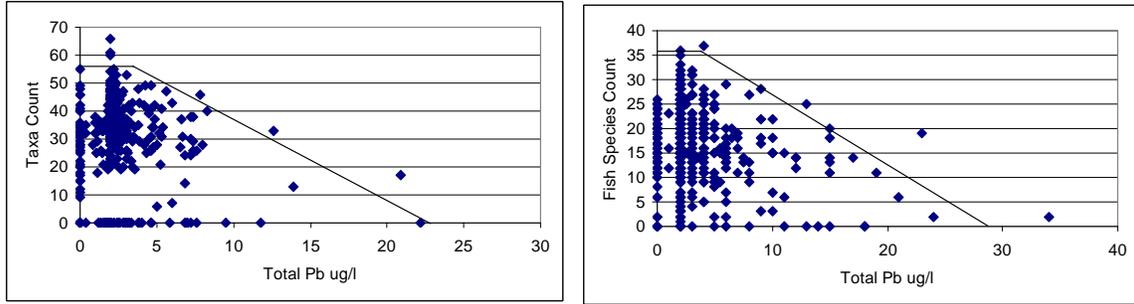
The effect of multiple stressors is assumed to be additive. Therefore, the total risk is the joint probability of extinction given by Equation 3.1.

$$p_E^{(fish)} = \prod_{i=1}^N (1 - p_S^{(risk\_i|hab\_i)}) \quad (\text{Eq 3.1})$$

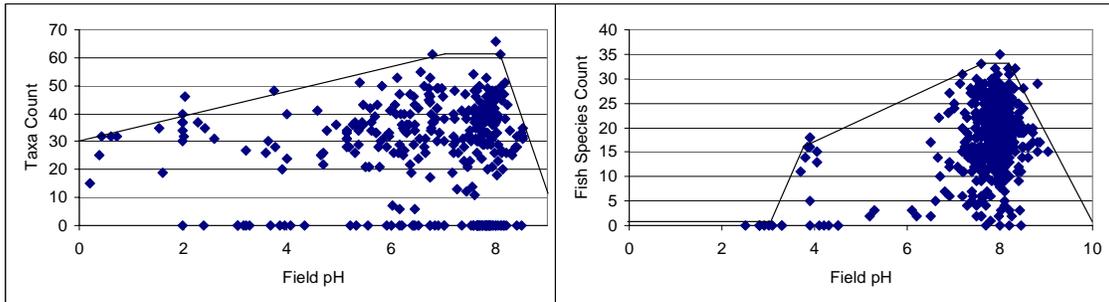
Where  $p_E^{(fish)}$  is the joint probability of fish extinction,  $p_S^{(fish|hab\_i)}$  is the probability of fish survival given habitat conditions  $i$ , and  $N$  is the total number of stressors influencing the fish. MRSLs and risk calculations were performed for the parameters, cadmium, copper, lead, zinc, dissolved oxygen, and pH for fish and taxa. Additionally, the relationships between some metric values and stressor concentration were examined.

### 3.2 Risk Calculation Using Maximum Species Richness Lines

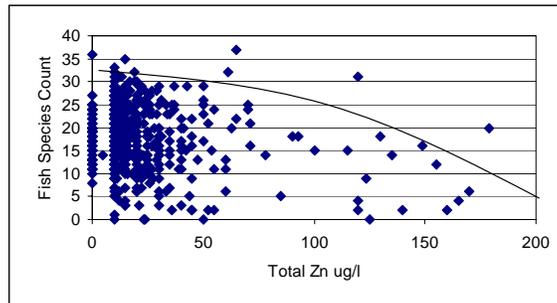
MSRLs can take on many shapes depending on the type of stress being applied. The most common response is a decrease in species count as parameter concentration increases as shown in Figure 3.3. Other shapes are possible, such as a peak in species count which is common for pH or flow, or a non linear response; see Figures 3.4 and 3.5 respectively for examples.



**Figure 3.3 Maximum species richness for lead: Count decreasing with increasing concentration**

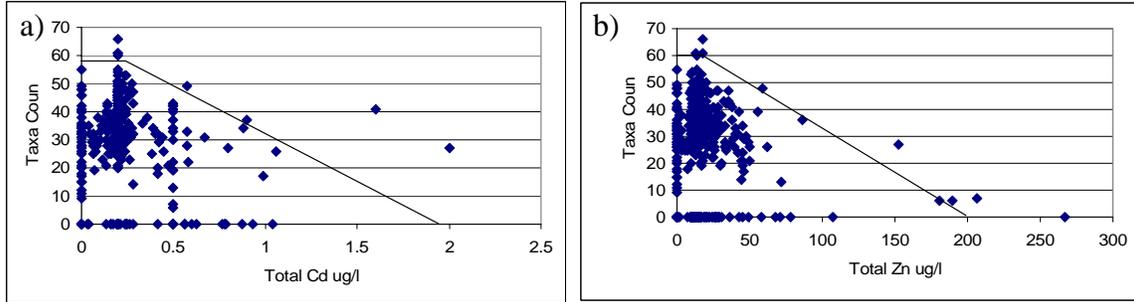


**Figure 3.4 Maximum species richness for pH: MSRL displays a peak at pH 7-8**



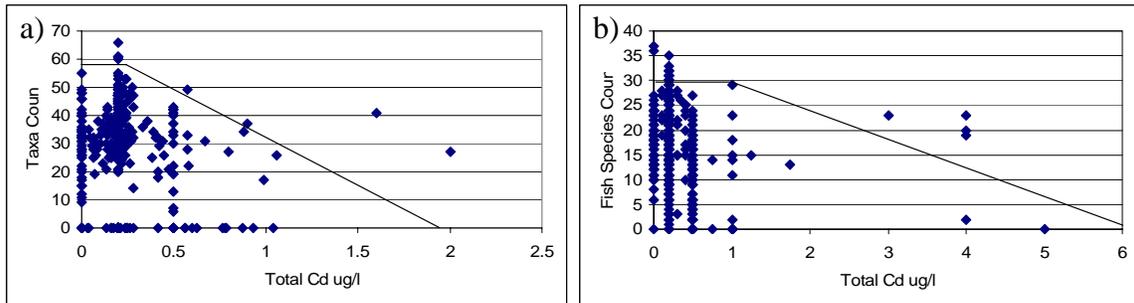
**Figure 3.5 Maximum species richness for fish vs zinc: Non-linear response**

The slope of the MSRL indicates the level of response to the stressor. That is, a steep slope indicates a great increase in risk with a small increase of stress while a shallow slope implies that the species richness is relatively insensitive to the parameter and an increase in the parameter results in a relatively small increase in risk. The results from working with  $LC_{50}$ s showed that copper presented a higher risk to fish than zinc even at a lower concentration. MSRLs for zinc and copper can be compared to show a similar relationship. The slope of the copper MSRL is much steeper than the zinc MSRL, which has a threshold concentration higher than that describing all copper response. Figure 3.6 compares copper and zinc MSRLs.

