Development of Biological Criteria for Tiered Aquatic Life Uses

Fish and macroinvertebrate thresholds for attainment of aquatic life use goals in Minnesota streams and rivers

Minnesota Pollution Control Agency

June 2016

Authors

R. William Bouchard, Jr.

Contributors/acknowledgements

Minnesota Pollution Control Agency:

Joel Chirhart, Brenda DeZiel, Mike Feist, John Genet, Dan Helwig, Ben Lundeen, Scott Niemela, John Sandberg, Bryan Spindler, Kevin Stroom, Mark Tomasek

Midwest Biodiversity Institute:

Ed Rankin, Chris Yoder

The MPCA is reducing printing and mailing costs by using the Internet to distribute reports and information to wider audience. Visit our website for more information.

MPCA reports are printed on 100% post-consumer recycled content paper manufactured without chlorine or chlorine derivatives.

Recommended citation:

MPCA (2016) Development of biological criteria for tiered aquatic life uses: Fish and macroinvertebrate thresholds for attainment of aquatic life use goals in Minnesota streams and rivers. Minnesota Pollution Control Agency, Environmental Analysis and Outcomes Division, St. Paul, MN

Minnesota Pollution Control Agency

520 Lafayette Road North | Saint Paul, MN 55155-4194 | www.pca.state.mn.us | 651-296-6300

Toll free 800-657-3864 | TTY 651-282-5332

This report is available in alternative formats upon request, and online at www.pca.state.mn.us .

Document number: wq-bsm4-02

Contents

List of tables

List of figures

List of acronyms

AWC	Altered Watercourse
BCG	Biological Condition Gradient
CWA	Clean Water Act
FIBI	Fish Index of Biological Integrity
GP	Low Gradient (Glide/Pool Habitat)
HDS	Human Disturbance Score
IBI	Index of Biological Integrity
IWM	Intensive Watershed Monitoring
LG	Low Gradient Streams
MIBI	Macroinvertebrate Index of Biological Integrity
MPCA	Minnesota Pollution Control Agency
MSHA	Minnesota Stream Habitat Assessment
NC	Northern Coldwater
NFGP	Low Gradient Northern Forest Streams
NFR	Northern Forest Rivers
NFRR	High Gradient Northern Forest Streams
NH	Northern Headwaters
NPDES	National Pollutant Discharge Elimination System
NR	Northern Rivers
NS	Northern Streams
PFR	Prairie Forest Rivers
PSGP	Low Gradient Prairie Streams
RR	High Gradient (Riffle/Run Habitat)
RCP	Reference Condition Percentile
SC	Southern Coldwater
SFGP	Low Gradient Southern Forest Streams
SH	Southern Headwaters
SR	Southern Rivers
SS	Southern Streams
SSRR	High Gradient Southern Streams
TALU	Tiered Aquatic Life Use
TMDL	Total Maximum Daily Load
UAA	Use Attainability Analysis
WQS	Water Quality Standards

Glossary of terms

Antidegradation: The part of state water quality standards that protects existing uses, prevents degradation of high quality water bodies unless certain determinations are made, and which protects the quality of outstanding national resource waters. (Currently nondegredation in Minnesota)

Beneficial Use: Desirable uses that acceptable water quality should support. Examples are drinking water supply, primary contact recreation (such as swimming), and aquatic life support.

Best Management Practice (BMP): An engineered structure or management activity, or combination of these that eliminates or reduces an adverse environmental effect of a pollutant, pollution, or stressor effect.

Biological Assessment: An evaluation of the biological condition of a waterbody using surveys of the structure and function of a community of resident biota; also known as bioassessment. It also includes the interdisciplinary process of determining condition and relating that condition to chemical, physical, and biological factors that are measured along with the biological sampling.

Biological Criteria (Biocriteria):

> **Scientific meaning:** quantified values representing the biological condition of a waterbody as measured by the structure and function of the aquatic communities typically at reference condition; also known as biocriteria.
>
> **Regulatory meaning:** narrative descriptions or numerical values of the structure and function of aquatic communities in a waterbody necessary to protect a designated aquatic life use, implemented in, or through state water quality standards.

Biological Condition Gradient (BCG): A scientific model that describes the biological responses within an aquatic ecosystem to the increasing effects of stressors.

Biological Integrity: The ability of an aquatic ecosystem to support and maintain a balanced, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitats within a region (after Karr and Dudley 1981).

Biological Monitoring: The use of a biological entity (taxon, species, assemblage) as an indicator of environmental conditions. Ambient biological surveys and toxicity tests are common biological monitoring methods; also known as biomonitoring.

Clean Water Act (CWA): An act passed by the U.S. Congress to control water pollution (formally referred to as the Federal Water Pollution Control Act of 1972). Public Law 92-500, as amended. 33 U.S.C. 1251 et seq.; referred to herein as the Act.

Criteria: A limit on a particular pollutant or condition of a waterbody presumed to support or protect the designated use or uses of a waterbody. Criteria may be narrative or numeric and are commonly expressed as a chemical concentration, a physical parameter, or a biological assemblage endpoint.

Designated Use: see Beneficial Use.

Ecoregion: A relatively homogeneous geographical area defined by a similarity of climate, landform, soil, potential natural vegetation, hydrology, or other ecologically relevant variables; ecoregions are portioned at increasing levels of spatial detail from Level I to Level IV.

Index of Biological Integrity (IBI): IBI refers to the index developed by Karr (1981) and explained by Karr et al. (1986). The IBI is a numerical index that is comprised of various measures of the biological community (called metrics) that are assigned a score (typically 0-10) based on their deviation from reference and summed to provide an integrative expression of site condition. It has been used to express the condition of fish, macroinvertebrate, algal, and terrestrial assemblages throughout the United States and in each of five major continents.

Macroinvertebrates: Animals without backbones, living in or on the substrates, of a size large enough to be seen by the unaided eye, and which can be retained by a U.S. Standard No. 30 sieve (0.595 mm openings). Also referred to as benthos, infauna, or macrobenthos.

Narrative Biocriteria: Written statements describing the narrative attributes of the structure and function of aquatic communities in a waterbody necessary to protect a designated aquatic life use.

Natural Condition: This includes the multiplicity of factors that determine the physical, chemical, or biological conditions that would exist in a waterbody in the absence of measurable impacts from human activity or influence.

Numeric Biocriteria: Specific quantitative measures of the structure and function of aquatic communities in a waterbody necessary to protect a designated aquatic life use.

Reference Condition: The condition that approximates natural, unimpacted to best attainable conditions (biological, chemical, physical, etc.) for a waterbody. Reference condition is best determined by collecting measurements at a number of sites in a similar waterbody class or region under minimally or least disturbed conditions (by human activity), if they exist. Since undisturbed or minimally disturbed conditions may be difficult or impossible to find in some states, least disturbed conditions, combined with historical information, models or other methods may be used to approximate reference condition as long as the departure from natural or ideal is known. Reference condition is used as a benchmark to establish numeric biocriteria and can be further described as follows:

> **Minimally Disturbed Condition (MDC)** – This term describes the condition of the biota in the absence of significant human disturbance and it is the best approximation of biological integrity.
>
> **Historical Condition (HC)** - The condition of the biota at some point in its history. It may be a more accurate estimator of true reference condition (i.e., biological integrity) if the historical point chosen is before the effect of any adverse human disturbance. However, more than one historical reference point is possible (e.g., pre-industrial, pre-Columbian).
>
> **Least Disturbed Condition (LDC)** – Least disturbed condition is found in conjunction with the best available physical, chemical, and biological habitat conditions given today's state of the landscape.
>
> **Best Attainable Condition (BAC)** – This is the expected condition of least disturbed sites under the implementation of best management practices for a sufficient period of time. This is a condition that results from the convergence of management goals, best available technologies, and a public commitment to achieving environmental goals (e.g., as established by WQS) under prevailing uses of the landscape. BAC may be equivalent to either MDC or LDC depending on the prevailing level of human disturbance in a region.

Reference Site: A site selected to represent reference condition. For the purpose of assessing the ecological condition of other sites, a reference site is a specific locality on a waterbody that is minimally or least disturbed and is representative of the expected ecological condition of similar waterbodies.

Regional Reference Condition: A description of the chemical, physical, or biological condition based on an aggregation of data from reference sites that are representative of a waterbody type within a region (e.g. ecoregion, subregion, bioregion, or major drainage unit).

Stressors: Physical, chemical, and biological factors that can adversely affect aquatic organisms. The effect of stressors is apparent in the biological responses.

Use Attainability Analysis (UAA): A structured scientific assessment of the physical, chemical, biological or economic factors affecting attainment of the uses of waterbodies.

Use Classes: A broad capture of a designated use for general purposes such as recreation, water supply, and aquatic life.

Tiered Aquatic Life Uses (TALUs):

> **As defined:** The structure of designated aquatic life uses that incorporates a hierarchy of use subclasses and stratification by natural divisions that pertain to geographical and waterbody class strata. TALUs are based on representative ecological attributes reflected in the narrative description of each TALU tier and embodied in the measurements that extend to expressions of that narrative through numeric biocriteria and by extension to chemical and physical indicators and criteria.
>
> **As used:** TALUs are assigned to water bodies based on the protection and restoration of ecological potential. This means that the assignment of a TALU tier to a specific waterbody is done with regard to reasonable restoration or protection expectations and attainability. Hence knowledge of the current condition of a waterbody and an accompanying and adequate assessment of stressors affecting that waterbody are needed to make these assignments.

Total Maximum Daily Load (TMDL): The maximum amount of a pollutant that a body of water can receive while still meeting water quality standards. Alternatively, a TMDL is an allocation of a water pollutant deemed acceptable to attain the designated use assigned to the receiving water.

Water Quality Standards (WQS): A law or regulation that consists of the designated use or uses of a waterbody, the narrative or numerical water quality criteria (including biocriteria) that are necessary to protect the use or uses of that particular waterbody, and an antidegradation policy.

Water Quality Management: A collection of management programs relevant to water resource protection that includes problem identification, the need for and placement of best management practices, pollution abatement actions, and measuring the effectiveness of management actions.

1. Overview

This report documents the development of biological criteria or biocriteria used to assess attainment of Minnesota's aquatic life use goals including the General Use goal and Tiered Aquatic Life Use goals. More detailed descriptions of biomonitoring, bioassessment, and Tiered Aquatic Life Use components related to the development of biocriteria including biological assessment guidance (MPCA 2012), stream classification (MPCA 2014b, a), human disturbance score (MPCA 2014c), biological condition gradient (BCG) (Gerritsen et al. 2013), and Indices of Biological Integrity (IBI) (MPCA 2014b, a) can be found in other documents. Minnesota has used IBI and chemical measures together to assess the integrity of streams since the mid-1990s. Both biological and chemical monitoring efforts are integral to the assessment of Minnesota's beneficial uses, including aquatic life uses. Monitoring programs for the protection of aquatic life that do not monitor biological communities are at risk of missing impairments. Biological assessments are a particularly powerful tool as they provide a more accurate measure of the condition of the biological communities and are a direct determinant of the attainment of aquatic life uses. As a result, the development and implementation of a robust biological monitoring and assessment program is integral to Minnesota's goals of protecting and restoring the integrity of aquatic resources.

Minnesota is an ecologically diverse state with water resources spanning a wide range of conditions. This diversity presents management challenges and as a result, Minnesota's current one-size-fits-all approach (i.e., General Use alone) results in over or under protection of some waters. Tiered Aquatic Life Uses or TALUs provide the framework to designate uses that are attainable thereby giving greater protection to high quality waters and setting appropriate goals for systems impacted by legacy uses (e.g., channelization). A TALU framework results in more accurate assessments as they are defined by attainable conditions in Minnesota's streams.

The development of biocriteria in Minnesota used a multiple lines of evidence approach which relied most heavily on Reference Condition and the BCG. The Reference Condition is the traditional approach for setting biocriteria, but this methodology alone was not sufficient for setting accurate TALU biocriteria that reflect Minnesota's aquatic life use goals. As a result, both methods were used together to strengthen Minnesota's approach to setting biocriteria. A comparison of the biological thresholds developed using each method demonstrated that the results were similar which resulted in greater confidence in the biocriteria. This document details the development of these approaches and how they were used together to develop Exceptional, General, and Modified Use biocriteria for Minnesota streams.

2. Introduction

2.1. The need for biological criteria

The objective of the Clean Water Act (CWA) is to "*restore and maintain the chemical, physical, and biological integrity of the Nation's waters"* (U.S. Code title 33, section 1251 [a]). Although this statement is central to the CWA, interpreting this language and putting this into practice is more difficult. Following adoption of the CWA, a debate began regarding how to define and measure "biological integrity". From this discussion a definition of biological integrity was put forward by Karr and Dudley (1981) as:

> ***"the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region."***

This definition continues to be widely accepted and serves to guide protection and maintenance of the integrity of waters in the United States. In addition to this objective, the CWA provides an interim goal for the Nations waters:

> ***"wherever attainable, an interim goal of water quality which provides for the protection and propagation of fish, shellfish, and wildlife and provides for recreation in and on the water" (U.S. Code title 33, section 1251 [a] [2])***

Given these goals and descriptions, it is then possible to develop water quality standards to protect aquatic life uses. Development of biocriteria for Minnesota streams will result in unambiguous goals and provide a more direct assessment of biological condition, thereby resulting in better outcomes for Minnesota's waters. The development of tiered statewide biocriteria for streams in Minnesota is a further refinement to Minnesota's water quality standards which recognizes that there are differences in the potential for restoration and protection among waters. Under a TALU framework, biocriteria serve two main purposes: 1) determining the beneficial use of a waterbody and 2) determining attainment of the beneficial use. In addition to these central goals, the data collected to support a TALU framework also provides information that can enhance other watershed protection tools such as water quality standards, stressor identification, Total Maximum Daily Load (TMDLs), watershed planning, and National Pollutant Discharge Elimination System (NPDES) permitting.

2.2. Minnesota's water quality standards

States, tribes, and territories are responsible for adopting, revising, and implementing water quality standards. Water quality standards (WQS) are comprised of three main elements: 1) Beneficial Uses, 2) Numeric and Narrative Criteria, and 3) Antidegradation. Beneficial uses and criteria define who or what we are protecting and the criteria define the conditions that are protective of those uses. Antidegradation provides additional protection to existing uses especially high quality and unique waters. In Minnesota, beneficial uses include drinking water, aquatic life, recreation, and agricultural uses; however the beneficial use that is most relevant to biocriteria and TALU is Class 2: Aquatic Life and Recreation. This use class is defined in rule as:

> ***"Aquatic life and recreation includes all waters of the state that support or may support fish, other aquatic life, bathing, boating, or other recreational purposes and for which quality control is or may be necessary to protect aquatic or terrestrial life or their habitats or the public health, safety, or welfare." [Minn. R. ch. 7050.0140 subp. 3]***

Minnesota's narrative standards for the protection of aquatic life uses in Class 2 waters are as follows:

> ***"For all Class 2 waters, the aquatic habitat, which includes the waters of the state and stream bed, shall not be degraded in any material manner, there shall be no material increase in undesirable slime growths or aquatic plants, including algae, nor shall there be any significant increase in harmful pesticide or other residues in the waters, sediments, and aquatic flora and fauna; the normal fishery and lower aquatic biota upon which it is dependent and the use thereof shall not be seriously impaired or endangered, the species composition shall not be altered materially, and the propagation or migration of the fish and other biota normally present shall not be prevented or hindered by the discharge of any sewage, industrial waste, or other wastes to the waters."***
> ***[Minn. R. ch. 7050.0150 subp. 3]***

To protect or restore aquatic life and other beneficial uses, criteria are used to define the conditions that will be protective and thereby sets the goals for waters. These criteria can be chemical, physical, or biological. The use of biocriteria has the advantage of directly measuring attainment of the aquatic life use and is less likely to miss impairments that chemical or physical measures alone may not identify (Yoder 1995). This is driven by two major attributes of biological communities:

1) Biological communities such as fish and macroinvertebrates are relatively long lived so stresses in the environment, even if they are intermittent and/or short lived, are often reflected in the condition of biological communities.
2) Biological communities integrate the effects of multiple stressors over time so impacts that might be missed because the relevant chemical or physical parameter was not measured will be identified by changes in these communities.

The use of biological communities in assessments also has the advantage of translating the condition of a waterbody into terms that are more relatable to the public. As a result, biocriteria along with chemical criteria are integral to a state's CWA program which seeks to protect and restore the integrity of its waters. In addition, the U.S. Environmental Protection Agency's (EPA) policy is that states incorporate biological assessments into water quality standards programs (USEPA 1990, 2011).

Minnesota's current WQS framework is a one-size-fits-all approach which applies a General Use for the protection of aquatic life to all streams and rivers of the state. The recommended revised framework includes four tiers for the protection of aquatic life: Exceptional, General, and Modified. These tiered uses are described in Yoder (2012) and the narratives are as follows:

> ***Exceptional*** *- These are waters that exhibit the highest quality of "exceptional" assemblages (as measured by assemblage attributes and indices) on a Minnesota Biological Condition Gradient (BCG) basis; narrative descriptors such as "exceptional" can be used as the distinguishing descriptors in the designated use narrative, but other descriptive terms are possible. These communities have minimal changes in structure of the biotic assemblage and in ecosystem function which is the ultimate goal of the CWA. It functions as a preservation use, which means it is intended for waters that already exhibit or have the realistic potential to attain an exceptional quality as measured by the biological criteria.*

> ***General*** *– These are waters that harbor "typically good" assemblages of freshwater organisms (as measured by assemblage attributes and indices) and that reflect the lower range of the central tendency of "least impacted" regional reference condition. In the language of the BCG, they are communities that can be characterized as possessing "*overall balanced distribution of all expected major groups; ecosystem functions largely maintained through redundant attributes*". As such this use represents the minimum CWA goal attainment threshold and it serves as the principal*

restoration use for management programs. It also serves as the "triggering threshold" for when a UAA is required to determine the attainability of this designated use tier for specific river or stream segments.