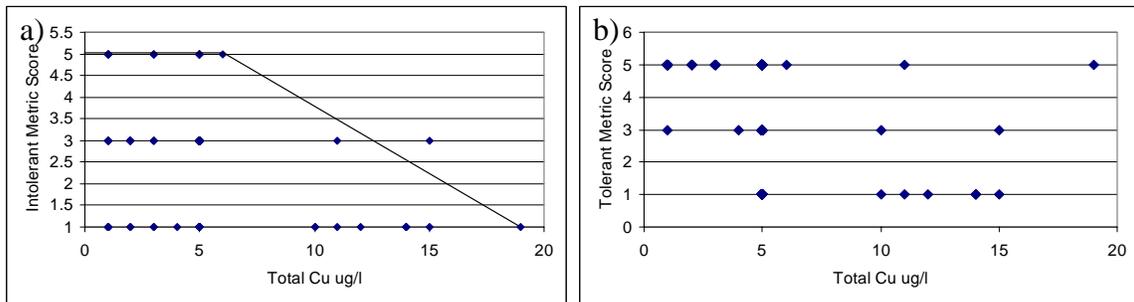


**Figure 3.6** Maximum species richness for a) copper and b) zinc

Similarly, the shape of the MSRL can be used to compare responses of different groups of organisms to the same stressor. In general, fish species react more strongly to stressors in the water than benthic invertebrates, as shown in Figure 3.7. This is probably because the stressor concentrations in the water column may not be entirely representative of the stress influence in the sediment where the invertebrates live. The response of different fish metric groups can also be compared. The Ohio database included individual metric scores. These scores plotted versus stressor concentration show the maximum score given the amount of stressor. Figure 3.8 shows the difference in response of intolerant and tolerant species to copper concentration. The metric score for intolerant species is reduced as copper increases but the tolerant score is not affected.



**Figure 3.7** Comparing cadmium MSRLs for a) invertebrates and b) fish



**Figure 3.8** Comparing responses of a) intolerant and b) tolerant fish to copper

Risk can be calculated from MSRLs at any concentration of interest. The risk represents the probability of extinction. Although individual MSRLs must be made for each stressor, the risk from multiple stressors can be calculated. As previously discussed, the risk of extinction from multiple stressors is one minus the joint probability of survival. Using

this approach, any combination of stressors can be examined assuming their effects are additive. This provides an approach to examine multiple stressors without having the response suppressed which was a problem encountered by other, similar approaches.

## 4 Artificial Neural Networks

### 4.1 Introduction to Artificial Neural Networks

Artificial neural networks have been used in many fields for performing pattern assessment. They are a powerful tool because of their ability to model non-linear relationships between inputs and targets as well as their ability to generalize inputs. Generalization allows networks to adapt to input combinations they have not been exposed to and still provide a reasonable prediction. This property is essential for ecological work because of the numerous possible combinations of conditions.

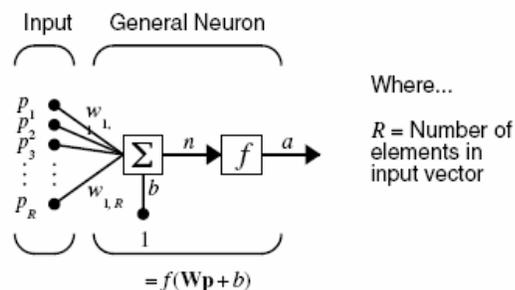
#### *Goal of Study*

The goal of the study was to develop an ANN that could predict IBI, the endpoint of the hierarchical model presented in Figure 1.3, when provided with water, habitat, and sediment quality. Considering this, the most complete dataset available at the time of the study was the database described in Section 3.2, the database developed by the Ohio EPA and provide by Edward Rankin. A database for watersheds in the Upper Illinois River Basin is being developed; when the database is complete, the network will be tested on data from that region.

#### 4.1.1 Network Architecture

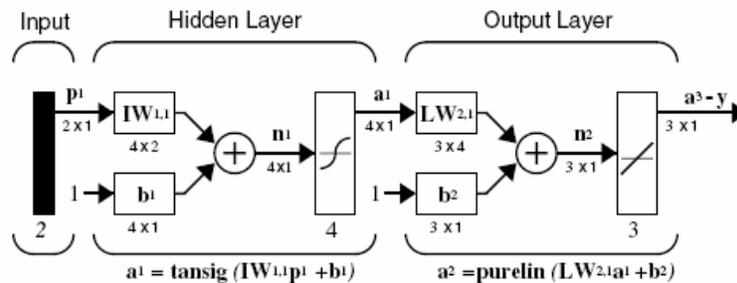
There are two basic approaches to artificial neural networks (ANN), supervised and unsupervised learning. Unsupervised learning can be used to find patterns and develop relationships in the data. In supervised learning, which is used in this study, the network is trained from a dataset including both inputs and targets. A well-trained network can then be used to predict the target value given just the inputs (Demuth and Beale 1992).

The architectures tested here were all variations of a layered, feedforward network using backpropagation. The variables involved in structuring such a network include the number of layers, the number of neurons in each layer, the transfer functions used in the layers, and the training function. Backpropagation networks are preferable for this application because of their ability to generalize. Generalization is the ability of a well trained network to produce a reasonable output for a set of inputs it has never encountered. It can be trained with a representative set of inputs and targets, but does not have to be exposed to every possible situation. Backpropagation networks with biases, a sigmoid layer, and linear output layer are capable of approximating any function, linear or nonlinear, with a finite number of discontinuities (Demuth and Beale 1992).



**Figure 4.1** A model neuron (Demuth and Beale 1992)

The basic elements of a network structure are shown in Figure 4.1. Each input is weighted with  $w$ . The weighted inputs and biases,  $b$ , are summed and sent through a transfer function. The transfer function can be any differentiable function. After initialization, the weights and biases are determined by training which makes changes based on the error measured between the network output and the target outputs. The network can have multiple neurons in a layer as well as multiple layers. In a multi-layer network, the output of one layer is the input for the next, which has another set of weights and biases associated with it. This continues until the final layer where the output should approximate the target. The layers between the input and the output layer are referred to as hidden layers. Figure 4.2 illustrates a multi-layered network.