Modified *– These are waters that have been extensively altered and currently exhibit legacy physical modifications that pre-date the November 28, 1975 existing use date in the Federal Water Quality regulations (40CFR Part 131). These waters have been determined to be in non-attainment of the General use biological criteria and have been determined to be incapable of attaining those criteria via a UAA. The biological criteria for the Modified use are established based on a separate population of "modified reference sites" that exhibit these types of modifications with little presence of other types of stressors. Possible subcategories include channelization for flood control and agricultural drainage and impoundments created by run-of-river low head dams. Separate reference populations are needed to derive the numeric biocriteria for each subcategory.*

These refined uses will result in the protection of good and high quality waters while setting attainable goals for waters impacted by legacy impacts such as channelization. Protection of these uses will be implemented through the application of tiered biological criteria and for some pollutants, tiered chemical criteria. As a result, biological criteria are needed to set minimum goals for each of these tiered uses so that nonattainment or attainment can be determined for management of these waters. The process for developing tiered biocriteria for Minnesota streams is described in this document.

2.3. Indices of biological integrity and biocriteria in Minnesota

The Minnesota Pollution Control Agency (MPCA) has been collecting biological data to determine the condition of waters in Minnesota since the establishment of the MPCA in the 1960s so there is a long history of using biological communities to monitor the condition of waters in the state. This experience has been important for developing a robust biological assessment program. Since the 1990s the MPCA has routinely monitored two biological communities in streams for the purpose of biological assessment: fish and macroinvertebrates. These two groups were selected as a result of the long history and knowledge with using these assemblages in Minnesota and by other states and tribes to measure biological condition. The use of two assemblages is preferred because each group may respond to different forms of stress (USEPA 2013). Therefore, an assessment program that uses two assemblages provides a more comprehensive evaluation of biological condition and is less likely to miss impairment when it actually exists.

To translate biological data into a form that can be used to determine attainment of aquatic life use goals in assessments, the MPCA uses indices of biological integrity or IBIs to measure biological condition. IBIs are the most common analytical tools in the United States used to measure the condition of aquatic communities. The formal development of IBIs in Minnesota began in the 1990s. During this period, the biomonitoring program was expanded and the collection of more data allowed development of watershed specific IBIs in the 1990s and early 2000s (e.g., Bailey et al. 1993, Niemela et al. 1999, Niemela & Feist 2000, Niemela & Feist 2002, Chirhart 2003, Genet & Chirart 2004). Using these watershed IBIs, numeric translators for the narrative criteria (Minn. R. ch. 7050.0150 subp. 6) were used to assess conditions of biological communities. Biocriteria were developed using two different methods. For the Red River basin, the IBI was divided into five 20-point intervals that corresponded to condition classes. The threshold between fair and poor (40) was used to assess attainment of aquatic life. For the St. Croix and Upper Mississippi basins, a reference condition approach was used to develop biocriteria for the protection of aquatic life use goals.

The biological data collected in the 1990s and the subsequent implementation of the intensive watershed monitoring (IWM) framework resulted in a dataset sufficient to revise and improve Minnesota's IBIs (see MPCA 2014b, a). This dataset included biological samples from a range of ecotypes, thermal regimes, stream sizes, stream gradients, and disturbance across Minnesota. Sampling was limited to perennial streams and streams that were wetted for a sufficient time period to permit rapid recolonization. Therefore, the resulting IBI models are only applicable to streams that meet these criteria and should not be used in ephemeral streams without additional testing or model development. The expanded statewide dataset allowed the MPCA to further refine the IBI stream classification framework by identifying natural differences in biological communities related to regional variation and physical stream features that improved the ability to detect anthropogenic disturbance. To develop a framework for the IBI models, a stream typology was created using several cluster analysis techniques to identify groups of sites with similar fish and macroinvertebrate communities and then to associate these groups with stream size, gradient, thermal regime, site habitat conditions (e.g., presence of riffle), and longitude/latitude (MPCA 2014a, b). This analysis resulted in nine distinct stream types for fish and nine similar, but not identical stream types for macroinvertebrates. The differences between the fish and macroinvertebrate frameworks were the result of variation between each assemblage's responses to environmental factors. For example, fish distributions may be affected by landscape features such as major waterfalls, but such features will not influence macroinvertebrate communities (MPCA 2014a). An IBI model was developed for each stream type (i.e., 18 total IBI models; MPCA 2014a, b) using the approaches described by Whittier et al. (2007). This statewide framework of IBI models accounts for natural differences in biological assemblages related to regional variation and physical stream features to minimize the effects of natural factors and maximize detection of anthropogenic stressors.

Using these new refined IBIs, the MPCA developed class-specific biocriteria based on robust reference datasets to manage Minnesota's aquatic life use goals. This effort resulted in nine different IBIs for each biological assemblage (18 total IBIs; Table 1) which are tailored to different ecological regions and waterbody types in Minnesota. The nine stream classes between fish and macroinvertebrates are not parallel because these communities are influenced by different natural factors across the Minnesota landscape. For example, fish distributions are more affected by barriers and watershed area than invertebrates. These different IBIs were developed such that the effects of natural differences on index scores are minimized while the signals from human-caused stressors are maintained. The ability to isolate the impacts of anthropogenic stressors to biological communities makes these indices effective measures of attainment of Minnesota's aquatic life use goals. A detailed description of these IBIs and their development can be found in MPCA (2014b, a).

Table 1. Fish and macroinvertebrate stream classes (Abbreviations: RR = high gradient, GP = low gradient)

Class #	Class Name	Class #	Class Name
	Fish		**Invertebrates**
1	Southern Rivers	1	Northern Forest Rivers
2	Southern Streams	2	Prairie Forest Rivers
3	Southern Headwaters	3	Northern Forest Streams RR
4	Northern Rivers	4	Northern Forest Streams GP
5	Northern Streams	5	Southern Streams RR
6	Northern Headwaters	6	Southern Forest Streams GP
7	Low Gradient Streams	7	Prairie Streams GP
10	Southern Coldwater	8	Northern Coldwater
11	Northern Coldwater	9	Southern Coldwater

3. Development of tiered biocriteria in other states

The approaches used to develop tiered biocriteria in other states have helped to inform Minnesota's process. Most states use biological communities to some degree to determine attainment of aquatic life goals, but few states have TALUs formally adopted into rule (although several states are in the process of developing these tools). The exceptions to this are Maine, Ohio, and Vermont which have formally adopted TALUs and biocriteria into rule. These states each used different methods to develop biocriteria, but all three states used the BCG or a form of the BCG as part of their biocriteria setting process. Two states, Ohio and Vermont, use the Reference Condition to set the biocriteria with the BCG, or a BCG-like tool, used as a check on the Reference Condition. When necessary, the BCG is used to modify the Reference Condition methods. These two states use IBIs as their assessment tool which makes them more similar to Minnesota in this regard. Maine uses the BCG to empirically develop biocriteria, but Maine's methods are less applicable to Minnesota because they use a probability-based multivariate analysis (i.e., Discriminant Function Analysis) rather than IBIs for assessment. More detailed descriptions of the biocriteria for these three states are provided below:

Maine: The biocriteria developed by Maine are rooted in the BCG although Maine's BCG was developed using a different approach than Minnesota's BCG. Using Maine's BCG, sites are placed into different aquatic life use tiers. In Maine's case, sites that meet BCG Levels 3/4 are considered benchmarks for their streams which represent attainment of the CWA interim goal (Class C). Maine also has two use tiers that exceed the CWA interim goal. Class B is equivalent to BCG Levels 2/3 and Class AA/A is consistent with BCG Levels 1/2. These sets of sites are then used as the "reference" set which includes several different levels of condition ranging from natural to the CWA minimum. This is different from the usual use of the term "reference" as there are several different levels of reference sites which correspond to Maine's TALUs. A linear discriminant model which uses a large number of biological metrics is run to determine the probability of a test site belonging to the different tiers. The probabilities are then used to determine if the site is in compliance with WQS.

Ohio: Ohio uses a Reference Condition approach which is informed by a BCG-like framework to develop their biocriteria (Ohio EPA 1987, 1989, Yoder & Rankin 1995). A 25th percentile of the reference sites for a given stream class is used to set biocriteria. The BCG-like framework was part of biocriteria development and helped ensure that the biocriteria developed from the reference sites were above the interim CWA minimum goal. Essentially this tool was used to gage reference condition on a gradient of naturalness to ensure that protective criteria were developed. As a result, if a threshold developed using the 25th percentile was low due to overall poor conditions in a given region, some modification to this percentile was made. For example, the Huron-Erie Lake Plain (HELP) ecoregion uses the 90th percentile of all sites to set biocriteria because the reference sites in this ecoregion fell below the CWA interim goal as defined by the BCG. Biocriteria for Exceptional Use waters is calculated as the 75th percentile of all reference sites across the state. A separate set of modified reference sites is used to set the biocriteria for the Modified Warmwater Use. The Modified Use reference sites met similar criteria to the General Use reference sites with the exception that the habitat is modified through channelization.

Vermont: Vermont uses a fish IBI and a macroinvertebrate multimetric index as numeric biocriteria developed from Regional Reference conditions. Guidelines have been developed to determine water quality standards attainment using both the fish community IBI, and the macroinvertebrate community metrics. A percentile approach was used to set thresholds for attainment across tiered use classes (Vermont Depatment of Environmental Conservation 2004).

4. Minnesota's approach to developing biological criteria

The biocriteria for Minnesota streams is based on data collected over a 15 year period (1996-2010) from more than 2,800 sampling sites. The dataset includes not only biological data (i.e., fish and macroinvertebrates), but chemical, physical, and land use data that were integral to developing protective goals for Minnesota streams. Experience from other states also provides a conceptual approach to developing biocriteria in Minnesota although the final biocriteria are tailored specifically to Minnesota's resources and goals.

For all three TALU aquatic life use class tiers (i.e., Exceptional, General, and Modified), a multiple lines of evidence approach was used to develop protective and attainable biocriteria. Two lines of evidence were most important: the BCG and the Reference Condition. The Reference Condition is the traditional approach used to identify biological thresholds. It includes the well accepted method of using an independent, *a priori* non-biological measure to select reference sites (e.g., an index of human activity in a watershed) which represent attainment of aquatic life use goals. The biological communities from these reference sites are then used to set goals for streams with an unknown condition. The BCG, on the other hand attempts to describe how biological communities change along a gradient of disturbance. The BCG approach relies on our fundamental understanding of fish and macroinvertebrate life history requirements and how disturbances from humans are known to impact their physiological and community level functions (e.g. spawning, reproductive success, feeding, etc.). The BCG is based on the ecological theory that water bodies with higher levels of effective anthropogenic stress have biological communities with lower condition compared to water bodies with less effective stress (Davies & Jackson 2006). Development of BCG models provides a common framework to interpret changes in biological condition regardless of geography or water resource type. More detailed descriptions of the BCG can be found in EPA (2005) and Davies and Jackson (2006).

In the process of assessing each approach, it was determined that the Reference Condition approach by itself was problematic for some regions of Minnesota because of the degree to which these regions had already been impacted. Specifically, southern streams had few sites that could be considered "least disturbed". In other states, biocriteria in these heavily impacted regions were based on a higher percentile of the reference sites or alternatively an 'all sites' approach was used (e.g., Ohio used the 90^{th} percentile of all sites for one ecoregion). Minnesota chose not use this approach, considering it inappropriate to make an *a priori* decision that some known proportion of streams is impaired. Instead, the BCG was relied on more heavily for these classes to establish biocriteria. While there is still a need to choose an impairment threshold along the BCG the decision is informed by aligning known ecological endpoints (i.e. BCG levels) with Minnesota's aquatic life use goal narratives. To do this, classes with a sufficiently large reference site sample size (i.e., northern and statewide classes) were used to determine the relationship between the Reference Condition and BCG level threshold could be applied to the other classes to determine thresholds. Finally the draft biocriteria for all stream classes were based on statistics derived from the BCG to ensure consistency for goals across stream classes and across the state. Despite limitations of the Reference Condition for some classes, these two approaches largely identified similar thresholds which provided better confidence in the final biocriteria.

5. Datasets used to develop biocriteria

The macroinvertebrate and fish data used to develop biocriteria were the result of extensive surveys in Minnesota from 1996 through 2010. The field sampling protocols for collection of biological data can be found in MPCA reports (MPCA 2002, 2004, 2009). Different datasets were used to develop biocriteria for each TALU tier. The analyses for the General and Exceptional uses included only sites from reaches that were considered to have natural channels (i.e., <50% channelized) as determined by a site visit and aerial photography. The Modified Use analyses included sites from both natural and channelized stream reaches. Some additional screening was performed to remove anomalous samples or sites. Sites that were close to lakes or large rivers were not included due to the possible influence of these water bodies on the biological communities. In addition, samples that were collected during periods of high or low flows were not included in these datasets. Datasets included all samples that met the above criteria which in some cases resulted in multiple samples from a small subset of sites. These additional samples were included to increase samples sizes. Sample sizes and disturbance as measured by the HDS varied between stream classes (Figures 1 and 2). The large river classes and coldwater classes had fewer samples which is a reflection of the relative abundance of these habitats in Minnesota. Northern classes had more sites with less disturbance (i.e., higher HDS scores) whereas southern class were more disturbed with only a small proportion of sites scoring higher than an HDS of 60.

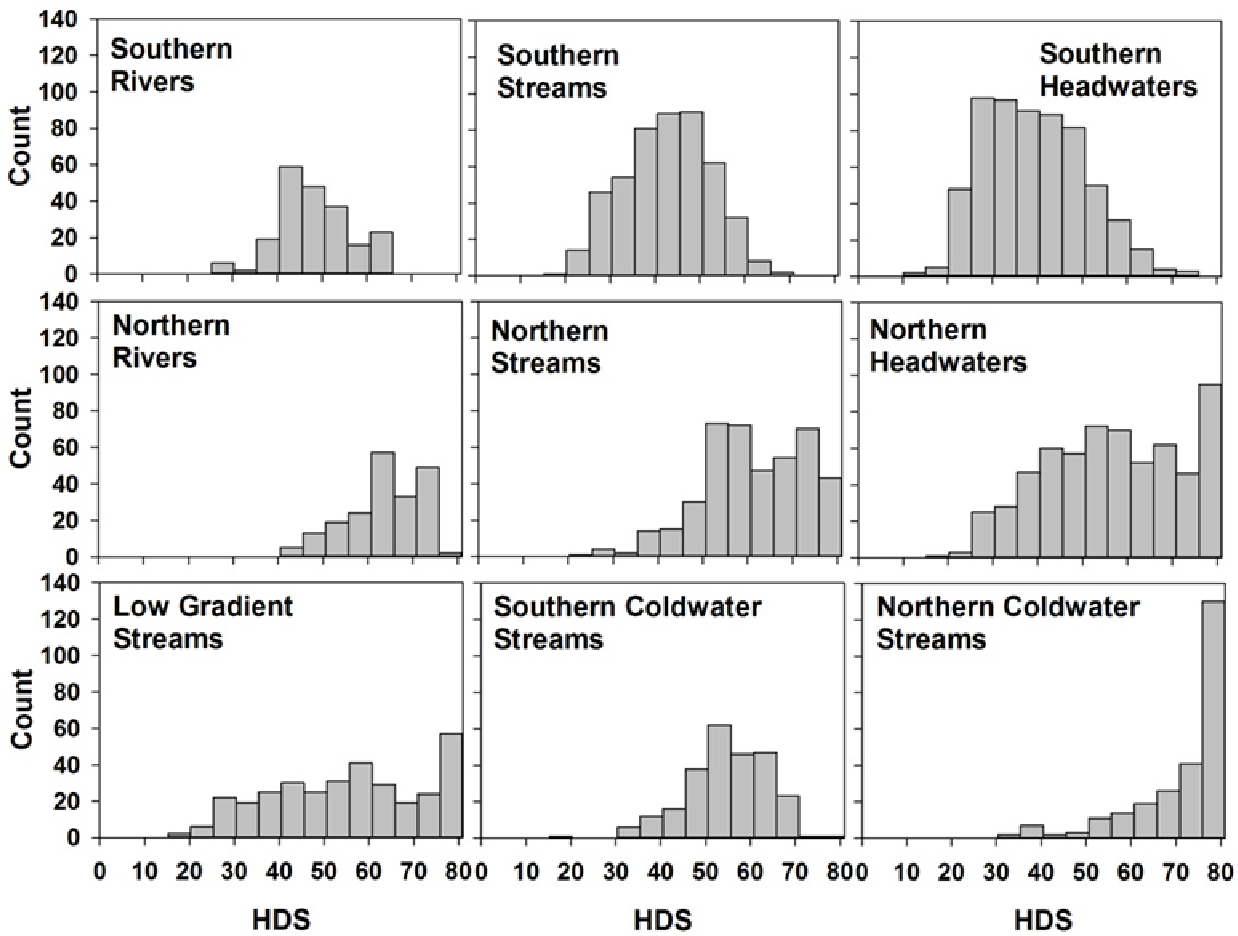

Figure 1. Histograms of Human Disturbance Scores for fish classes.