**Figure 4.2** A model of a layered network (Demuth and Beale 1992)

The three transfer functions that are most commonly used in backpropagation networks are log-sigmoid, tan-sigmoid, and linear. Log-sigmoid and tan-sigmoid both have limited ranges of outputs, zero to one and negative one to one respectively. They are commonly used in the hidden layers of a network. The linear transfer function can have output of any value and so it is usually used in the output layer when target values are not in the range of negative one to one.

#### 4.1.2 Developing Network Structure

Choosing the best structure for a particular problem is largely a trial and error process (Demuth and Beale 1992). Possible structures are virtually endless and it is difficult to predict the level of complexity required for a given application. Often, an overly complex network does not add much to performance but can take an excess amount of time to train. This study began with simple structures and slowly increased their complexity to determine the best structure for the application as well as examine the influence of structure features on the results.

The neural network was built using the Matlab Neural Network Toolbox. The toolbox contains many predefined training functions that can be used to train a network. In general, training functions change the weights and biases to minimize the difference between the outputs and the targets, which is usually defined by the mean squared error. Basic backpropagation training functions move weights in the direction of the negative gradient. In more complex training functions variables can be changed to increase the speed of convergence. The two most common variables used to increase convergence are the learning rate and momentum. The learning rate is multiplied by the negative gradient to determine the change in the weights. The larger the learning rate, the bigger the step taken in each passes. Care must be taken when working with learning rates as too large a

step may cause the network training to become unstable while too small a step takes too long to converge. Momentum allows the training to respond to both the local gradient and the recent trends in the error surface. This protects the network from getting hung up in local minimums of error surface. It accomplishes this by making the weight change equal to the sum of the change suggested by local gradient and a fraction of the previous weight change. Several different predefined Matlab training functions were tested with varied learning rates and momentum.

Training is considered complete when the network reaches its target error or the error ceases to decrease with continued training. If training continues beyond the point where performance is not improved the error may increase because of a loss in the ability to generalize. This condition is referred to as over-training and is usually avoided by performing cross-validation at the same time as training. Cross-validation uses a small set of the input and output data separate from the training data to evaluate the performance of the network and determine the appropriate stop time.

The networks built used four different Matlab backpropagation training algorithms, 'traingd', 'traingda', 'traingdm', and 'traingdx.' Traingd trains using batch steepest decent and a constant learning rate. Traingda and traingdm are similar to traingda but traingda has an adaptive learning rate and traingdm uses momentum. Traingdx uses both an adaptive learning rate and momentum. All of the algorithms were able to reach maximum training value in short time periods, usually less than 100 epochs. The networks built had 2 to 5 hidden layers each of which had multiple neurons which used either a tan-sigmoid, Matlab function 'tansig', or log-sigmoid, Matlab function 'logsig', transfer functions, and an output layer with a single neuron using a linear transfer function, Matlab function 'purelin'. Linear transfer functions were used in the output layer because the target value range was greater than -1 to 1.

#### 4.1.3 Simulations

Once the network was trained to the appropriate level, it was tested with the remaining dataset to which it had not yet been exposed. The output generated by the network was compared with the known target values to determine the quality of the network predictions. The simulation automatically compared the test output of the network, A, and target IBI values, T, by plotting them against each other. The plot shows the linear best-fit line for the data and the target, one-to-one line along which the data should fall, as well as the correlation coefficient, R. Ideally, the equation for the best-fit line should have a slope of one and a y-intercept of zero. This would indicate perfect prediction. Realistically, these values are a goal and not a requirement for acceptable prediction

## 4.2 Network Data

Supervised neural network development requires two types of data each separated into three groups for the different phases. Data are divided first into groups of inputs and the output. The input data include and parameter which is felt to influence the output data. The output data is the parameter for which network prediction is being developed, sometimes referred to as the target. In this study, the input data were various sets of

parameters including information about water, habitat, and sediment quality. The output or target of prediction was the biological endpoint IBI. The input and output data were divided into three groups for the different stages of network development, training, cross-validation, and simulation.

IBI was measured at most stations in the Ohio database two or three times per year, usually between June and September. Fish sampling was generally not done on the same days as water quality sampling. Two approaches were taken for matching IBI and water quality data. The first was to assume the value of IBI was representative for the month in which it was measured and apply the single IBI value to all dates of water quality measurement in the same month. Using this approach the water quality input comes from a single measurement. However, water quality can have unexpected high or low values that are not representative of longer term water quality. To account for this, a second approach was also developed using the median value of the water quality measurements taken over the same month as the IBI measurement. Data prepared using both methods were used separately to train and test networks; however, results tended to be better for those developed by the first approach, probably because extreme values are often the most informative in an ecological setting.

The habitat in Ohio was evaluated once a year at stations throughout the state. Since habitat is relatively stable, the single measurement of habitat quality was duplicated for all measurement dates at the station over the year. The habitat was evaluated by the Qualitative Habitat Evaluation Index (QHEI). In the database there are nine parameters for habitat, the seven index scores, percent embeddedness, and the total QHEI score. See Table 4.1 for a list of habitat parameters. The total score for QHEI was used to represent the quality of habitat as an input to the network. This was done to reduce the number of inputs presented to the network and therefore reduce the complexity.

**Table 4.1 Habitat Parameters Available in Ohio Database**

Parameter Name	Description
EMBEDDED	Embeddedness
COVER	Cover Metric Score For QHEI
CHANNEL	Channel Metric Score For QHEI
RIPARIAN	Riparian Metric Score For QHEI
POOL	Pool Metric Score For QHEI
RIFFLE	Riffle Metric Score For QHEI
GRADIENT_S	Gradient Metric Score For QHEI
SUBSTRATE	Substrate Metric Score For QHEI
QHEI	Qualitative Habitat Evaluation Index

There was no direct measure of sediment quality in the Ohio database. Therefore, the invertebrate community index (ICI) was used to represent sediment quality. This decision was based on the assumption that ICI is directly influenced by sediment quality. This assumption may be valid because by definition the benthic invertebrates in a stream live primarily in the sediment environment and good sediment quality must be maintained to have a good community index value. ICI was also measured about once per year. As with

QHEI, ICI values were duplicated over all other measurements made at the station during the year.

There are 48 parameters measured for water quality in the database. Any measurement that was listed as being at or below the detection limit was replaced with the concentration equal to half the listed detection limit. Except for the temperature parameter, any value listed as zero was removed. Zero was used in the database where measurements were not made. Water quality was measured year round at many stations; however, only those measurements taken in the same month as an IBI measurement were use. Other dates were removed from the dataset because they could not be associated with a target IBI value. One factor that makes neural networks popular is their ability to process a large number of input parameters and determine their relative influence on the output. Because of this, the first group of parameters was limited primarily on data availability. This was necessary because not all parameters were measured on all dates. Only station and dates with complete records, i.e. no missing values, were used for training and testing of the networks.

It is advised for a neural network to have data with a limited range of values. In this case, inputs were normalized to values between zero and one. Once the parameters were selected, the log of the reported value was calculated and then all values were normalized using Equation 4.1.

$$input = \frac{X - \min}{\max - \min} \quad (\text{Eq. 4.1})$$

Where the *input* is the value used by the network, X is the log of the original parameter value, min is the minimum value of X, and max is the maximum value of X. Although the outputs can have any range in value, the performance of the network is usually improved if the range of these values is also reduced. IBI values range from 12-60 (Ohio EPA 1987). The best results were obtained when the target values were IBI values were divided by ten; however, normalizing them from zero to one, in the same way as the inputs, was also tested.

The data were separated into three groups for the purpose of training and testing the network. The training dataset was the largest, using about 70% of the full dataset. All training datasets had over 600 complete records. The other 30% was used to test the data in two ways. A small set, about 10%, was extracted to perform validation during the training. The remaining data was used to test the performance of the network once the training was complete. All sub-datasets were randomly generated from the full dataset.

### **4.3 Preliminary Artificial Neural Network Analysis**

Several different supervised, backpropagation networks were built, trained, and tested with stream quality data from the Ohio database. The stream quality data included parameters that measured chemical, habitat, and macroinvertebrate quality. Several different groups of inputs were used to train and test the networks separately in an attempt to improve prediction of IBI.

The analysis presented represents preliminary work with neural networks and has not explored the limits of their application. Current work in unsupervised network analysis is being done to improve understanding of the parameters in the Ohio database and their relationship with biological integrity. The results of this work will be used to determine which parameters have the most influence over in stream biology.

Several different groups of data were used to train and test numerous different networks. The parameter groups were each a different subset of parameters available in the Ohio database. Initially, data sets consisted primarily of chemical parameters and used only QHEI to represent habitat quality and ICI for sediment. This was done to reduce the number of parameters in the input and maintain some simplicity in the simulation.

#### 4.3.1 Initial Results

Unfortunately the results from the simulations involving the initial sets were less than satisfactory. Results tended to either cluster well around a centerline but not have the target slope and intercept, or have a best fit line closer to the one to one but have more highly scattered data. Consideration of both the correlation coefficient and the quality of the best-fit line should be included in the evaluation the quality of any results.

The best results were obtained using the training functions traingda and traingdx. Although the additional groups were made in an attempt to increase performance, the first group of parameters, which had the most parameters, developed the best performing networks. Figure 4.3 shows the results of simulation four networks using Group 1 parameters. Networks A and B used the traingda training algorithm and Network C and D used traingdx. The structures of the networks are listed below:

**Table 4.2 Network Training Group 1**

Parameter Name	Description
P10	TEMPERATURE (CENTIGRADE)
P299	DISSOLVED OXYGEN (MG/L)
P400	PH, FIELD
P530	TOTAL SUSPENDED SOLIDS (MG/L)
P610	TOTAL AMMONIA (MG/L)
P1045	TOTAL IRON (UG/L)
P310	BOD, 5-DAY, (MG/L)
P900	HARDNESS
P940	CHLORIDE (MG/L)
P945	SULFATE (MG/L)
P1027	TOTAL CADMIUM (UG/L)
P1051	TOTAL LEAD (UG/L)
P1092	TOTAL ZINC (UG/L)
ICI	INVERTEBRATE COMMUNITY INDEX
QHEI	QUALITATIVE HABITAT EVALUATION INDEX

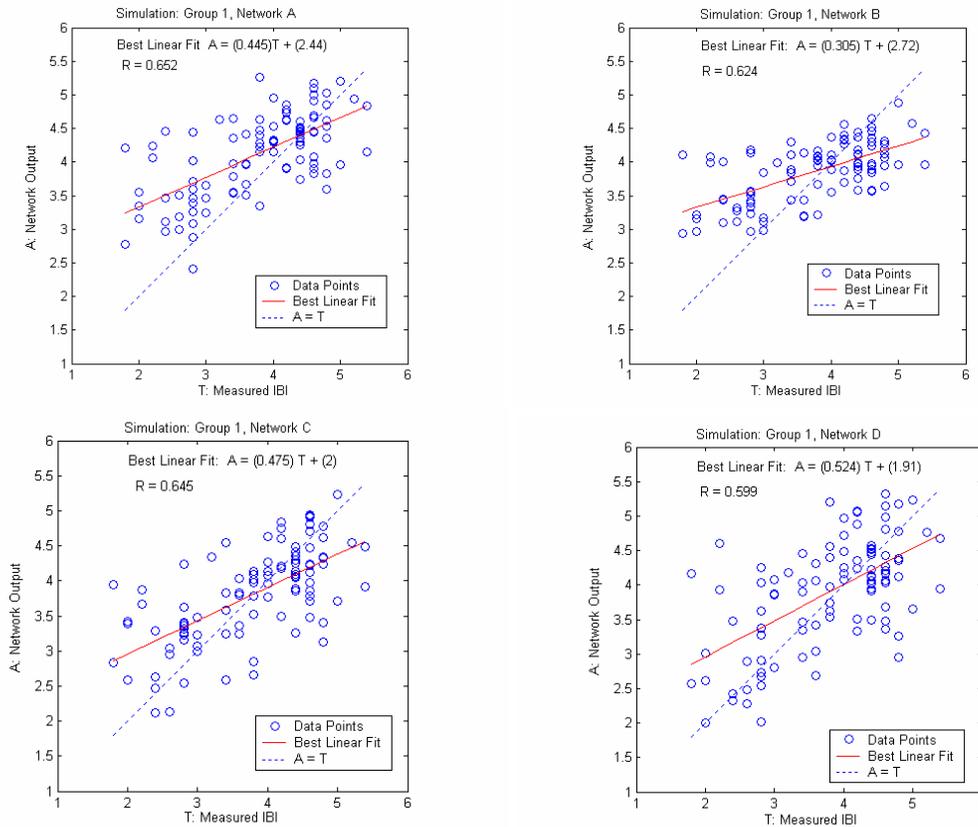
Network A: net = newff(minmax(p), [100, 50, 1], {'logsig', 'logsig', 'purelin'}, 'traingda')