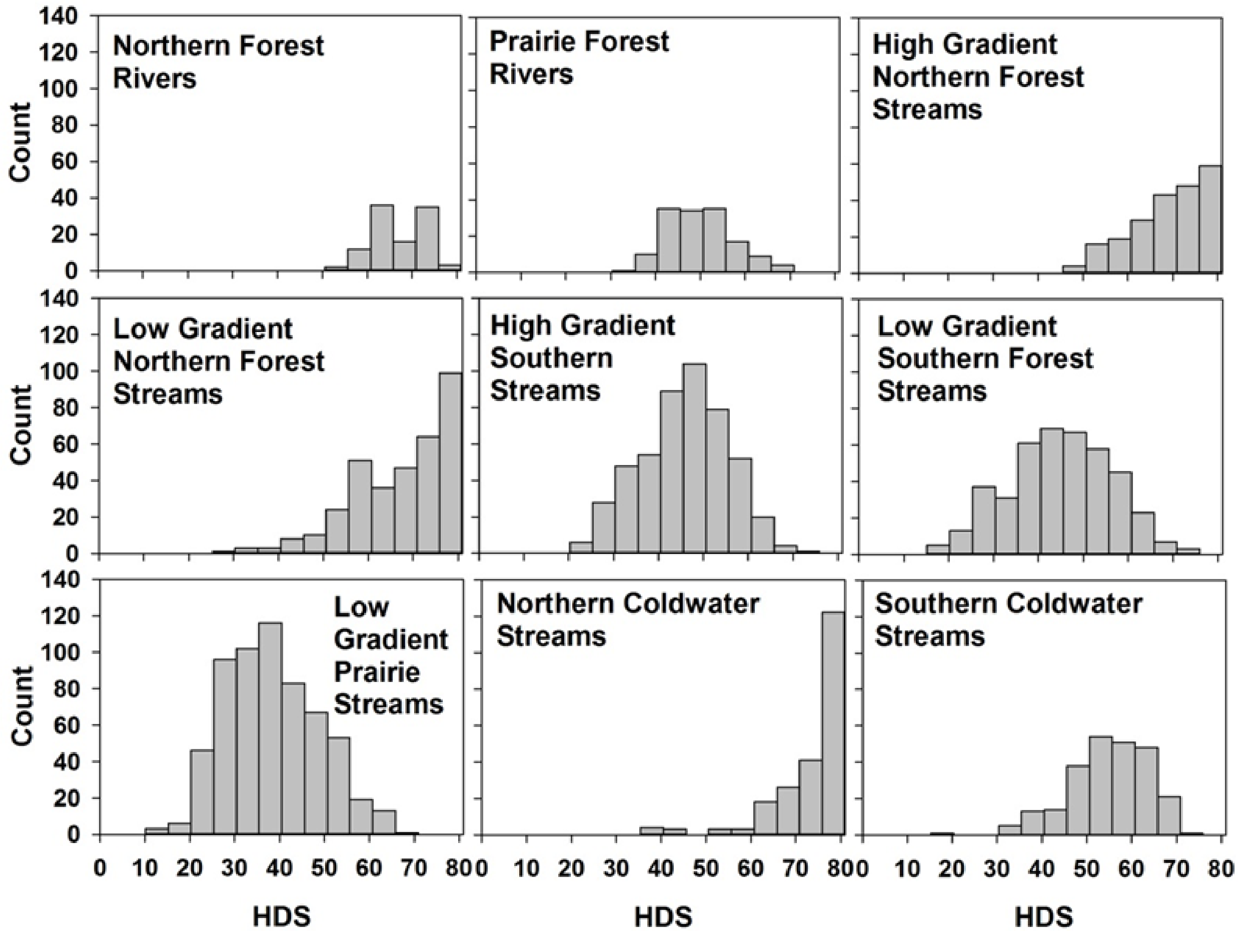

Figure 2. Histograms of Human Disturbance Scores for macroinvertebrate classes.

6. Development of Minnesota's reference condition

There are many approaches that have been proposed and employed to determine attainable conditions that support aquatic life use goals (Hughes et al. 1986). However, the development of statewide goals in Minnesota limits these approaches to those that are effective for a state with diverse aquatic resources and for both point and nonpoint sources of pollution. As a result the most effective approach that Minnesota can use is the regional reference site approach (see Hughes et al. 1986). The regional reference site approach is used to develop biocriteria by states with biocriteria in water quality standards (e.g., Ohio, Vermont). This approach involves the selection of reference or benchmark sites from homogenous regions and waterbody types that approximate biological integrity and therefore represent attainment of aquatic life use goals for those classes of water bodies (Hughes et al. 1986, Gibson et al. 1996). IBI scores are then calculated for the reference sites and a percentile of IBI scores for each set of reference sites is chosen to represent the true reference condition. Most commonly, the 25th or 10th percentiles of IBI scores are used to address uncertainty regarding relative impacts to lower scoring sites. The elimination of the lower quartile or decile removes the effect of outliers and provides a degree of safety as the reference site selection process is imperfect and likely includes some sites that are not truly of reference quality. The decision of which percentile to apply is based on the overall condition of the class; where the 10th percentile of reference site IBI scores is appropriate in a class with many “minimally disturbed” sites (see Stoddard et al. 2006). In contrast a waterbody class with only “least disturbed” sites will require the use of the 25th percentile (Figure 3). The use of different percentiles is determined by how confident you are that the reference site population represents attainment of aquatic life use goals. Regardless of the statistic used, the resulting value represents the threshold or biocriteria which is used to determine if sites are considered to be in attainment of aquatic life use goals. It also addresses attainability issues by incorporating the majority of what have been defined as reference and eliminating the circularity of alternate and *post priori* approaches.

The most important step of the Reference Condition approach is establishing or defining the Reference Condition. The approach described above is a brief overview where a sufficient number of sites that represent the attainability and attainment of aquatic life use goals can be identified. In heavily disturbed areas or regions, it can be difficult to find sites that represent attainment of biological goals or protection of biological integrity (Gibson et al. 1996). As a result, an alternative or modified approach is needed to preclude setting a biocriterion too low resulting in an underestimation of potential aquatic life use goals. If a stream class has overall poor condition (i.e., poorer than least disturbed), then thresholds developed for that class are likely to be under protective (Figure 3). There are a number of modifications or methods that can be used to modify the biocriteria to different stream classes so that they are not under or over protective. In cases where reference sites are defined as “best available”, even the 25th percentile may still result in under protective biocriteria (Figure 3). This scenario requires more creative approaches such as using the 90th percentile of all sites as in the HELP ecoregion in Ohio (Ohio EPA 1987, 1989). In such a case, additional information is needed to support a method that differs from the standard approach. In the case of Ohio, a BCG-like tool was used to develop biocriteria differently for the HELP ecoregion.

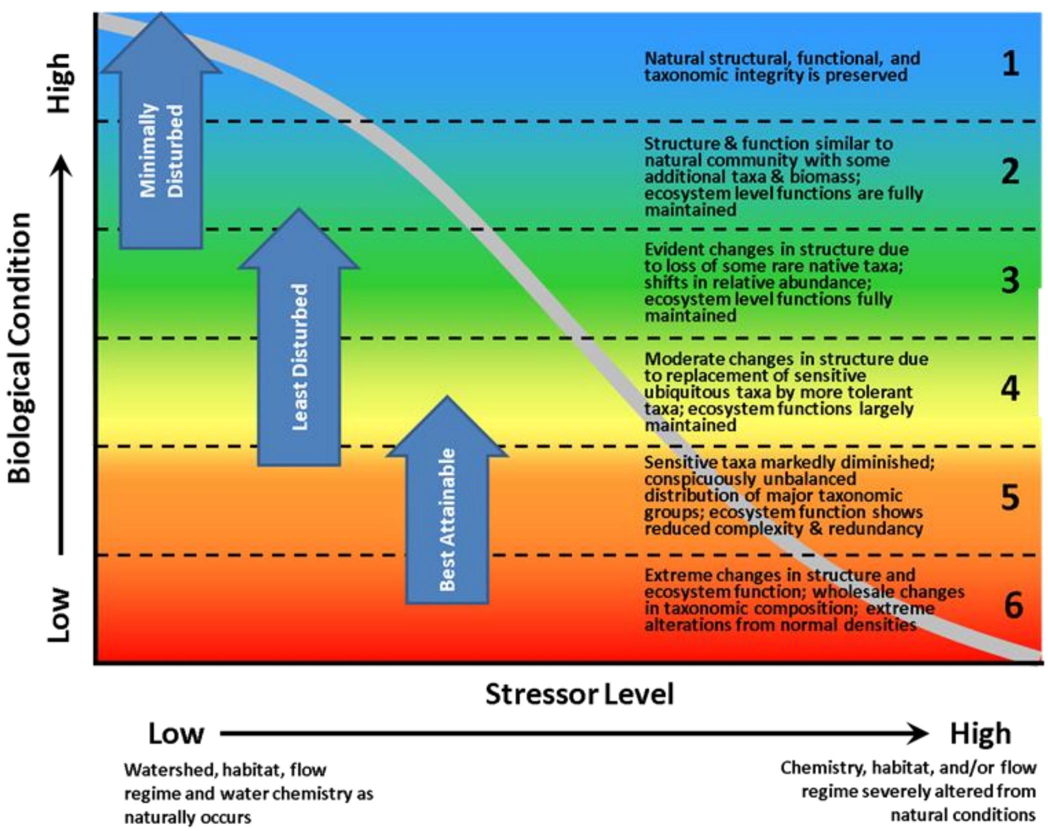

Figure 3. The distribution of minimally disturbed, least disturbed, and best attainable Reference Condition along the axis of biological condition against the level of stress (adapted from Stoddard et al. [2006]). Minimally disturbed, least disturbed, and best attainable are shown as they relate to their position in the Biological Condition Gradient (BCG).

Central to developing a Reference Condition is the ability to select stream sites that are least or minimally disturbed using an *a priori* measure of condition that is independent of the biology. Generally these models are not based on water quality or biological parameters, but rather employ land use and other measures of human activity in a watershed or stream reach. The MPCA has developed the Human Disturbance Score (HDS) (MPCA 2014c), an index to measure the degree of human activity upstream of and within a stream monitoring reach. The HDS includes both watershed and reach level measures of human disturbance which receive a score of 0-10 (Table 2). Additional adjustments are made for watershed and reach-level factors which can negatively impact waterbody condition. These metrics and adjustments together have a maximum score of 81 (Table 2). Minnesota stream reference sites were identified as those with an HDS score of 61 or greater (i.e., the upper 25% of the HDS distribution). Once sites were selected based on their HDS score, several additional filters were applied to remove sites disparately influenced by nearby stressors. All sites in close proximity to urban areas (site within or adjacent to urban area), feedlots (feedlot at or immediately upstream of site [only streams >50 mi^2]), or point sources (continuous point source <5 mi upstream of site) were removed. Sites meeting these criteria and receiving an HDS score of 61 or greater were consistent with other criteria for Reference Condition sites including low human population density, low agricultural activity, and no nearby NPDES discharges (Gibson et al. 1996). Sites meeting these criteria were considered to be minimally or least disturbed and therefore potentially representative of attainment of Minnesota's aquatic life use goals.

Table 2. Metrics and scoring for Minnesota's Human Disturbance Score

Human Disturbance Score Metric	Scale	Primary Metric or Adjustment	Maximum Score
Number of animal units per km^2	watershed	primary	10
Percent agricultural land use	watershed	primary	10
Number of point sources per km^2	watershed	primary	10
Percent impervious surface	watershed	primary	10
Percent channelized stream per stream km	watershed	primary	10
Degree channelized at site	reach	primary	10
Percent disturbed riparian habitat	watershed	primary	10
Condition of riparian zone	reach	primary	10
Number of feedlots per km^2	watershed	adjustment	-1
Percent agricultural land use on >3% slope	watershed	adjustment	-1
Number of road crossings per km^2	watershed	adjustment	-1 or +1
Percent agricultural land use in 100m buffer	watershed	adjustment	-1
Feedlot adjacent to site	reach (proximity)	adjustment	-1
Point source adjacent to site	reach (proximity)	adjustment	-1
Urban land use adjacent to site	reach (proximity)	adjustment	-1
		Maximum	**81**

A comparison of HDS metric values for natural channel reference and non-reference sites showed good separation between these stream sites for most metrics (Figure 4). These figures also provide a visualization of the relatively low levels of disturbance at the references sites and their upstream watershed for these measures. There is no difference between reference and non-reference sites for the degree of channelization at the site because the sites included in this analysis have natural channels so no difference would be expected.

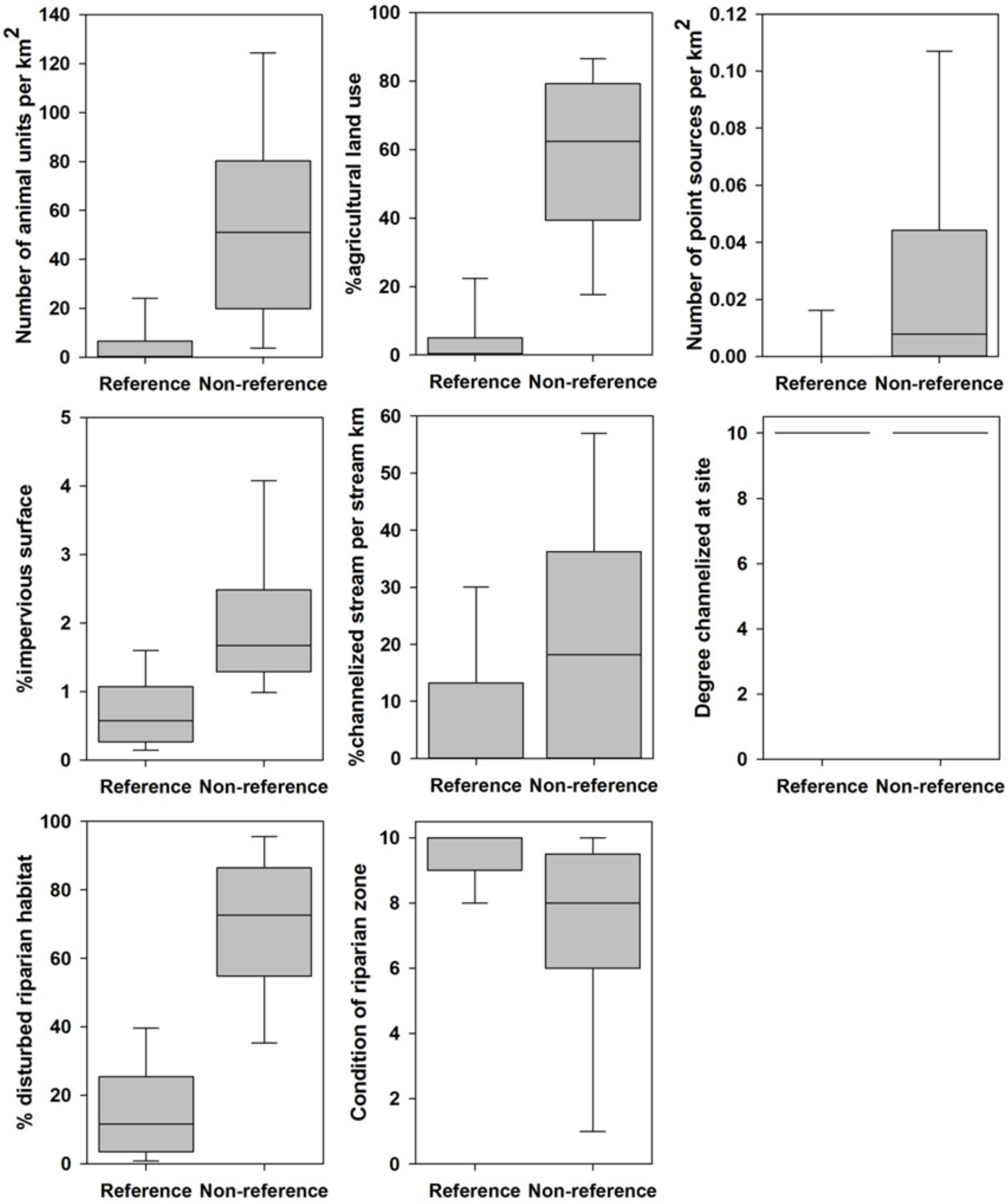

Figure 4. A comparison of Human Disturbance metric values for natural channel reference and non-reference sites. The degree of channelization is the proportion of reach that has a natural channel in 10% intervals (e.g., a score of 10 = 100% natural channel). Condition of the riparian zone is the average of % undisturbed from 0-30 m and 0-15 m buffers. Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 90th and 10th percentiles.

There was a distinct difference in the size of reference (i.e., minimally and least disturbed) datasets between stream classes with few reference sites present from the plains or southern stream classes (Table 3, Figure 5). Specifically, the southern warmwater classes for both assemblages had 25 or fewer sites. The southern coldwater classes had fewer sites than the northern coldwater classes, but both fish and macroinvertebrates had more than 50 sites. Northern and statewide classes had more than 100 sites with the exception of one northern class with 83 sites. The low number of sampled reference sites in the southern classes poses some problems for biocriteria development because small sample sizes can result in greater uncertainty in the statistics (e.g., 10th or 25th percentile) used to determine Reference Condition thresholds. In addition, the IBI scores from southern reference sites were lower than their northern counterparts (Figure 6) which reflects the overall poorer condition of streams in the plains ecoregions.