Network B: net = newff(minmax(p), [100, 25, 1], {'tansig', 'logsig', 'purelin'}, 'traingda')

Network C: net = newff(minmax(p), [50, 25, 1], {'logsig', 'logsig', 'purelin'}, 'traingdx')

Network D: `net = newff(minmax(p), [50, 25, 1], {'tansig', 'tansig', 'purelin'}, 'traingdx')`

Using Network A as an example, ‘newff’ creates a new feed-forward network with the defined properties, the minimum and maximum values for the input, `p`, are with defined or, in this case scanned for, the number of neurons in each layer is defined, three layers with 100, 50, and 1 neuron respectively, the transfer functions used by each layer are defined, and finally the learning function is defined.



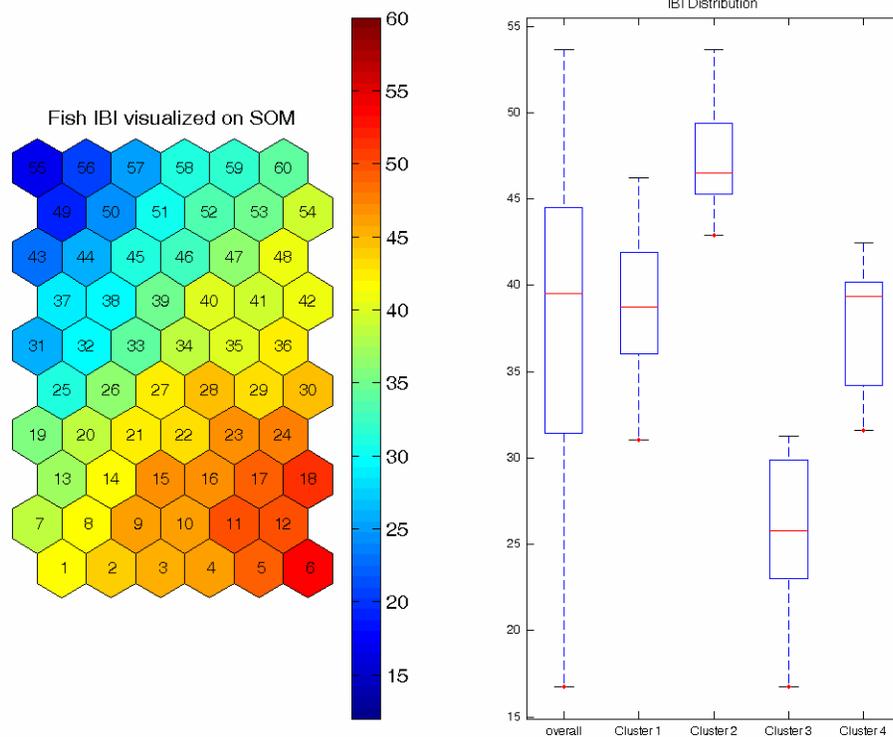
**Figure 4.3** Examples of simulations results for Group 1 parameters

The other groups used often had similar results with performances that were close to those obtained using Group 1.

### 4.3.2 Improving Network Results

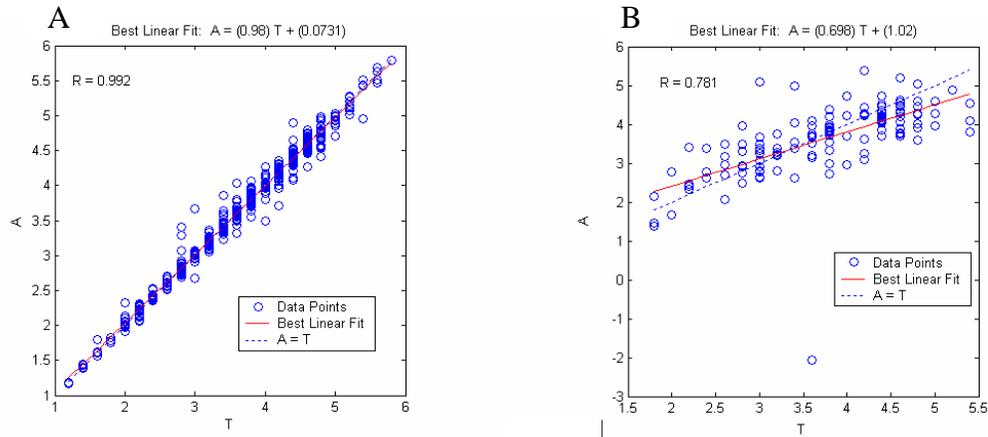
The quality of network predictions depends on the information provided to it for training. It can be difficult to identify those parameters which have the most influence on the output. For the preliminary analysis, the method for choosing parameters to be used for the training was based on availability of professional judgment. This approach is not ideal as it is possible to overlook strongly contributing parameters. Traditionally, unsupervised neural network analysis is performed before supervised learning. Unsupervised analysis can be used to identify relationships between parameters and biological integrity, the inputs and targets used in by the supervised network. Work is currently being done to understand the relationships in the Ohio database (Virani 2005).

Virani (2005) used Self Organizing Maps, SOMs, to pattern fish metrics in the Ohio database. The patterning organizes monitoring sites on a grid of neurons based on the similarities in the characteristics of the fish metrics measured at the sites. The results of the mapping were then analyzed to develop clusters of sites with similar responses. The range of IBI and water quality and habitat parameters, based on an average of all the sampling sites patterned at each neuron, of the clusters were then compared using canonical correspondence analysis, CCA, to define the most influential parameters. This information can then be used to select parameters for use in training the supervised network and is likely to results in improved predictions. Figure 4.4 shows an example of the clusters developed from a SOM.