Table 3. Numbers of reference sites for fish and macroinvertebrate stream classes (Abbreviations: RR = high gradient, GP = low gradient)

Class #	Class Name	Reference
1	Southern Rivers	18
2	Southern Streams	8
3	Southern Headwaters	15
4	Northern Rivers	116
5	Northern Streams	186
6	Northern Headwaters	215
7	Low Gradient Streams	111
10	Southern Coldwater	61
11	Northern Coldwater	196
1	Northern Forest Rivers	83
2	Prairie Forest Rivers	9
3	Northern Forest Streams RR	162
4	Northern Forest Streams GP	210
5	Southern Streams RR	15
6	Southern Forest Streams GP	25
7	Prairie Streams GP	13
8	Northern Coldwater	185
9	Southern Coldwater	60

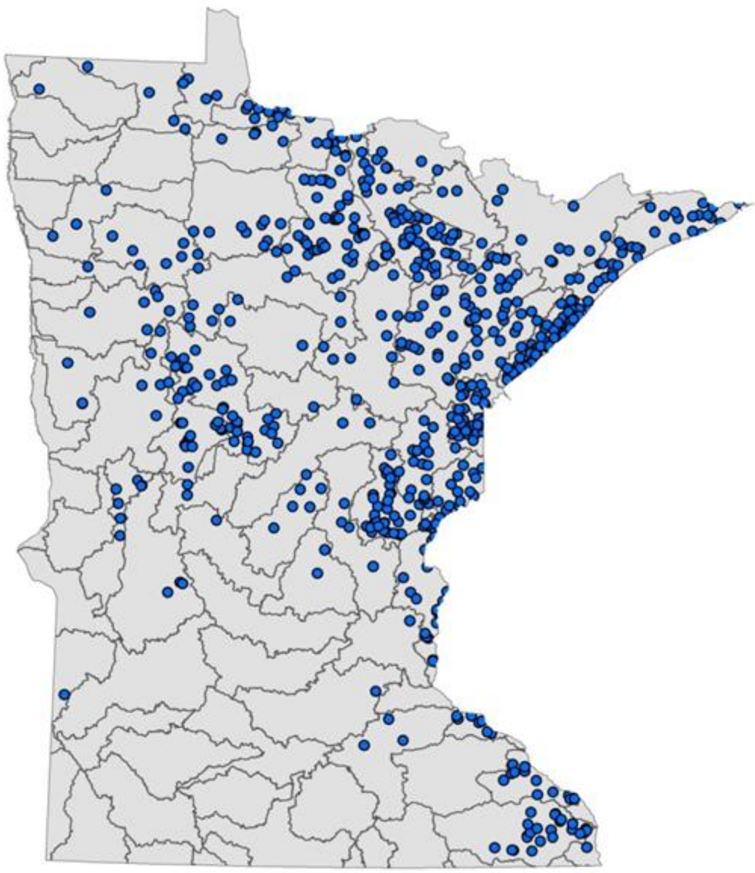

Figure 5. Map of minimally-disturbed and least-disturbed sites based on Human Disturbance Score criteria.

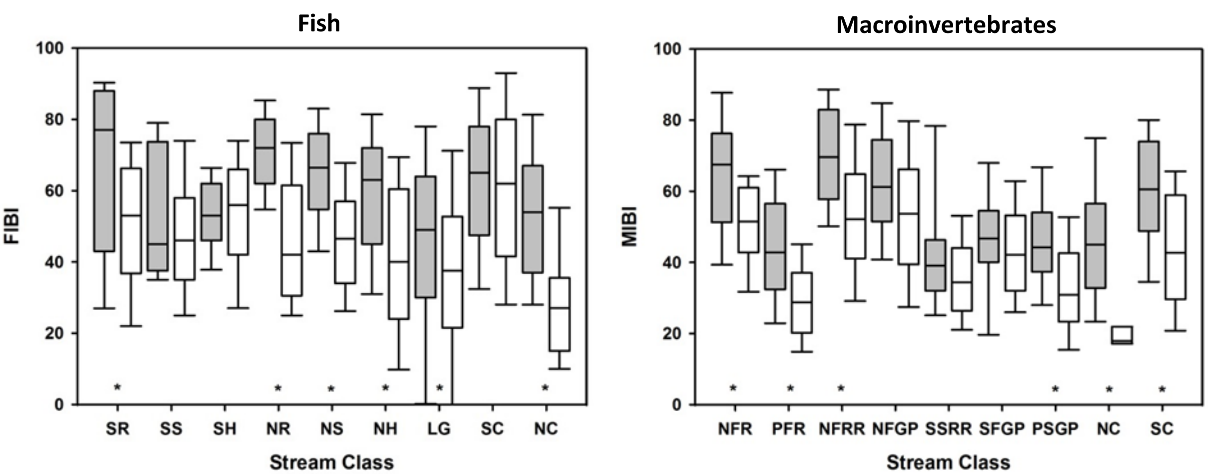

Figure 6. A comparison of reference (gray box plots) and non-reference (white box plots) site IBI scores for fish and macroinvertebrates from natural channel sites. Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 90th and 10th percentiles; * indicates significant difference at the α=0.05 level based on a Mann-Whitney rank sum test; Abbreviations: SR = Southern Rivers, SS = Southern Streams, SH = Southern Headwaters, NR = Northern Rivers, NS = Northern Streams, NH = Northern Headwaters, LG = Low Gradient Streams, SC = Southern Coldwater, NC = Northern Coldwater, NFR = Northern Forest Rivers, PFR = Prairie Forest Rivers, NFRR = High Gradient Northern Forest Streams, NFGP = Low Gradient Northern Forest Streams, SSRR = High Gradient Southern Streams, SFGP = Low Gradient Southern Forest Streams, PSGP = Low Gradient Prairie Streams.

Comparing IBI scores for reference and non-reference sites provides a way to assess if the reference sites selection process was effective. Box plots of IBI scores for reference sites versus non-reference were generated and these scores were compared for each class using a Mann-Whitney Rank Sum Test in SigmaPlot ver. 12 (Systat Software 2011) because most datasets were not normal. A comparison of the 25th percentiles for reference and non-reference sites indicated that the reference sites score higher and fall above the first or second quartile of the non-reference sites (Figure 6). In general, the 25th percentile of the reference sites for the northern classes falls above the second quartile of the non-reference sites whereas the southern classes fall between the first and second quartiles of the non-reference sites (Figure 6). The reference and non-reference IBI scores were not significantly different (α=0.05) for three fish and three macroinvertebrate. Five of these six classes were southern classes. This difference between regions is likely a reflection of the greater disturbance for the southern reference sites. This is apparent when the HDS values for reference sites are compared between northern and southern stream classes (Figure 7).

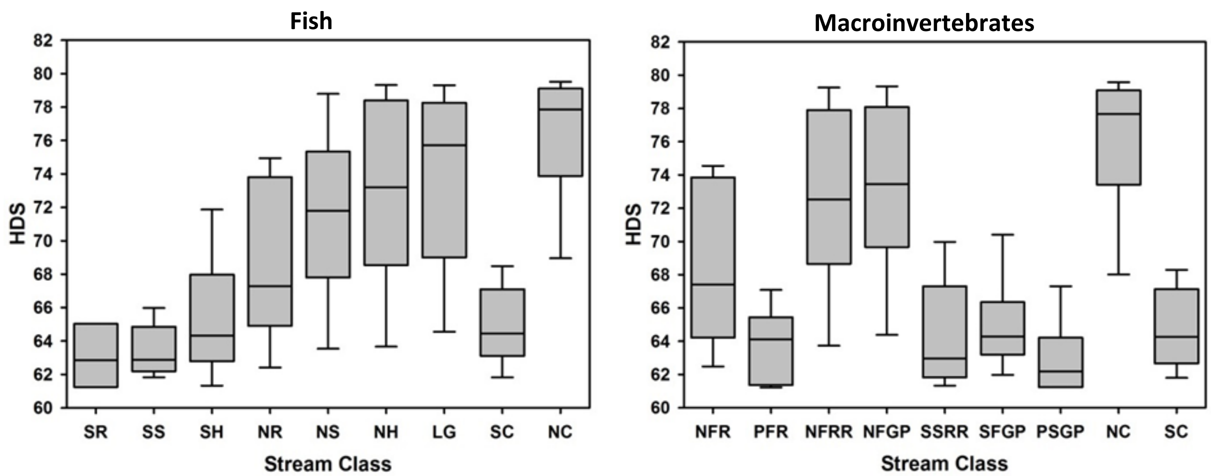

Figure 7. Human Disturbance Score distributions for reference sites from fish and macroinvertebrate classes. Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 90th and 10th percentiles; Abbreviations: SR = Southern Rivers, SS = Southern Streams, SH = Southern Headwaters, NR = Northern Rivers, NS = Northern Streams, NH = Northern Headwaters, LG = Low Gradient Streams, SC = Southern Coldwater, NC = Northern Coldwater, NFR = Northern Forest Rivers, PFR = Prairie Forest Rivers, NFRR = High Gradient Northern Forest Streams, NFGP = Low Gradient Northern Forest Streams, SSRR = High Gradient Southern Streams, SFGP = Low Gradient Southern Forest Streams, PSGP = Low Gradient Prairie Streams.

The interquartile ranges of IBIs for reference and non-reference sites overlap for most classes regardless of region although this overlap is less for northern classes (Figure 6). This is a result of the fact that metrics in the HDS index are good at measuring human activity and therefore anthropogenic stressors, but do an inadequate job of quantifying the extent that the impact of the activity has been reduced by management practices. As a result, the HDS alone is not a good measure of stress that can be used for assessment purposes. For example, the percent of agricultural land use does not take into account conservation measures or the intactness of riparian habitat that can mitigate the impacts of agriculture on streams. In addition, the HDS index is not sensitive to some broad-scale stressors such as connectivity which can negatively impact biota. As a result, some sites identified as non-reference score well and attain the beneficial use. These high scoring non-reference sites do provide some insight into the attainability of biological goals when appropriate restoration and conservation practices are employed.

Overall, the reference site selection process was effective in identifying higher performing sites although the reduced separation between reference and non-reference and the small sample sizes in the southern stream classes raised some concerns about the applicability of the reference condition approach in southern Minnesota. However, in the northern regions and in coldwater classes sufficient and effective reference datasets were developed which can be used to support development of biological criteria.

7. Minnesota's biological condition gradient

The BCG is a conceptual model of aggregated biological knowledge used to describe changes in biological communities along a gradient of increasing stress and is based on a combination of ecological theory and empirical knowledge. A number of indices have been developed to measure the biological condition in aquatic systems (e.g., IBI, RIVPACS; Karr et al. 1986, Hawkins et al. 2000, Whittier et al. 2007), but these measures are based on the available conditions that are used to develop the models. This can result in under-protective criteria for stream classes with overall poor condition. The BCG differs from these models in that it provides a common “yardstick” of biological condition that is rooted in the natural condition whether or not it presently exists. As a result, the BCG can be used to develop biocriteria that are consistent across regions and stream types in Minnesota. This is particularly important for a state such as Minnesota where the range of conditions is regionally distinct and extreme (i.e., relatively pristine to degraded). The BCG divides biological condition into six levels that are intended to provide a stepwise explanation about how a biological assemblage responds to a gradient of increased stressor effects (Figure 8). The BCG has been proven to be a valuable tool for those states that are in the process of developing biological criteria (USEPA 2011) and some of the states that have adopted or are in the process of adopting TALUs have developed BCGs or analogous models. More detailed descriptions of the BCG can be found in USEPA (2005) and Davies and Jackson (2006).

The development of Minnesota's warmwater BCG involved input from biological experts familiar with biological assemblages in Minnesota streams from the MPCA and Minnesota Department of Natural Resources (MDNR). BCG models were developed for fish and macroinvertebrates for each of the seven warmwater stream classes. BCG models were also developed for the coldwater stream classes which involved participation by experts from Minnesota, Wisconsin, Michigan, Fond du Lac Band of Lake Superior Chippewa, Oneida Nation, Little River Band of Ottawa Indians, and Red Lake Band of Chippewa. In Minnesota this included two classes each for fish and macroinvertebrates. A detailed description of how the BCGs were developed for Minnesota can be found in Gerritsen et al. (2013).

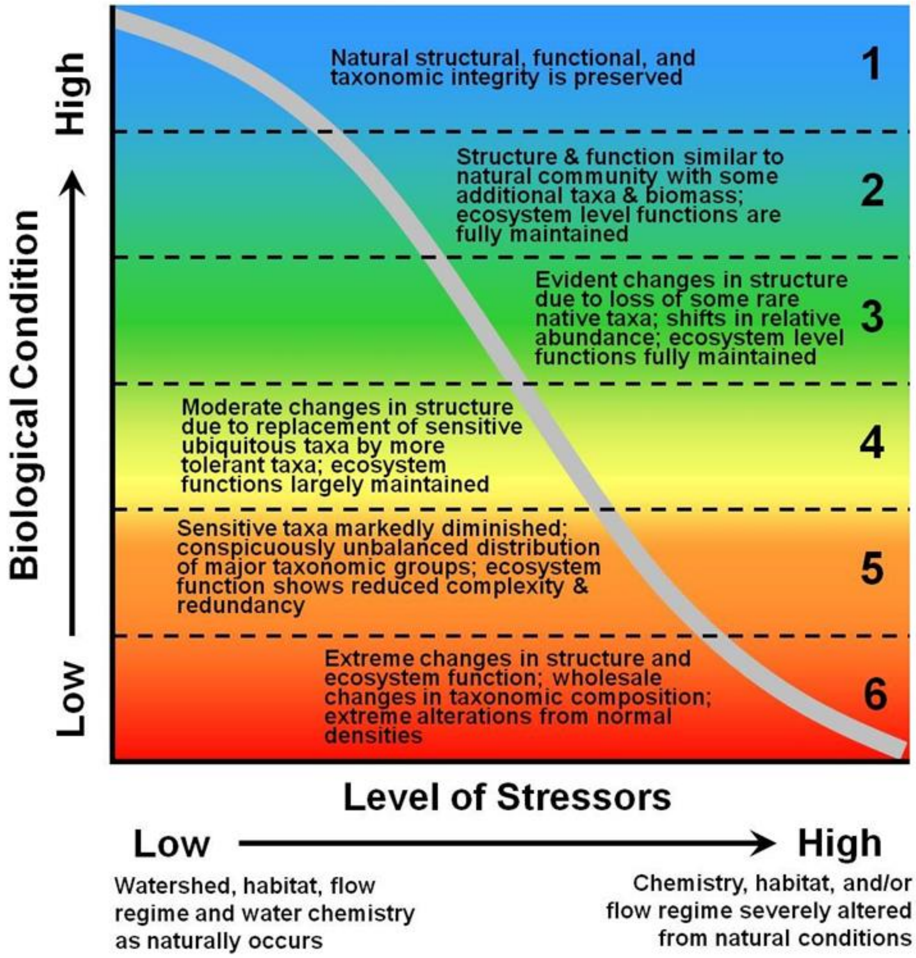

Figure 8. Conceptual model of the biological condition gradient (modified from Davies and Jackson [2006]).

BCG models were developed for all 18 fish and macroinvertebrate stream classes. However, not all six BCG levels were represented in the empirical datasets for all of the stream classes. Specifically, examples of BCG Level 1 were generally not identified for southern stream classes. In a number of the southern macroinvertebrate classes, BCG Level 2 streams were also absent. On the other end of the BCG scale, a number of northern classes lacked samples that corresponded to BCG Levels 5 or 6. The truncated BCG gradient for macroinvertebrate classes is in part a reflection of better historical knowledge for fish, differences in the geographic boundaries for stream classes, and different sensitivities to stressors between these groups. Some BCG levels did not fit the expected IBI-BCG pattern (e.g., Fish Northern Rivers BCG Level 6; Macroinvertebrate High Gradient Northern Forest Streams BCG Level 5). These anomalous levels were the result of small sample sizes and were generally not indicative of a deficiency in the models. A description of more detailed analysis of the performance of the Minnesota BCGs can be found in Gerritsen et al. (2013).

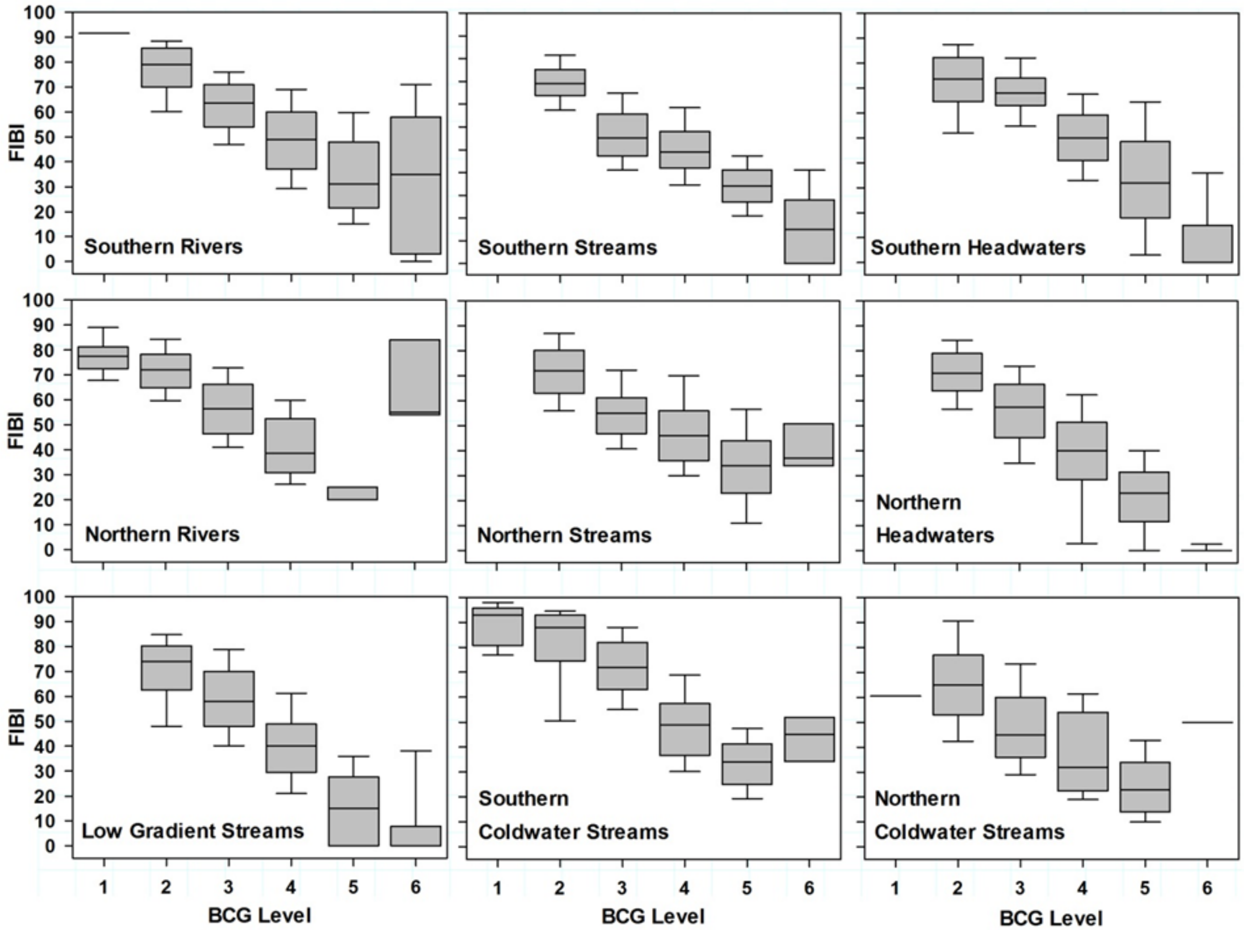

Figure 9. Frequency distributions of IBI scores by BCG level for fish class sites from natural channel streams sampled from 1996-2010 . Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 90th and 10th percentiles.

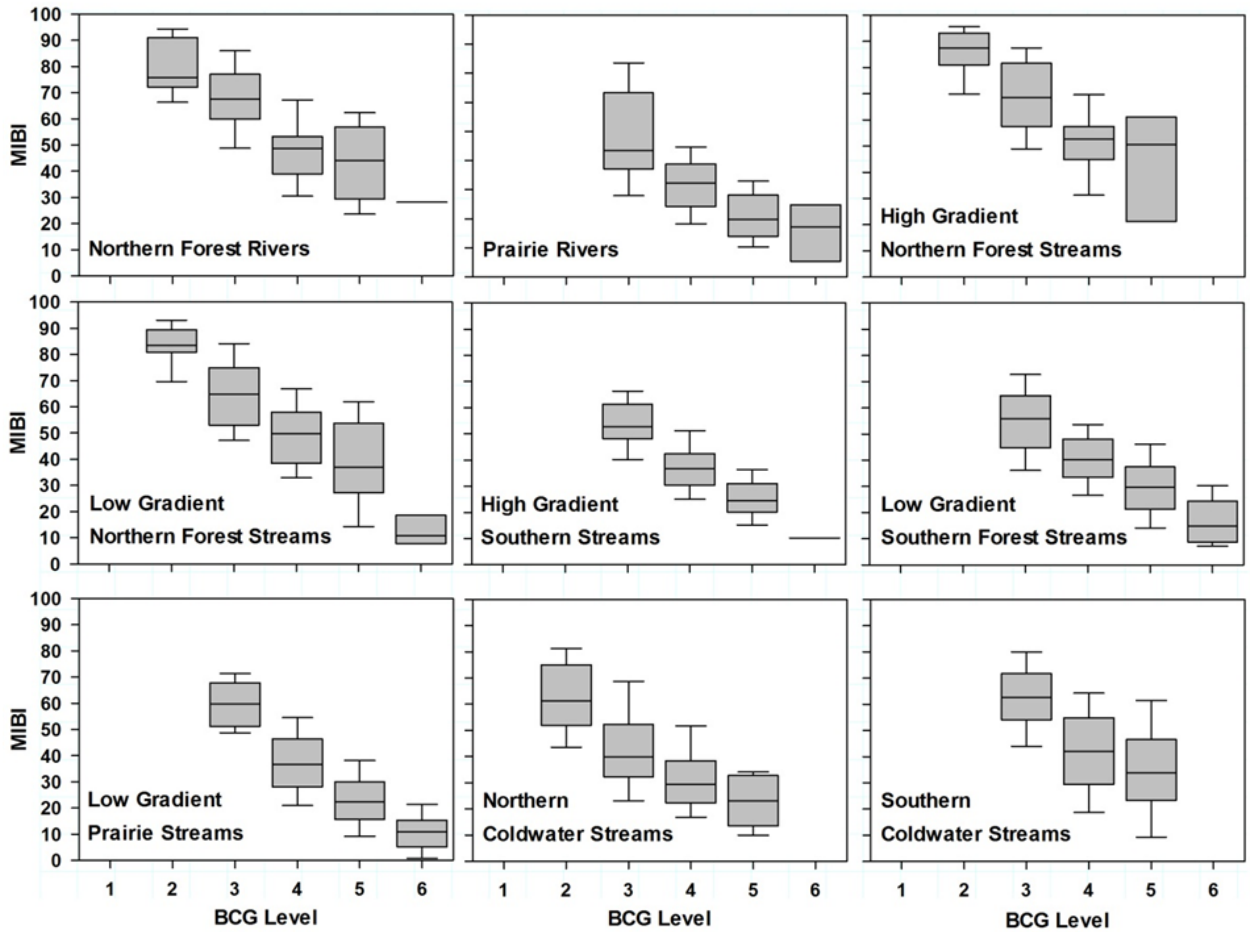

Figure 10. Frequency distributions of IBI scores by BCG level for macroinvertebrate classes using data from natural channel streams sampled from 1996-2010 . Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 90th and 10th percentiles.

The IBI scores and BCG levels for each stream class were compared using box plots of IBI scores for each BCG level. From these figures, it is apparent that IBI scores and BCG levels are correlated (Figures 9 and 10). In general, a downward step pattern is observed with the lower quartile of a BCG level being similar to the upper quartile of the next highest BCG level. For some BCG levels (often BCG Level 6) this pattern does not hold, but this is largely the result of one or more levels containing a small number of samples and the difficulty of a best professional judgment approach alone in discriminating the lower extremes of the BCG.

8. Development of general use biocriteria

Minnesota's General Use applies to waters that support "good" assemblages of freshwater organisms and that reflect the lower distribution of the "least impacted" regional Reference Condition. The General Use represents the minimum threshold for attainment and it serves as the goal for restoration for management programs when nonattainment is determined. As such it is an important trigger for further management actions and considerations such as use attainability analysis and TMDLs.

8.1. The BCG and the general use

The BCG was integral to the development of biocriteria for Minnesota streams. The fact that the BCG is anchored in the natural condition allows it to be used to set consistent biocriteria across a landscape with diverse conditions. However, this first required Minnesota's aquatic life use goals to be mapped to the BCG. Maine uses biological communities described as BCG Levels 3 and 4 to set their minimum biological goals (USEPA 2005). Vermont's biocriteria threshold for streams classified as "good" (Class B2/3) is also associated with BCG Level 4 as is Ohio's Warmwater Habitat use biocriteria (USEPA 2005). During the process of developing the BCG there was general consensus among the biological experts that Minnesota's aquatic life use goals are located in or near Level 4 on the BCG.

The narrative that describes BCG Level 4 is also relevant when considering placement of the Minnesota's aquatic life use goals. At this level, the structure and function of the ecosystem is largely maintained although there may be some moderate changes to species composition. This means that the minimal goal is not "pristine", but rather can reflect some anthropogenically caused changes to the biological assemblage. The narrative language for BCG Level 4 is: "*overall balanced distribution of all expected major groups; ecosystem functions largely maintained through redundant attributes*". In general, the ecosystem function of the community is maintained by redundancy in species composition. For example, some sensitive taxa may be replaced by intermediate or facultative taxa that fulfill similar ecological roles. Minnesota rule states that:

> *"For all Class 2 waters, the normal fishery and lower aquatic biota upon which it is dependent and the use thereof shall not be seriously impaired or endangered, the species composition shall not be altered materially..."* [Minn. R. ch. 7050.0150 subp. 3]

A biological community classified as Level 5 on the BCG has already undergone considerable structural and functional loss which is not consistent with Minnesota rule (BCG Level 5: "*conspicuously unbalanced distribution of major groups…; organism condition shows signs of physiological stress; ecosystem function shows reduced complexity and redundancy*"). Based on this information and the consensus formed by Minnesota biologists and biologists in other states, BCG Level 4 is consistent with attainment of Minnesota's aquatic life use goals.

8.2 Development of candidate general use biocriteria

To select biocriteria thresholds a multiple lines of evidence approach was used. This involved a quantification of numerous candidate IBI thresholds at different BCG levels followed by a comparison of the BCG derived IBI thresholds to the IBI thresholds established using the Reference Condition approach. The BCG and reference site datasets only included sites from reaches that were considered to have natural channels (i.e., <50% channelized) as determined by a site visit and aerial photography. For the Reference Condition, reference sites were selected using the criteria discussed in Section 6. The 25th percentile of IBI scores was then calculated for the reference sites from each of the stream classes to determine candidate biological thresholds. Candidate thresholds were developed for the General Use using several statistics from BCG Levels 4 and 3 to assess empirically the location of this goal on the BCG. To develop BCG thresholds, samples from sites determined to be BCG Level 4 were extracted and the 25th, 50th, and 75th percentiles of IBI scores was determined for these datasets. The 25th percentile of IBI scores for BCG Level 3 was also calculated and used in this analysis.

Sample sizes by class indicated that many more reference sites were present in the northern stream classes compared to the southern classes (Table 4). In fact, six southern classes had fewer than 20 reference sites which were potentially problematic for biocriteria development. Sample sizes were more evenly distributed between classes for BCG Levels 3 and 4. In fact, only a single class (Macroinvertebrate Low Gradient Prairie Streams) had fewer than 20 sites in BCG Level 3. For BCG Level 4, there were six classes with 20-50 reference sites. Five of these classes were northern classes. The remaining 12 BCG Level 4 classes had more than 50 sites. In general, the sample sizes for BCG Levels 3 and 4 were sufficient for biocriteria development (see Sample size sufficiency for developing biocriteria, pg. 44) although some classes were small enough to require additional assessment. For those classes where the samples size for these BCG levels was small, the reference site dataset was generally greater than 100 sites.

Table 4. Number of samples for datasets used to develop General Use candidate biological criteria. BCG and reference datasets include only sites with natural channels (Abbreviations: RR = high gradient, GP = low gradient).

Class #	Class Name	Reference	BCG4	BCG3
Fish				
1	Southern Rivers	18	82	61
2	Southern Streams	8	74	102
3	Southern Headwaters	15	183	49
4	Northern Rivers	116	28	47
5	Northern Streams	186	155	54
6	Northern Headwaters	215	37	127
7	Low Gradient Streams	111	67	52
10	Southern Coldwater	61	43	101
11	Northern Coldwater	196	21	118
Macroinvertebrates				
1	Northern Forest Rivers	83	25	56
2	Prairie Forest Rivers	9	81	28
3	Northern Forest Streams RR	162	39	159
4	Northern Forest Streams GP	210	63	135
5	Southern Streams RR	15	182	57
6	Southern Forest Streams GP	25	69	57
7	Prairie Streams GP	13	78	12
8	Northern Coldwater	185	53	56
9	Southern Coldwater	60	109	73

8.3. Comparison of BCG and reference condition derived thresholds

A comparison between the threshold derived using the Reference Condition and those developed using the different BCG statistics indicated that the median of BCG Level 4 was most similar to the Reference Condition (Table 5, Figure 11). In general, the 25th percentile of BCG Level 4 was much less protective than the Reference Condition with 16 of 18 classes less protective including 10 classes with a difference of 10 points or more. The 75th percentile of BCG Level 4 and the 25th percentile of BCG Level 3 were much more protective than the Reference Condition with 13-16 of the 18 classes more protective including 7-10 classes with differences of more than 10 IBI points (Table 5). In general, the median of BCG Level 4 tended overall to be equivalent to the Reference Condition thresholds. Eight of the 18 stream classes were more protective using the median of BCG Level 4; however, there were only two classes where this difference was more than 10 points. One macroinvertebrate class, the low Gradient Northern Forest class, had no difference between the median of BCG Level 4 and the reference condition. The nine classes where the median of BCG Level 4 was less protective were for the fish classes the Northern Rivers, Northern Streams, Northern Headwaters and Northern Coldwaters and for the macroinvertebrates the Northern Forest Rivers, Prairie Forest Rivers, High Gradient Northern Forest Rivers, Northern Coldwaters, and the Southern Coldwaters. Eight of the nine had a difference between the median of BCG Level 4 and the Reference Condition of 10 points or fewer. The median of BCG Level 4 for the Northern Rivers fish class was 25 points less than the Reference Condition. This difference was a result of the overall high condition of fish communities in this class (see Figure 6) and may be more of an indication of the applicability of the HDS in large rivers where impacts are mitigated through dilution and greater overall stream stability. For example, large rivers can perform better than the HDS score might indicate because many of the HDS metrics are at the watershed level and may not reflect reach scale conditions in a large river. Although stressors far up in a watershed count against a site's HDS score, the effects of these stressors may be localized, diluted, or mitigated upstream of the sample reach. Interestingly, the 50th percentile of BCG Level 4 is more protective than the Reference Condition for the Northern Forest Rivers macroinvertebrate class. This may be the result of differences between fish and macroinvertebrate community patterns where fish generally increase in species richness in larger streams whereas macroinvertebrate richness may peak at mid-order streams (Vinson & Hawkins 1998).

Minnesota's approach to biocriteria development is novel compared to other states with TALU biocriteria although similar tools (i.e., BCG and Reference Condition) were used in the development process. To develop biocriteria that are protective of the structural and functional health of biological communities, we use the median of BCG Level 4 to set biocriteria. Communities at the middle of this level can be best characterized as possessing "*overall balanced distribution of all expected major groups; ecosystem functions largely maintained through redundant attributes*" which is in line with the language in Minnesota Rule [Minn. R. ch. 7050.0150 subp. 3]. Analysis of Minnesota's Reference Condition is most closely aligned to thresholds developed using the median of BCG Level 4. There are also several examples from other states that have placed their General aquatic life use goal thresholds within or near Level 4 on the BCG (e.g., Maine, Ohio, and Vermont).

Table 5. Candidate General Use biocriteria with a comparison of BCG and Reference Condition derived thresholds. Numbers in parentheses are the difference between each BCG candidate criteria and the Reference Condition criteria (Abbreviations: %ile = percentile, RR = high gradient, GP = low gradient).