**Figure 4.4 Clusters resulting from Self Organizing Map analysis of Ohio database from Virani (2005)**

The results of the unsupervised network analysis of the Ohio database suggested that habitat parameters are very influential in determining the biological community quality. Over half of the top 15 parameters were habitat measures. Based on these results, input parameters used in the supervised network analysis was changed to include many of the habitat parameters available in the database. This produced distinct improvements in the results of prediction. Figure 4.5 below, shows the calibration of a network using Group 2 parameters and the results of the simulation for the same network. Barring one outlier, the results are encouraging.



**Figure 4.5** A) Calibration and B) simulation results for network trained with Group 2

**Table 4.3** Network training Group 2

Description
RIFFLE
POOL
EMBEDEDDNESS
CHANNEL
SUBSTRATE
RIPARIAN
CALCIUM
PERCENT FORRESTED
HARDNESS
TOTAL IRON (UG/L)
BOD, 5-DAY, (MG/L)
DISSOLVED OXYGEN (MG/L)
INVERTEBRATE COMMUNITY INDEX
TOTAL KEJDAL NITROGEN

The preliminary results from network analysis suggest that prediction of biotic integrity within a reasonable range is possible. Continued investigation into the appropriate input parameters to be used for training a network is likely to result in more accurate predictions. Applications for predictions of biotic integrity have a wide range, from analyzing the effects of new development to determining the most efficient way to use rehabilitation funds. It is likely that neural networks could be applied in many other ecological setting and be expected to have positive results.



## **5 Conclusions and Recommendations**

Protection of ecological integrity in waters of the United States is often approached through three different paths, chemical, physical, and biological. For a large portion of history, watershed managers were primarily concerned with chemical integrity as it was conceived to be the limiting factor in watershed health. However, as treatment improved, continued problems with physical and biological integrity surfaced. Chemical and physical integrity are relatively independent of each other, but biological integrity is largely contingent on the local physical and chemical quality. More recent assessments have begun to focus on biological integrity as a more comprehensive measure of watershed health.

Biological integrity is an endpoint that can be affected by all watershed features. Although most actions in a watershed do not affect the biology directly, their effects propagate through a hierarchy of watershed layers that eventually reach biology. A change in a watershed feature, or root stressor, changes the input into the waterbody which changes the in-stream chemical or physical quality. Such changes present the possibility, or risk, that the biological integrity will be altered.

The ability to calculate the risk to biological integrity from root stressors would aid watershed decision makers and planners who could use the information to choose alternatives that minimize risk upfront. The information could also be used to evaluate restoration alternatives to select the most effective plan. Current research has shown that increased urbanization, using as % imperviousness as a surrogate measure, can be correlated with a reduction in biological integrity. However, this does not provide the specificity needed to make comparisons of the risk of one project over another or the ability to identify possibilities for restoration.

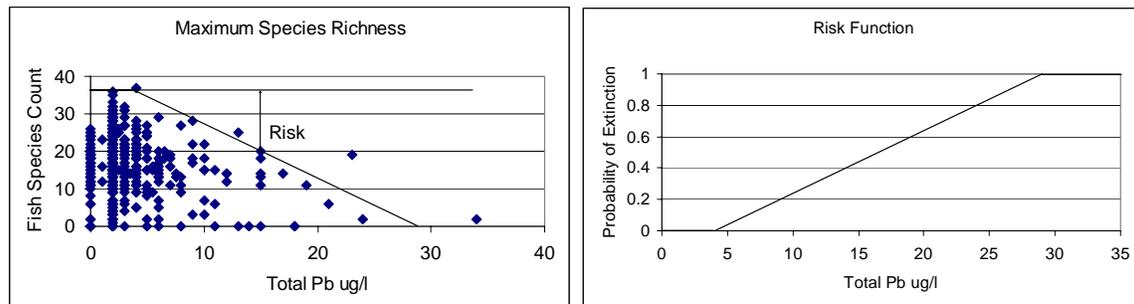
As a starting point, this study calculated risk to biology from in-stream measurements. Calculations were made using two methods which can be used to determine risk from single stressors. Multiple stressors were also considered for the prediction of biological integrity represented by IBI. These approaches could be used to estimate projects' effects on aquatic biology. Future work should be done to connect both approaches to root stresses in the watershed.

### **5.1 Risk Calculation**

Laboratory toxicity data and field measurements were both used in different approaches to calculate risks to in-stream biology. They both measure risk based on single stressors, but their results may be different. This is because they measure very different kinds of risk.

When risk was calculated using LC<sub>50</sub>s, two distributions were considered; the distribution of effects and the distribution of concentration. It represents the risk that a single organism will be adversely affected by the ongoing conditions in a stream. The risk calculated using Maximum Species Richness Lines, MSRLs, was the risk to the entire

group being examined and not just a single organism within the group. Moreover, it did not consider the distribution of stressor concentrations; rather it is the risk from a single constant concentration. If the risk was defined as a distribution as it was with  $LC_{50}$ s, it would be possible to consider the risk from a distribution of concentrations. This could be done by defining the cumulative distribution of risk as the inverse of the MSRL. Figure 5.1 illustrates this concept. The probability of extinction using to plot the risk function is equivalent to plotting the percent of species affected, which was used to describe toxic response. The WERF methodology could then be used to determine the risk as a joint probability of the distribution of concentrations of the stressor and the risk function developed from the MSRL (Parkhurst et al. 1996a).



**Figure 5.1** Developing risk functions from Maximum Species Richness

## 5.2 Addressing Limitations of Risk Calculation

The risk calculation methods that were examined can only consider risk from a single stressor at a time. This is the primary limitation with current risk estimation methods. When multiple stressors are present in a system, they assume that the risks from each individual stressor is additive. This may not in fact be the case for all stressor combinations. It is possible that certain combinations of stressors are synergistic and present more than just the addition of their individual risks. Conversely, stressors may interfere with each other and present less than additive risk in combination.

There are also limitations specific to each method. One of the major difficulties with using toxicity data to calculate risk arose from the range of concentrations at the sample site. Ambient concentrations are usually much lower than  $LC_{50}$ s, even for the most sensitive species. This makes risk calculation very sensitive to values in the lower tail of the distribution of  $LC_{50}$ s, which is also a very difficult area to define because of the lack of data. Furthermore, the differences in distributions are often very high in this region, making risk calculation very sensitive to the choice of distribution. However, if a distribution can adequately describe the lower region of concentration then there can be high confidence in the results because the data used to develop the distributions come from laboratory tests performed under the same conditions. There is also some minor geographic limitation to this method. The risk can be calculated for the distribution of concentration in any stream as long as the species used to develop the distribution of effects are native to that region.