Class #	Class Name	Reference Condition	BCG4 25th %ile	BCG4 50th %ile	BCG4 75th %ile	BCG3 25th %ile
Fish						
1	Southern Rivers	43	37 (-6)	49 (6)	61 (18)	54 (11)
2	Southern Streams	38	41 (4)	50 (13)	58 (21)	47 (10)
3	Southern Headwaters	46	45 (-1)	55 (9)	62 (16)	64 (18)
4	Northern Rivers	62	30 (-32)	38 (-25)	52 (-11)	46 (-16)
5	Northern Streams	55	37 (-18)	47 (-8)	59 (4)	47 (-8)
6	Northern Headwaters	45	31 (-14)	42 (-3)	56 (11)	48 (3)
7	Low Gradient Streams	30	31 (1)	42 (12)	54 (24)	49 (19)
10	Southern Coldwater	48	37 (-11)	50 (3)	59 (12)	63 (16)
11	Northern Coldwater	37	26 (-12)	35 (-2)	55 (18)	37 (0)
Macroinvertebrates						
1	Northern Forest Rivers	51	39 (-12)	49 (-3)	53 (2)	60 (9)
2	Prairie Forest Rivers	32	24 (-8)	31 (-1)	38 (6)	37 (5)
3	Northern Forest Streams RR	58	43 (-14)	53 (-5)	58 (0)	58 (0)
4	Northern Forest Streams GP	52	39 (-12)	51 (0)	58 (7)	56 (4)
5	Southern Streams RR	32	31 (-1)	37 (5)	44 (11)	49 (17)
6	Southern Forest Streams GP	40	38 (-3)	43 (3)	50 (10)	48 (8)
7	Prairie Streams GP	37	33 (-5)	41 (4)	47 (10)	58 (21)
8	Northern Coldwater	33	22 (-11)	32 (-1)	42 (9)	32 (-1)
9	Southern Coldwater	49	30 (-19)	43 (-6)	56 (7)	54 (5)

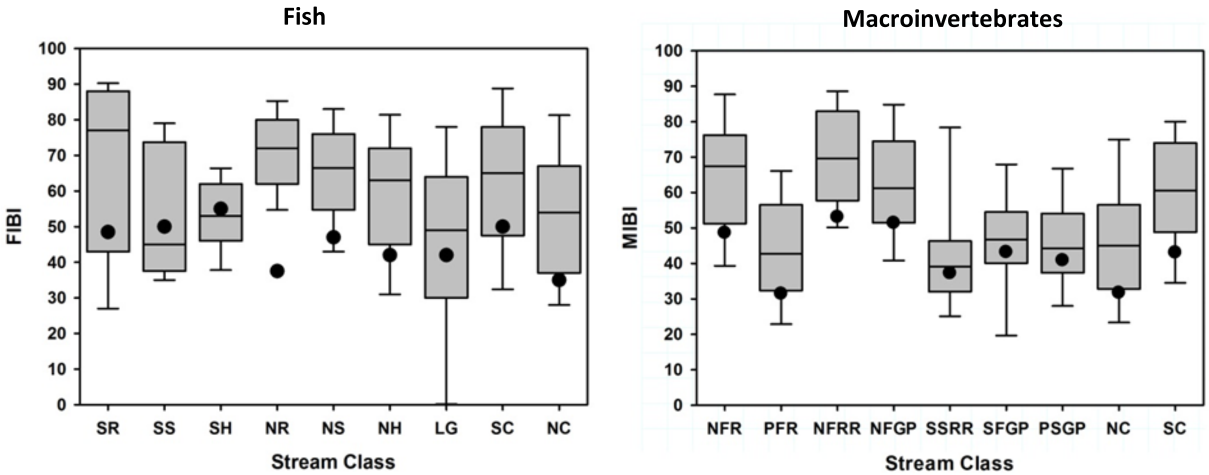

Figure 11. Frequency distribution of fish IBI (FIBI) and macroinvertebrate IBI (MIBI) scores at reference sites in Minnesota by stream class. The General Use biocriterion (●) is the median of the class-specific BCG Level 4 for all classes. Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 90th and 10th percentiles; Abbreviations: SR = Southern Rivers, SS = Southern Streams, SH = Southern Headwaters, NR = Northern Rivers, NS = Northern Streams, NH = Northern Headwaters, LG = Low Gradient Streams, SC = Southern Coldwater, NC = Northern Coldwater, NFR = Northern Forest Rivers, PFR = Prairie Forest Rivers, NFRR = High Gradient Northern Forest Streams, NFGP = Low Gradient Northern Forest Streams, SSRR = High Gradient Southern Streams, SFGP = Low Gradient Southern Forest Streams, PSGP = Low Gradient Prairie Streams.

The use of the median of BCG Level 4 will produce consistently protective biocriteria for streams across Minnesota that will not result in regions with heavy overall disturbance to be held to a lower standard. Most importantly, the BCG permits Minnesota to set criteria that will be at least protective of the Minnesota's aquatic life use goals in regions were too few minimally or least disturbed reference sites are available. By using the median of BCG Level 4 as a threshold we are recognizing the fact that the biologists involved with BCG development have placed the goal between Levels 3 and 4. Also, the use of the median allows for some uncertainty or variation within the BCG level. The BCG is in reality a continuum even though discrete levels are portrayed along the gradient for the sake of clarity and communication (Figure 12). Specifically the IBIs within BCG levels fulfill this continuum and provide a continuous measure of this intra-level gradient. We consider the narrative associated with each BCG level to apply best to the center of the respective level. Toward the margins of each level, characteristics of the adjacent levels can become apparent. As such, biological communities toward the bottom of BCG Level 4 are starting to show some negative attributes observed in BCG Level 5. Therefore, locating the goals at the bottom of BCG Level 4 will likely result in under protective biocriteria. In contrast, the use of the median allows sufficient protection of Minnesota's aquatic life use goals and is an additional safety factor in the criteria setting process.

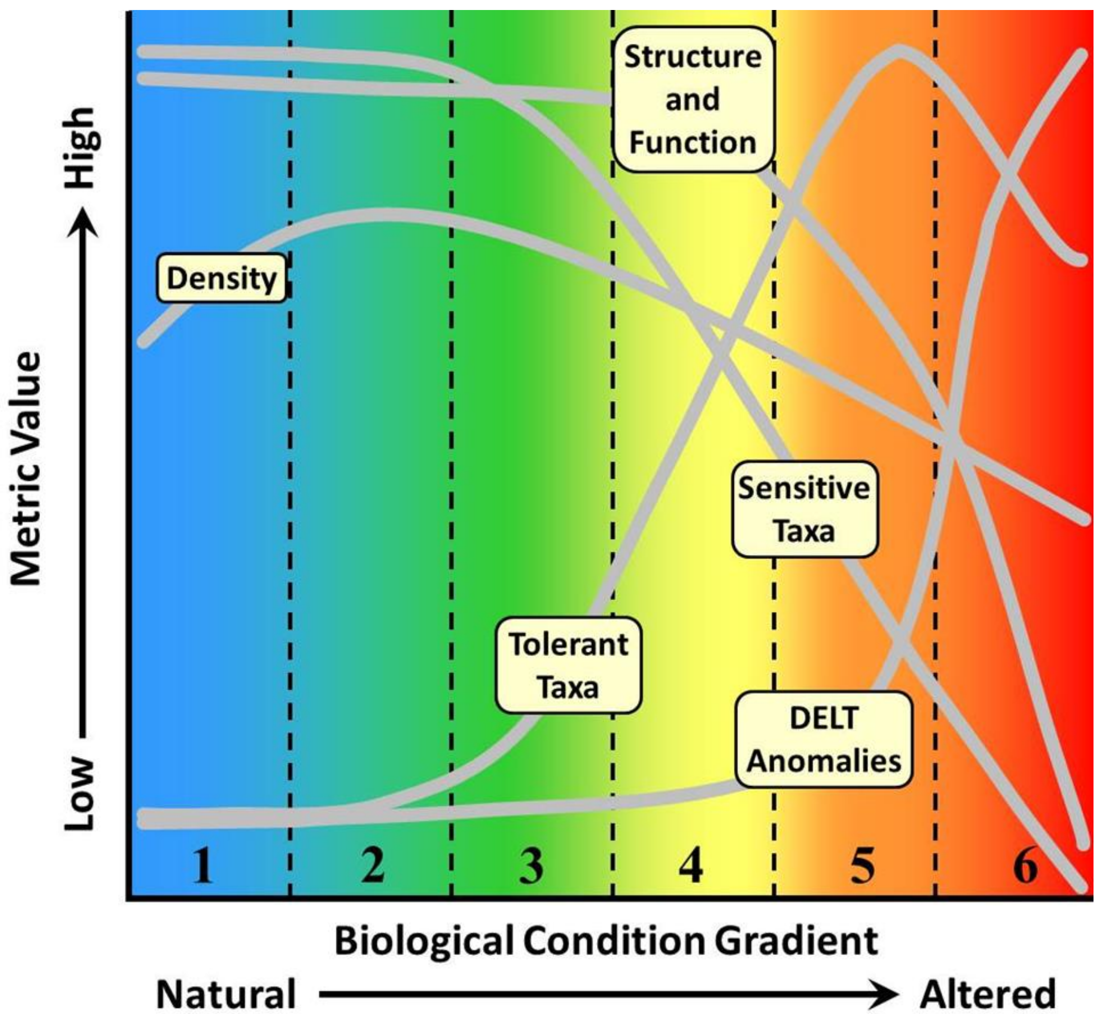

Figure 12. Conceptualization of biological community characteristic changes along the biological condition gradient (DELT = deformities, erosions, lesions, or tumors).

9. Exceptional use biological criteria development

Exceptional Use waters exhibit the highest quality assemblages (as measured by assemblage attributes and indices) in Minnesota. These communities have minimal changes from the natural condition in the structure of the biological assemblage and in ecosystem function. The designation functions as a preservation use, which means it is intended for waters that already exhibit or have the realistic potential to attain an exceptional quality as measured by the biological criteria. On the BCG, Exceptional Use extends from BCG Level 1 into Level 3 (see Figure 3). There are few examples from other states that can be used as a model for the development of Exceptional use biocriteria in Minnesota. The best model for developing Minnesota's Exceptional Use is Ohio. To set biocriteria for Exceptional Use waters, Ohio used the 75th percentile of all reference sites. However, this is not feasible in Minnesota because each index is calibrated independently using datasets with different ranges of condition (i.e. BCG level) between classes. Therefore, IBI scores are not equivalent between classes which results in the need for a class-by-class approach. A potential limitation of a class-by-class approach is that in classes where the overall condition is poor there may be too few minimally disturbed sites which would leave these classes without goals for exceptional streams. As with the development of General Use biocriteria, the BCG provides a tool to set biocriteria for stream classes with a limited Reference Condition dataset.

9.1. Development of candidate exceptional use biocriteria

As with the General Use, a multiple lines of evidence approach was used to develop biocriteria for Exceptional Use streams. Based on narrative expectations for an Exceptional Use, biocriteria should fall in BCG Level 2 or 3. BCG Level 3 is described as "*Evident changes in structure due to loss of some rare native taxa; shifts in relative abundance; ecosystem level functions fully maintained*". BCG Level 2 is similar to BCG Level 3 in that the ecosystem functions are maintained. These levels differ in that the presence of all native taxa is maintained in BCG Level 2. BCG Level 2 is described as "*Structure & function similar to natural community with some additional taxa & biomass; ecosystem level functions are fully maintained*". Both levels could describe exceptional communities in Minnesota streams so they are both included in analyses. As with the development of General Use biocriteria, several statistics from the BCG and Reference Condition were analyzed to determine the most appropriate thresholds for attainment of an Exceptional Use. The BCG and reference site datasets only included sites from reaches that were considered to have natural channels (i.e., <50% channelized) as determined by a site visit and aerial photography. From the BCG models, the 50th and 75th percentile of IBI scores for sites within BCG Level 3 and the 25th percentile of IBI scores for sites within BCG Level 2 were calculated. The 75th percentile of IBI scores for least disturbed reference sites (HDS >61) was also calculated as a Reference Condition-based threshold (see Section 6 for a description of reference site selection). The Reference Condition thresholds were then compared to the BCG derived thresholds for each stream class.

Table 6. Number of samples for datasets used to develop Exceptional Use candidate biological criteria. BCG and reference datasets include only sites with natural channels (Abbreviations: RR = high gradient, GP = low gradient)

Class #	Class Name	Reference	BCG3	BCG2
	Fish			
1	Southern Rivers	18	61	17
2	Southern Streams	8	102	22
3	Southern Headwaters	15	49	7
4	Northern Rivers	116	47	106
5	Northern Streams	186	54	149
6	Northern Headwaters	215	127	120
7	Low Gradient Streams	111	52	17
10	Southern Coldwater	61	101	14
11	Northern Coldwater	196	118	71
	Macroinvertebrates			
1	Northern Forest Rivers	83	56	15
2	Prairie Forest Rivers	9	28	0
3	Northern Forest Streams RR	162	159	11
4	Northern Forest Streams GP	210	135	14
5	Southern Streams RR	15	57	0
6	Southern Forest Streams GP	25	57	0
7	Prairie Streams GP	13	12	0
8	Northern Coldwater	185	56	31
9	Southern Coldwater	60	73	0

Counts of samples from BCG Level 3 were sufficient (see Sample size sufficiency for developing biocriteria, pg. 44) for all classes with the exception of the macroinvertebrate class for low gradient prairie streams (Class 7) which had only 12 samples (Table 6). In contrast, many of the classes for both fish and macroinvertebrates had too few BCG Level 2 samples (Table 6). In fact only a single macroinvertebrate class and five fish classes had at least 20 samples in these datasets. Most of the reference site datasets had more than 20 samples with the exception of the four macroinvertebrate southern warmwater classes (Prairie Forest Rivers, High Gradient Southern Streams, Low Gradient Southern Forest Streams, and Low Gradient Prairie Streams) and the three southern warmwater fish classes (Southern River, Southern Stream and Southern Headwater) (Table 6).

Three different BCG statistics were calculated for the Exceptional Use biocriteria analysis: 1) the 25th percentile of IBI scores for BCG2, (2) the 75th percentile of IBI scores for BCG3, and 3) the 50th percentile of IBI scores for BCG3. These three BCG statistics were compared to the 75th percentile of IBI scores from reference sites. For the Reference Condition, reference sites were selected using the criteria discussed in Section 6.

9.2. Comparison of BCG and reference condition derived thresholds

Due to the small sample sizes for most of the classes, the 25th percentile of BCG Level 2 was difficult to assess (Table 6). This was especially true of the macroinvertebrate classes although this statistic seemed to be a reasonable estimate of the Reference Condition derived threshold for the fish classes (

Table 7). Due to the data limitations and the fact the 25th percentile of BCG Level 2 and the 75th percentile of BCG Level 3 are often similar (Figures 9 and 10) this statistic was not considered further. The 50th percentile of BCG Level 3 was, with the exception of four southern classes, considerably less protective than the Reference Condition (

Table 7). As a result it appears to be an unsuitable statistic for setting Exceptional Use biocriteria. A comparison between candidate thresholds derived using the Reference Condition and those developed using the different BCG statistics indicated that the 75th percentile of BCG Level 3 was most similar to the Reference Condition (

Table 7).

The similarity of 75th percentile of BCG Level 3 to the 75th percentile of the Reference Condition varied from class to class with four fish classes and three macroinvertebrate classes with at least a 10 point difference between these datasets. However the reference thresholds for two fish classes and three macroinvertebrate classes (all southern classes) had Reference Condition sample sizes of less than 25 so these differences could be explained by the limited reference dataset. The remaining two classes that had differences of at least 10 points were in northern classes where the reference threshold was more protective. The BCG Level 3 datasets in these classes had sample sizes of 42-61 samples indicating that differences were probably not the result of insufficient datasets. These threshold differences may be a reflection of the high quality of the streams in these classes that results in over protective thresholds when using a percentage of the reference condition. It is possible that for classes dominated by minimally disturbed conditions like water bodies in northern Minnesota, a 50th percentile of the Reference Condition is more appropriate. Based on these comparisons, the 75th percentile of BCG Level 3 is the most appropriate threshold for setting the Exceptional Use biocriteria. It is largely comparable to the Reference Condition, but does not suffer from the small sample sizes that are observed with many of the Reference Condition datasets. Use of the BCG to set Exceptional Use biocriteria is consistent with the General Use thresholds which were also derived using the BCG. A comparison of the candidate Exceptional Use criteria and distribution of IBI scores for the reference sites is provided in Figure 13.

Table 7. Candidate Exceptional Use biocriteria with a comparison of BCG and Reference Condition derived thresholds. Numbers in parentheses are the difference between each BCG candidate criteria and the Reference Condition thresholds (Abbreviations: %ile = percentile, RR = high gradient, GP = low gradient).

Class #	Class Name	Reference Condition	BCG3 50th %ile	BCG3 75th %ile	BCG2 25th %ile
Fish					
1	Southern Rivers	88	64 (-24)	71 (-17)	70 (-18)
2	Southern Streams	74	55 (-19)	66 (-8)	74 (0)
3	Southern Headwaters	62	69 (7)	74 (12)	73 (11)
4	Northern Rivers	80	57 (-23)	67 (-13)	65 (-15)
5	Northern Streams	76	55 (-21)	61 (-15)	63 (-13)
6	Northern Headwaters	72	60 (-12)	68 (-4)	64 (-8)
7	Low Gradient Streams	64	58 (-6)	70 (6)	59 (-6)
10	Southern Coldwater	78	72 (-6)	82 (4)	75 (-4)
11	Northern Coldwater	67	47 (-21)	60 (-7)	54 (-13)
Macroinvertebrates					
1	Northern Forest Rivers	76	68 (-9)	77 (1)	72 (-4)
2	Prairie Forest Rivers	57	44 (-13)	63 (7)	-
3	Northern Forest Streams RR	83	70 (-13)	82 (-1)	81 (-2)
4	Northern Forest Streams GP	74	67 (-8)	76 (2)	81 (6)
5	Southern Streams RR	46	54 (8)	62 (16)	-
6	Southern Forest Streams GP	55	58 (3)	66 (11)	-
7	Prairie Streams GP	54	61 (7)	69 (15)	-
8	Northern Coldwater	57	40 (-17)	52 (-4)	52 (-5)
9	Southern Coldwater	74	63 (-11)	72 (-2)	-

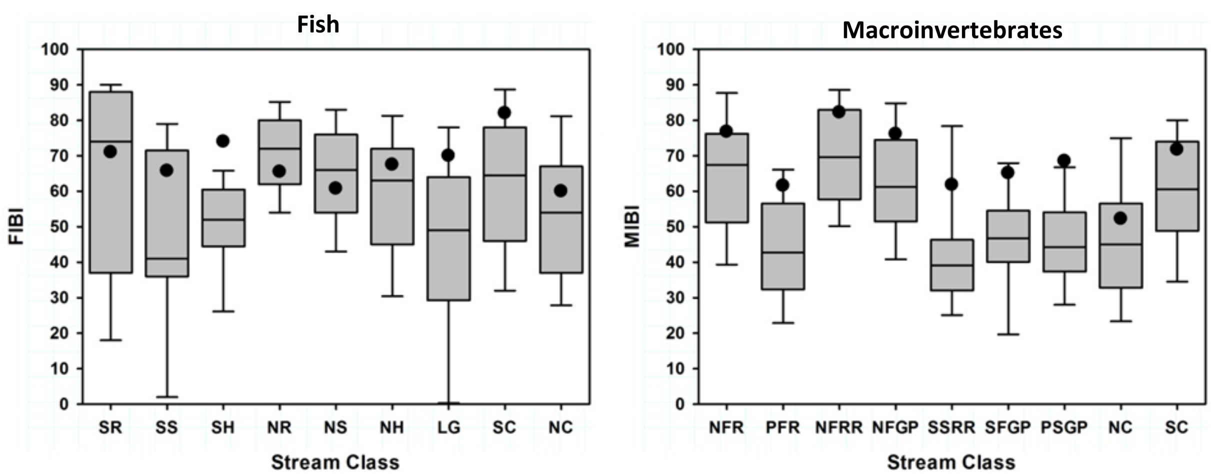

Figure 13. Frequency distribution of fish IBI (FIBI) and macroinvertebrate IBI (MIBI) scores at reference sites in Minnesota by classification strata. The Exceptional Use biocriterion (●) is set at the 75th percentile of the class-specific BCG Level 3 for all classes. Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 95th and 5th percentiles; Abbreviations: SR = Southern Rivers, SS = Southern Streams, SH = Southern Headwaters, NR = Northern Rivers, NS = Northern Streams, NH = Northern Headwaters, LG = Low Gradient Streams, SC = Southern Coldwater, NC = Northern Coldwater, NFR = Northern Forest Rivers, PFR = Prairie Forest Rivers, NFRR = High Gradient Northern Forest Streams, NFGP = Low Gradient Northern Forest Streams, SSRR = High Gradient Southern Streams, SFGP = Low Gradient Southern Forest Streams, PSGP = Low Gradient Prairie Streams.

10. Modified use biological criteria development

Some activities in Minnesota have resulted in legacy impacts to streams that currently have difficulty meeting Minnesota's aquatic life General Use goals. These activities include stream channelization that was performed under Minnesota Drainage Law (Minnesota Statute 103E). The relationships between aquatic communities and reduced habitat condition have been well documented (Gorman & Karr 1978, Karr et al. 1986, Schlosser 1987). The biological limitation of these streams is imposed by insufficient habitat for supporting aquatic communities that meet Minnesota's General Use goals. Despite these limitations, when these watersheds are managed appropriately (i.e., maintaining buffers, etc.) these systems should still be expected to meet some goal below General Use, and not be written off as waters that are incapable of supporting aquatic life or providing beneficial uses other than drainage. In fact, biological data collected by the MPCA clearly demonstrates that some of these channelized waterways have the potential to meet the General Use goals. Under TALU they will be held to a reasonable goal that accounts for the loss of habitat and is reflective of the biological potential of a properly managed channelized stream. In accordance with the CWA, to determine when a Modified Use applies, a Use Attainability Analysis (UAA) will be performed to determine that the system cannot meet the General Use and that habitat is limiting this use. In cases where the habitat is deemed to be limiting, an evaluation is then required to determine if the habitat condition is the result of legal activities and that it cannot be restored (Yoder 2012). If these criteria are met, the stream would be eligible for a Modified Use. It is an objective of Minnesota and the CWA that these modified systems will ultimately be able meet at least General Use goals when the technology makes attainment of these goals feasible (i.e., multiuse drainage ways). In this regard, the Modified Use can be considered a temporary use until these technologies are developed and proven to be feasible and effective.

10.1. Development of candidate modified use biocriteria

As with the other use tiers, the Reference Condition was compared to BCG statistics to determine where the Reference Condition falls on the BCG and to ensure development of consistent biocriteria across Minnesota. Developing Modified Use biocriteria required selection of a set of "reference channelized streams" that represent systems that are managed appropriately (i.e., maintained proper buffer width and other best management practices [BMPs]). These were selected using landscape measures as surrogates of these activities and some water quality measures to filter out sites impacted by upstream chemical stressors.

10.1.1. Selection of modified "Reference Sites"

The first criterion that needed to be met for Modified Use "reference sites" was that the sampling reaches were channelized. If a sampling reach was more 50% channelized as determined from assessments of aerial photography and site visits, it was considered channelized. Candidate reference sites for the Modified Use were identified using watershed and reach level measures of riparian condition. Sites were considered for inclusion if less than 80% of the riparian was disturbed at both the reach level and the watershed level (Table 8). These criteria were intended to match the required permanent 16.5 foot buffer strips of perennial vegetation along drainage ways as required by Minnesota's Drainage Law (Minn. Stat. 103E.021). For the reach level, the percent disturbance was visually estimated as an average of the disturbance (crop, turf grass, roads, etc.) in 15m and 30m buffers on both sides of the stream. The use of the average at two buffer scales gave more weight to the near-stream buffer. At the watershed level, percent disturbance was determined within a 100m buffer for streams upstream of the sampling site using GIS. In addition to these criteria, sites in close proximity to point sources, feedlots, or urbanization were excluded. To filter out sites that are impacted by more than channelized habitat, measures of dissolved oxygen were used. Sites with dissolved oxygen below 4 mg L^{-1} or greater than 12 mg L^{-1} were excluded from the modified reference site dataset.

Table 8. Criteria used to select "reference" modified stream sites.

Metric	Scale	Criteria
% Disturbed Riparian	watershed	<80%
% Disturbed Riparian	reach	<80%*
Dissolved Oxygen (mg/L)	reach	4-12

* Reach riparian condition measured as average of disturbance in 15m and 30m buffers.

A comparison between the percent of disturbed riparian at the watershed and reach level and dissolved oxygen identified the expected responses with declining condition associated with more riparian disturbance and with high or low levels of dissolved oxygen (Figure 14). These analyses were limited to stream reaches that were determined to be more than 50% channelized so the disturbance gradient is truncated and not as distinct when comparing all streams. However, it is apparent that sites with less riparian disturbance perform better biologically. Furthermore, when these three measures are used together along with the three stressor proximity scores to filter out reference sites, these Modified Reference sites perform statistically better biologically than the Modified Non-reference sites for most stream classes (Figure 15). The classes in Figure 15 are limited to the classes with Modified Reference samples sizes of 15 or more samples. A Mann-Whitney Rank Sum test run in SigmaPlot ver. 12 (Systat Software 2011) determined that a significant difference was present between the Modified Reference and Non-reference samples for eight of the nine stream classes. The one class that did not have a significant difference was the macroinvertebrate Southern Streams High Gradient class. However, a difference between the Modified Reference and Non-reference samples for this class is apparent and the non-significant result was in part due to the low power of the test associated with a small samples size. The difference in the 25^{th} percentile of IBI scores between Modified Reference and Non-reference samples for the nine stream classes ranged from 3-22 points with an average difference of 12 points.

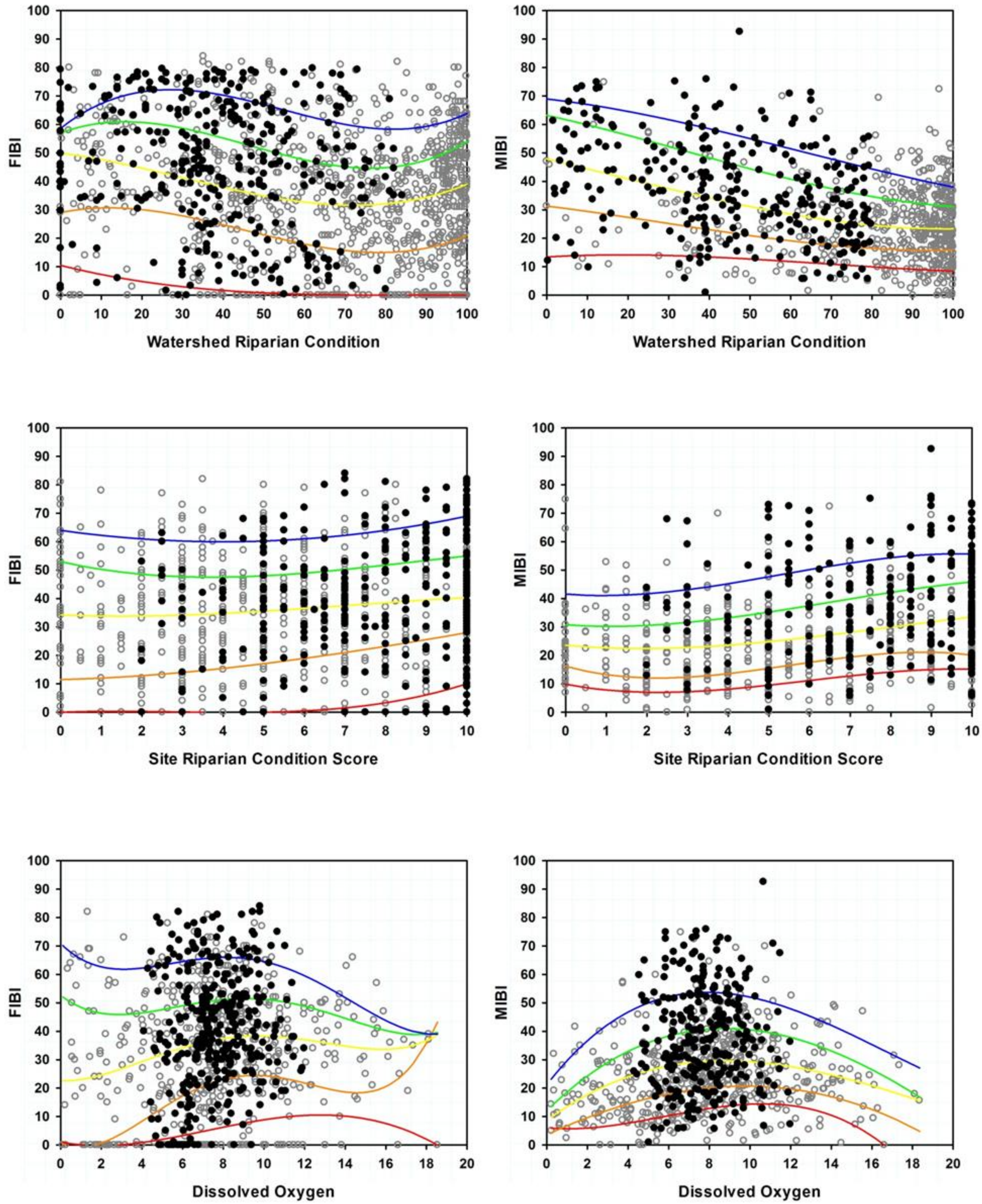

Figure 14. Relationships between watershed riparian condition, sample reach riparian condition, and dissolved oxygen and fish and macroinvertebrate IBIs for stream reaches that are >50% channelized. Black circles are modified reference site samples and open gray circles are modified non-reference site samples. Regressions are quantile regression smoothing fits at the 90th, 75th, 50th, 25th, and 10th percentiles performed in R v. 2.10.0 (R Development Core Team 2009) using "rq" in the "quantreg" package (Koenker 2009) and "bs" in the "splines" package.

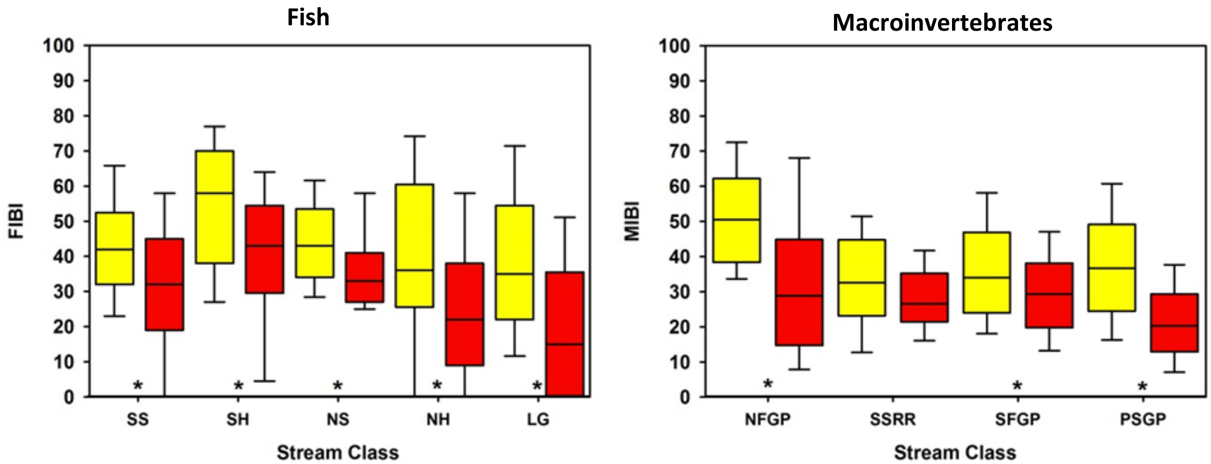

Figure 15. Comparison of IBI scores from modified reference and modified non-reference sites. Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 95th and 5th percentiles, yellow boxes = reference site samples, red boxes = non-reference site samples, * indicates significant difference at the α=0.05 level using a Mann-Whitney Rank Sum test; Abbreviations: R = Reference, NR = Non-reference, SR = Southern Rivers, SS = Southern Streams, SH = Southern Headwaters, NS = Northern Streams, NH = Northern Headwaters, LG = Low Gradient Streams, NFGP = Low Gradient Northern Forest Streams, SSRR = High Gradient Southern Streams, SFGP = Low Gradient Southern Forest Streams, PSGP = Low Gradient Prairie Streams.

10.1.2. Calculation of candidate biocriteria

As with the General and Exceptional Uses, a multiple lines of evidence approach was used to develop biocriteria for Modified Use streams. Based on narrative expectations for a Modified Use, biocriteria should fall in BCG Level 4 or 5. BCG Level 5 is described as:

"*Sensitive taxa markedly diminished; conspicuously unbalanced distribution of major taxonomic groups; ecosystem function shows reduced complexity & redundancy*"

Although this condition is not acceptable for the General Use, it accurately describes a system with reduced habitat diversity which leads to a community with less taxonomic complexity and function (Gorman & Karr 1978). These systems often support more tolerant taxa that are dominated by omnivores and generalists and may have greater biomass due to increased productivity (Yoder & Rankin 1995). Not only will a channelized stream lose species, but changes in ecological function (e.g., nutrient assimilation) is also likely in these systems (Yarbro et al. 1984). Although these changes are not desirable, they reflect current technology of the operation of channelized streams for drainage. Despite these limitations, channelized systems can support beneficial aquatic communities and goals for these systems should reflect what is attainable when these systems are managed appropriately. For example, goals for channelized streams should not allow a nearly complete loss of function and diversity and in extreme cases of biomass. As result, streams that support biological communities that fall into BCG Level 6 would not be considered to be in attainment of aquatic life use goals.

Table 9. Number of samples for datasets used to develop Modified Use candidate biological criteria. BCG datasets included both natural channel and channelized stream sites (Abbreviations: RR = high gradient, GP = low gradient).

Class #	Class Name	Modified Reference	BCG5	BCG4
		Fish		
1	Southern Rivers	12	41	95
2	Southern Streams	33	214	114
3	Southern Headwaters	19	100	371
4	Northern Rivers	3	7	30
5	Northern Streams	53	78	201
6	Northern Headwaters	77	217	73
7	Low Gradient Streams	41	83	114
10	Southern Coldwater	1	62	50
11	Northern Coldwater	6	35	25
		Macroinvertebrates		
1	Northern Forest Rivers	0	6	25
2	Prairie Forest Rivers	9	25	91
3	Northern Forest Streams RR	3	3	42
4	Northern Forest Streams GP	49	43	99
5	Southern Streams RR	18	132	217
6	Southern Forest Streams GP	39	157	112
7	Prairie Streams GP	68	284	183
8	Northern Coldwater	7	10	59
9	Southern Coldwater	1	35	118

For this analysis, IBI scores corresponding to the 25th, 50th, and 75th percentiles of BCG Level 5 and the 25th percentile of BCG Level 4 were calculated for each stream class. The 25th percentile of IBI scores for the Modified Reference Condition (see Section 10.1.1) was also determined for this analysis. The BCG Level 4 and 5 sites included both channelized and natural reaches whereas the Modified Reference dataset included only channelized reaches. All 18 classes in BCG Level 4 had 20 or more samples (Table 9). For BCG Level 5, 14 of 18 classes had dataset sizes of 20 or more samples. The classes with low samples sizes for BCG Level 4 were the fish Northern River class and the macroinvertebrate Northern Forest River, High Gradient Northern Forest Streams, and Northern Coldwater classes. In contrast, the reference dataset had only seven classes with 20 or more samples (Table 9). The large river classes and the coldwater classes in particular had small sample sizes which make development of Modified Use criteria technically difficult or impossible. As a result, Modified Use biocriteria will not be developed for large river and coldwater classes. In addition, the High Gradient Northern Forest Stream class for macroinvertebrates had a very small sample size due to a lack of channelized streams in this class and streams that scored a BCG level of 5 or lower. Therefore, it is also not appropriate to consider a Modified Use for this class of streams.