MSRLs are developed from field measurements and are therefore dependent on the range of quality encountered in the field. This means that there are usually a large number of measurements available at low concentrations of stress and fewer at high concentrations. It can be difficult to develop reliable MSRLs for higher concentrations with few measurements. However, this also means that the response is very well defined in the ranges of concentration most commonly encountered. MSRLs are most applicable to the area where the field measurements they were developed from were taken; however, they are relatively easy to develop from new areas if the data is available. If the data is not available MSRLs from similar watersheds may be applicable though they should be used with some caution.

Overall, both approaches to calculating risk provide quality measurements when used appropriately. They can both be used to give a reliable measurement of risk from single stressors. They also have applications beyond just calculating risk. By using particular groups of organisms such as groups based on tolerance of feeding function, the reactions of the groups can be understood and compared. This reveals information about the way in which stressors affect the environment. For example, if insectivores have a stronger reaction of a certain stressor it may be because the stressor affects the food source before it reaches a concentration that would affect the fish directly. Such information could be used to help build links to the next layer in the hierarchy of stressor effects.

### **5.3 Artificial Neural Networks**

In a preliminary analysis, artificial neural networks were used to predict IBI based on measurements of water and habitat quality. If IBI could be predicted with confidence, the predictions could be used to evaluate projects in the watershed without needing to go through the sometimes lengthy process of calculating and interpreting risk. Neural network predictions also remove the question of the reactions of stressors when combined. Any non-linear relationships between stressors would be learned by the network directly from the training data.

The results are positive, particularly in those cases where individual habitat parameter were included, and indicate the further study in this area should be done. They have shown that the network is able to learn the general trends in the ecological data. Continued work on developing the self-organizing maps (SOMs) can be used to determine the most appropriate parameters to use as input to the supervised network. It is likely additional improvements can be made through other changes or additions to the input parameter.

### **5.4 Linking Biology and Watershed**

The ultimate goal is to be able to predict biological integrity from watershed features and actions. This is very difficult because of there are often numerous changes taking place at a time. This may require a layering of predictions similar to the layers suggested for risk propagation through the watershed hierarchy. Water body quality is constrained by the geomorphic features of the region and if affected by the actions in the contributing

region. Petts (1980) describes the impact of human alteration in three stages; change in pollutant loads leads to a change in water quality and stream morphology and eventually affect biology. The current network models the final stage of impact. If models were developed for the beginning two layers they could all be linked by having the output of one model act as the input to the next. A successful model would make it possible to predict the affect of human activity on in-stream biology.

Current research has already identified many links between watershed qualities and activities and biological integrity. However, much like risk, the links have only been made for one feature at a time and some research has had conflicting results. Moreover, even though some specific metric values show a relationship to certain watershed features; the relationships are usually best expressed by local habitat variables rather than more distant watershed uses.

Urbanization and other land uses are threatening water bodies across the United States. In recent years the focus of water quality management has shifted from viewing the water body as an independent entity to viewing it as part of the watershed system. This shift has brought recognition of the complexity of the interactions between landscape and stream integrity. Traditional deterministic models are limited in their ability to estimate environmental risk on a broad scale because of the uncertainties in the process and the interaction of multiple stressors. Land use/cover data, as well as physical and chemical data, have been shown to be correlated with the biological integrity of streams. Development of new models that can cope with such diverse influences is necessary for water reclamation and protection to advance.

## **5.5 Recommendations for Future Work**

There are several steps that can be taken that may significantly improve the network results. The quality of any network prediction is largely dependent on the quality of the data and parameters used to train it. If the correct parameters are not used in training, the network will not be able to predict the target. Identifying the most influential parameters is an important step in network development.

SOMs are currently being developed for the Ohio database and are being used identify clusters among the monitoring sites (Virani 2005). This information can be used to identify the most influential parameters available in the database. Once identified, they should be used to train networks for IBI prediction. Preliminary results from this work have shown that the individual habitat metrics, such as those measuring riffle and pool quality, and embeddedness are some of the most influential parameters. Additional work in this area may reveal more about which parameters are most appropriate for training.

It may also be important to add some inputs that are not currently considered. Fragmentation is considered a major category of stress for biology and is not given much emphasis by the parameters available in the Ohio database. Fragmentation is considered in part by the habitat metric that measures channelization, but a more direct measure may be helpful. Impoundments are the primary form of fragmentation and can have significant

effects on biological integrity (Luttrell et al. 1999; Mammoliti 2002). Because they have such a strong influence over biological communities, they should be included in neural network model whenever possible. The input to the network may be binary, presence or absence of impoundments at the site, or possibly some representation of distance to impoundments. Information on the location of impoundments in the United States is available from in the National Inventory of Dams (U.S. Army Corps of Engineers 1999).

Improvements might also be made by using a different type of input. In the proposed hierarchy, biological inputs are linked to the risk from in-stream parameters. A network could be developed that uses risks as inputs instead of parameter measurements. The risk values could be calculated using either toxic response or MSRLs. Using a network to analyze the risk values from multiple stressors may reveal information about their interactions, i.e. additive, synergistic, etc.

In addition to improvements in the input, there may be an opportunity to improve results based on the target values. One of the primary problems with the current results is that the predictions are largely limited to the mid-range of IBI. It is likely that this occurs because the bulk of the training data has IBI values in this range. The extreme values of IBI may be underrepresented and therefore the network has difficulty generalizing in these ranges. The network predictions may improve if the training set is developed so that target values are more evenly distributed over the entire range of IBI.



## References

- Allen, T. F. H. and T. B. Starr (1982). *Hierarchy - Perspectives for Ecological Complexity*. Chicago, IL, The University of Chicago Press.
- Bartošová, A. (2002). *Estimating the Ecological Risk to Aquatic Biota from Physical and Chemical Impairment Due to Urbanization*. Department of Civil and Environmental Engineering. Ph.D. Dissertation. Milwaukee, WI, Marquette University.
- Davis, W. S. (1995). *Biological Assessment and Criteria: Building on the Past*. Biological Assessment and Criteria. W. S. Davis and T. P. Simon. Boca Raton, FL, Lewis Publishers: 15-29.
- Demuth, H. and M. Beale (1992). *Neural Network Toolbox User's Guide*, The MathWorks Inc. Natick, MA.
- Faribrother, A. and R. S. Bennett (2000). *Multivariate statistical applications for addressing multiple stresses in ecological risk assessments*. Multiple Stressors in Ecological Risk and Impact Assessment: Approaches in Risk Estimation. S. A. Ferenc and J. A. Foran. Pensacola, FL, SETAC Press: 69-115.
- Fausch, K. D., J. R. Karr and Y. P.Y. (1984). "Regional application of an index of biotic integrity on stream fish communities." *T. Am. Fish. Soc.* **115**: 39-55.
- Frissell, C. A., W. J. Liss, C. E. Warren and M. D. Hurley (1986). "A Hierarchical Framework for Stream Habitat Classification: Viewing Streams in a Watershed Context." *Environmental Management* **10**(2): 199-214.
- Karr, J. R. (1995). *Protecting Aquatic Ecosystems: Clean Water Is Not Enough*. Biological Assessment and Criteria. W. S. Davis and T. P. Simon. Boca Raton, FL, Lewis Publishers: 7-13.
- Karr, J. R. and E. W. Chu (1999). *Restoring Life to Running Waters*. Washington DC, Island Press.
- Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant and I. J. Schlosser (1986). *Assessing Biological Integrity in Running Waters: A Method and Its Rationale*. Illinois Natural History Survey. Champaign, IL. Special Publ. 5.
- Karr, J. R., R. C. Heidinger and E. H. Helmer (1985). "Effects of chlorine and ammonia from wastewater treatment facilities on biotic integrity." *Journal WPCF* **57**: 912-915.

- Leigh, D. S., B. J. Freeman, M. C. Freeman, E. A. Kramer, C. M. Pringle and A. D. Rosemond (1999). 1999 Progress Report: Land Use and Geomorphic Indicators of Biotic Integrity in Piedmont Streams.
- Leopold, L. B., M. G. Wolman and J. P. Miller (1994). *Fluvial Processes in Geomorphology*. New York, NY, Dower Publications.
- Luttrell, G. R., A. A. Echelle, W. L. Fisher and D. J. Eisenhour (1999). "Declining status of two species of the *Macrhybopsis aestivalis* complex (Teleostei: Cyprinidae) in the Arkansas River Basin and related effects of reservoirs as barriers to dispersal." *Copeia* **4**: 981-989.
- Mammoliti, C. S. (2002). "The Effects of Small Watershed Impoundments on Native Stream Fishes: A Focus on the Topeka Shiner and Hornyhead Chub." *Transactions of the Kansas Academy of Science* **105**(3): 219-231.
- Novotny, V. (2003). *Water Quality: Diffuse Pollution and Watershed Management*. Hoboken, NJ, J. Wiley.
- Novotny, V., A. Bartošová, N. O'Reilly and T. Ehlinger (2005). "Unlocking the relationship of biotic integrity of impaired waters to anthropogenic stresses." *Water Research* **39**: 184-198.
- Novotny, V., D. Clark, R. Griffin, A. Bartosova and D. Booth (2001). *Risk Based Urban Watershed Management - Integration of Water Quality and Flood Control Objectives, Final Report*. Institute for Urban Environmental Risk Management, Marquette University. Milwaukee, WI.
- Novotny, V. and J. W. Witte (1997). "Ascertaining aquatic ecological risks of urban stormwater discharges." *Water Resources* **31**(10): 2573-2585.
- Odum, E.P, J.T. Finn and E.H. Franz (1979). Perturbation theory and the subsidy-stress gradient. *Bioscience* **29**: 349-352.
- Ohio EPA. (1987). *Biological criteria for the protection of aquatic life: Volume II. Users manual for biological field assessment of Ohio surface waters*. Division of Water Quality Monitoring and Assessment, Surface Water Section. Columbus, Ohio.
- Parkhurst, B., W. Warren-Hicks, R. D. Cardwell, J. Volosin, T. Etchinson, J. B. Butcher and S. M. Covington (1996a). *Aquatic Ecological Risk Assessment: A Multi-Tiered Approach*. Water Environment Research Foundation. Alexandria, VA. Final Report, Project 91-AER-1.
- Parkhurst, B., W. Warren-Hicks and C. Creager (1996b). *Modification of the WERF Methodology for Aquatic Ecological Assessment for Assessing Watershed-Scale Aquatic Risks*. Watershed '96 Conference Proceedings, Baltimore, MD.

- Paulson, C. and B. R. Parkhurst (1996). Application of Aquatic Ecological Risk Methodology to Support Site-Specific Water Quality-Based Permitting. Watershed '96, Conference Proceedings, Baltimore, MD.
- Petts, G. E. (1980). "Long-term consequences of upstream impoundment." *Environmental Conservation* 7(4): 325-332.
- Rankin, E. T. (1995). Habitat Indices in Water Resource Quality Assessments. Biological Assessment and Criteria. W. S. Davis and T. P. Simon. Boca Raton, FL, Lewis Publishers: 181-208.
- Rankin, E. T. and C. O. Yoder (1999). Methods for Deriving Maximum Species Richness Lines and Other Threshold Relationships in Biological Field Data. Assessing the Sustainability and Biological Integrity of Water Resources using Fish Communities. T. P. Simon. Boca Raton, FL, CRC Press: 611-624.
- Schueler, T. (1994). "The importance of imperviousness." *Watershed Protection Techniques* 1(3): 100-111.
- Swanson, S. M. (2005). Multiple Stressors: Literature Review and Gap Analysis. Water Environment Research Foundation. Alexandria, VA. Project Report 00-ECO-2B.
- U.S. Army Corps of Engineers (1999). <http://www.nicar.org/data/dams/>. 2004
- USEPA (1994). Water Quality Standards Handbook: Second Edition. Office of Water. Washington, DC. EPA-823-B-94-005.
- USEPA (2000). National Water Quality Inventory: 2000 Report. Office of Water. Washington, DC. EPA-841-R-02-001.
- Virani, H. (2005). Self Organizing Feature Maps combined with ecological ordination techniques for effective watershed management. Department of Electrical and Computer Engineering. Masters Thesis. Boston, MA, Northeastern University.
- Yoder, C. O. and E. T. Rankin (1995). Biological Response Signatures and the Area of Degredation Value: New Tools for Interpreting Multimetric Data. Biological Assessment and Criteria. W. S. Davis and T. P. Simon. Boca Raton, FL, Lewis Publishers: 263-286.
- Younus, M., M. Hondzo and B. Engel (1999). "Stream temperature dynamics in upland agricultural watersheds: Measurements and modeling." *Journal of Environmental Engineering* 126(6): 518-526.