10.2. Comparison of BCG and reference condition derived thresholds

A comparison of candidate Modified Use thresholds were made between four different BCG statistics (25th percentile of BCG Level 4 and the 25th, 50th, and 75th percentile of BCG Level 5) and the 25th percentile of IBI scores from Modified Reference sites. The Modified Reference site dataset was described in Section10.1.1. A comparison of these statistics is provided in Table 10. Further discussion will not include large river, coldwater classes or the macroinvertebrate High Gradient Northern Forest Stream class as discussed in Section 10.1.2. The 25th percentile of BCG Level 5 was, without exception, less protective than the Modified Reference Condition (Table 10). As a result it is an unsuitable statistic for setting Modified Use biocriteria as it is not likely to be sufficiently protective. The remaining BCG statistics were similar when compared to the Modified Reference Condition (Table 10). The median of BCG Level 5 was generally slightly less protective than the Modified Reference Condition. The 75th percentile of BCG Level 5 and the 25th percentile of BCG Level 4 were both consistently more protective than the Modified Reference Condition. A class-by-class assessment indicated that the median of BCG Level 5 for had less than an 8 point difference from the Modified Reference Condition thresholds for all nine classes. The 75th percentile of BCG Level 5 and the 25th percentile of BCG Level 4 had three and seven classes with less than an eight point different from the Modified Reference Condition thresholds, respectively. Differences between the 25th percentile of the Modified Reference Condition and the median of BCG Level 5 ranged from 1-7 points (Table 10). Five of the nine classes were less protective than the Modified Reference Condition, but these differences were small (i.e., seven points or less). The thresholds calculated using the 75th percentile of BCG Level 5 were greater than the Modified Reference Condition thresholds for all classes. Four classes had a 10 point or greater difference between the methods. The thresholds calculated using the 25th percentile of BCG Level 4 were similar to the 75th percentile of BCG Level 5 with two classes with a 10 point difference from the Modified Reference Condition.

The comparison between candidate thresholds derived using the Modified Reference Condition and those developed using the different BCG statistics indicated that the 50th percentile of BCG Level 5 was most similar to the Modified Reference Condition (Table 10). This statistic is largely comparable to the Modified Reference Condition (Figure 16), and does not suffer from the small sample sizes that are observed in some of the Modified Reference Condition stream classes. The relatively high sample sizes should result in less error in the estimation of these thresholds (see Sample size sufficiency for developing biocriteria, pg. 44). Use of the BCG to develop Modified Use thresholds is consistent with the General and Exceptional Use thresholds which were also derived using the BCG.

Table 10. Candidate Modified Use biocriteria with a comparison of BCG and Modified Reference Condition (MRC) derived thresholds. Numbers in parentheses are the difference between each BCG candidate criteria and the MRC criteria. Thresholds are grayed out for classes where analysis indicated that a Modified Use was not appropriate. The total difference calculation was based only on classes where the Modified Use class was appropriate (Abbreviations: %ile = percentile, RR = high gradient, GP = low gradient).

Class #	Class Name	MRC 25th %ile	BCG5 25th %ile	BCG5 50th %ile	BCG5 75th %ile	BCG4 25th %ile
			Fish			
1	Southern Rivers	50	21 (-29)	28 (-22)	44 (-7)	37 (-13)
2	Southern Streams	32	27 (-5)	35 (3)	41 (9)	42 (10)
3	Southern Headwaters	38	18 (-21)	33 (-6)	49 (11)	41 (3)
4	Northern Rivers	47	20 (-27)	25 (-22)	25 (-22)	31 (-16)
5	Northern Streams	34	26 (-8)	35 (1)	45 (11)	36 (2)
6	Northern Headwaters	26	12 (-14)	23 (-3)	32 (6)	29 (3)
7	Low Gradient Streams	22	0 (-22)	15 (-7)	28 (6)	29 (7)
10	Southern Coldwater	-	25 (-)	34 (-)	41 (-)	37 (-)
11	Northern Coldwater	26	14 (-12)	23 (-3)	34 (8)	23 (-4)
			Macroinvertebrates			
1	Northern Forest Rivers	-	29 (-)	44 (-)	57 (-)	39 (-)
2	Prairie Forest Rivers	18	14 (-4)	21 (3)	28 (10)	24 (6)
3	Northern Forest Streams RR	29	21 (-7)	51 (22)	61 (33)	45 (16)
4	Northern Forest Streams GP	38	27 (-12)	37 (-1)	54 (16)	39 (0)
5	Southern Streams RR	23	20 (-3)	25 (1)	31 (8)	30 (7)
6	Southern Forest Streams GP	24	22 (-2)	30 (6)	38 (14)	34 (10)
7	Prairie Streams GP	24	16 (-9)	22 (-2)	30 (6)	28 (4)
8	Northern Coldwater	14	13 (0)	23 (9)	33 (19)	22 (8)
9	Southern Coldwater	-	23 (-)	34 (-)	47 (-)	29 (-)

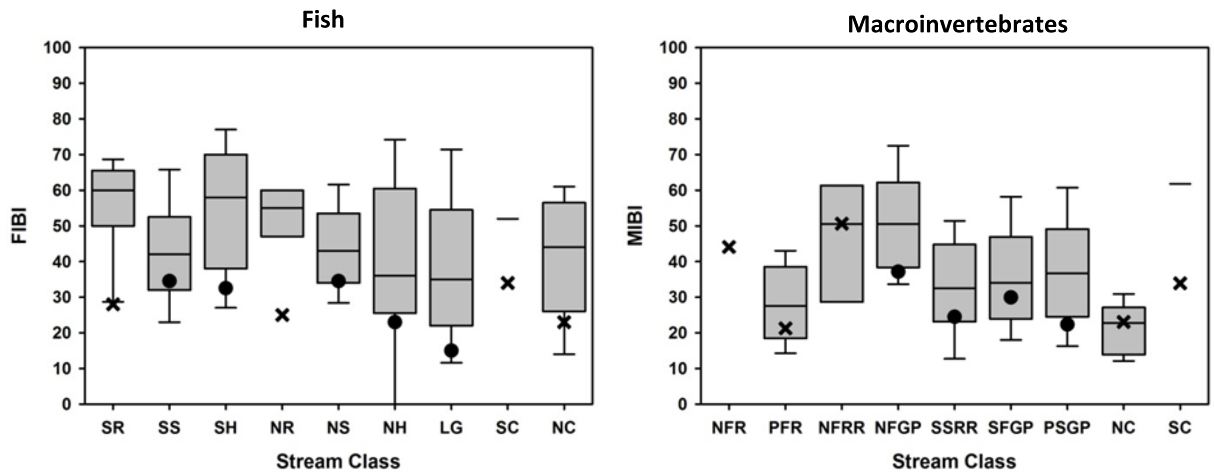

Figure 16. Frequency distribution of fish IBI (FIBI) and macroinvertebrate IBI (MIBI) scores at Modified Reference sites by classification strata. The Modified Use biocriteria (●) are set at the 50th percentile of the class-specific BCG Level 5 for all classes. The 50th percentile for classes for which Modified Use criteria were not developed are indicated by an "×". Symbols: upper and lower bounds of box = 75th and 25th percentiles, middle bar in box = 50th percentile, upper and lower whisker caps = 95th and 5th percentiles; Abbreviations: SR = Southern Rivers, SS = Southern Streams, SH = Southern Headwaters, NR = Northern Rivers, NS = Northern Streams, NH = Northern Headwaters, LG = Low Gradient Streams, SC = Southern Coldwater, NC = Northern Coldwater, NFR = Northern Forest Rivers, PFR = Prairie Forest Rivers, NFRR = High Gradient Northern Forest Streams, NFGP = Low Gradient Northern Forest Streams, SSRR = High Gradient Southern Streams, SFGP = Low Gradient Southern Forest Streams, PSGP = Low Gradient Prairie Streams.

11. Implementation of tiered aquatic life biocriteria

The tiered biological criteria will be integral to performing use designation reviews and assessments of attainment of aquatic life use goals in Minnesota streams. The first step for determining the appropriate use for a stream reach will be to review whether or not the biology meets biocriteria (Yoder 2012). In reaches where the General Use biocriteria are not attained, a UAA will be performed to determine if the habitat is limiting attainment of the biocriteria and if the poor habitat is the result of legal human activity.

Data collected as part of Minnesota's IWM strategy was reviewed to provide a preliminary assessment of the proportion and distribution of each tiered aquatic life use in Minnesota. These data were collected from 40 8-digit HUC watersheds which encompass the range of ecotypes and disturbance gradients found in Minnesota. It should be noted that these reviews were a preliminary assessment and are not final recommendations. The formal review of TALUs will include a process for public input and could include additional data if available. The process used for preliminary assignment of tiered aquatic uses to WIDs (i.e., waterbody IDs) largely followed the approach recommended by Yoder (2012). A description of this process follows.

Reviews of TALUs for WIDs utilized biological data, habitat assessments, chemistry data, the Altered Watercourse (AWC) layer (Krumrie et al. 2013), site visit photos, and aerial imagery. The first step was to review the available reportable biological data from each WID. If all biological visits or the preponderance of the visits indicated attainment of the Exceptional Use biocriteria then the WID was assigned Exceptional Use. If the visit indicated attainment of the General Use biocriteria but not the Exceptional Use, the WID was assigned General Use. This included reaches that were determined to be channelized based on the site visit and AWC layer.

For WIDs with biological data that did not meet the General Use biocriteria a preliminary use attainability analysis was performed. This involved a review of habitat data collected as part of the Minnesota Stream Habitat Assessment (MSHA). Models were developed to predict the probability of attaining biological criteria given a certain suite of habitat features (see MBI 2011). These models (see examples in Figure 17) were used to determine if habitat was likely contributing to nonattainment of the biocriteria. If the probably of attaining the biological criteria was less than 25% based on the number of good habitat attributes, the number of poor habitat attributes, or the ratio of poor/good habitat attributes and the MSHA score, it was considered to be limited by habitat. If this probability was between 25-50% it was considered to be possibly limited by habitat and additional information was considered such as whether or not both assemblages failed the General Use biocriteria to determine if the General was likely attainable. If all probabilities of attainment were above 50% then the WID was not considered to be limited by habitat and a General Use was assigned.

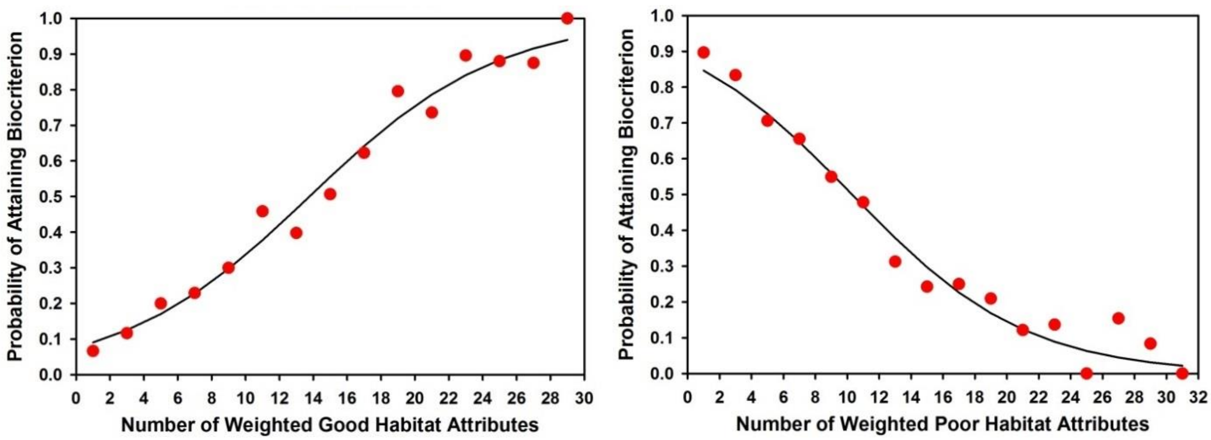

Figure 17. Probability of meeting the General Use biocriterion for fish against the number of good or poor habitat attributes in Northern Headwaters (fit is a logistic regression).

When the biological criteria did not meet the General Use and the habitat was determined to be limiting, a review of the channel status was performed. The purpose of this review was to determine if the reach was altered by channelization. Site visit determinations, aerial imagery, and the Altered Watercourse layer were used in this review. If a WID was determined to not be channelized it was assigned General Use. If it was channelized, then the biocriteria were compared against the Modified Use biocriteria. When the Modified Use criteria were attained, then the WID was assigned Modified Use. If the Modified Use biocriteria were not attained, then the reach was reviewed to determine if major alterations to the habitat were present (e.g., concrete revetments, extensive rip rap).

A total of 1,733 WIDs that comprised 12,472 stream miles were reviewed for TALUs in this preliminary assessment. From these reviewed WIDs, 39 (2%) WIDs were assigned Exceptional Use, 1305 (75%) were assigned General Use, and 389 (22%) were assigned Modified Use. These totals were somewhat different when based on stream miles with 3% (343 miles) assigned Exceptional Use, 84% (10,518 miles) to General Use, and 13% (1,610 miles) to Modified Use. As expected the General Use was the dominant use class. The relatively large number of Modified Uses was a reflection of the proportion of stream miles that are altered, which were determined by the Altered Watercourse study (Krumrie et al. 2013) to be 41,628 mi or 49% of Minnesota's stream miles. There was also a considerable difference in the percent of Modified Use streams between the WID count and the stream mile estimate. This was in part due to the lack of a Modified Use for large rivers which tend to have longer WIDs. The Exceptional Use streams were also a small percentage of the total reviewed WIDs. Most of the Exceptional Use streams (72%) were in northern Minnesota (Northern Lakes and Forests and Northern Minnesota Wetlands ecoregions; Omernik 2002). There were seven Exceptional Use WIDs in the Driftless Area ecoregion, three in the North Central Hardwoods ecoregion, and one in the Western Corn Belt Plains ecoregion.

The relatively small percentage of Exceptional Use waters may be the result the population of watersheds that have been sampled thus far. Most of the watersheds in northeast Minnesota have not been intensively monitored. As a result, the proportion of Exception Use streams is likely to increase when these relatively undisturbed watersheds are monitored.

12. Periodic review of biocriteria

Periodically the biocriteria should be reviewed as a result of incorporation of new data and due to changing conditions: either a result of improving or declining aquatic resource conditions or large-scale impacts such as climate change. In addition, changes to BMP technology will impact biological goals as higher biological condition becomes achievable in watersheds with considerable human activity. This will be especially true of Modified Uses as drainage management technology improves along with riparian and upland management. Routine resampling of designated reference sites as part of the rotating monitoring approach is the most common approach to assess changes over time. This approach had been discussed for Minnesota and it appears to be a good method to document long term changes in the Reference Condition. However, because the BCG is also used to develop biocriteria, there will need to be a discussion to determine how long-term monitoring is used to revise criteria. It may be as simple as repeating the biocriteria process with the updated dataset. However, some consideration should be given to whether or not the BCG model will need to be revisited. A reasonable timeframe to revisit biological criteria is 10 years as this will fit with the 10 year rotating cycle of the intensive watershed approach.

As part of periodic reviews it may also be necessary to develop IBIs and biocriteria for new stream classes or aquatic life use tiers. For example, it may be determined through additional sampling that the current framework of 18 stream classes for fish and macroinvertebrates, does not sufficiently address natural variation of streams in Minnesota. As a result new stream classes could be identified which will need IBIs and biocriteria to be incorporated into Minnesota's watershed management programs.

13. TALU biological criteria: summary

Using a robust dataset of biological, physical, chemical, and land-use data, a framework of biocriteria were developed for tiered aquatic life uses that are protective of Minnesota's aquatic life use goals. These biocriteria and will result in improved management of streams and rivers in Minnesota (Table 11, Figure 18). The development of 18 fish and macroinvertebrate stream classes across a diverse landscape posed some obstacles, but ultimately this refined classification system permitted the development of more accurate and appropriate goals for Minnesota streams. A specific challenge that resulted from this classification system was the lack of true "reference sites" in the southern classes which resulted in the need to rely more on the BCG to develop protective goals for this region. The BCG provided a common "yardstick" across stream classes of varying condition and allowed consistent and protective goals to be developed for Minnesota streams and rivers. As a result, impairment decisions across the state will be based on thresholds that represent similar levels of impairment as measured by the biota. This refined method will provide a comparable measure of condition status regardless of a streams geographic locality. The BCG also offers narrative descriptors to the biological criteria developed for Minnesota Streams. These are as follows:

> **Exceptional Use:** "*Evident changes in structure due to loss of some rare native taxa; shifts in relative abundance; ecosystem level functions fully maintained.*"
>
> **General Use:** "*Overall balanced distribution of all expected major groups; ecosystem functions largely maintained through redundant attributes.*"

Modified Use: *"Sensitive taxa markedly diminished; conspicuously unbalanced distribution of major taxonomic groups; ecosystem function shows reduced complexity & redundancy."*

These narratives are also consistent with recommended descriptions for these uses (see Yoder 2012). These narratives and their associated biocriteria are consistent with Minnesota's aquatic life use goals. Furthermore, the consistency in the associated biocriteria across stream classes provided by the BCG will ensure that Minnesota is compliant with Minnesota rules regardless of the location of a stream reach. It should be noted that the narrative descriptors of these thresholds represent minimum acceptable conditions for these aquatic life uses. Many streams will exceed the biological goals associated with their designated use and Minnesota rules support maintenance of these waters that exceed minimum goals. As a result, the draft biocriteria do not represent "pollute-down-to" goals for waters that exceed these thresholds. Waters that exceed these goals should be maintained or if possible improved. In practice and as part of water quality standards, Minnesota's antidegradation (Minn. R. ch. 7050.0185) rules protect streams that exceed these minimum goals. Implementation of these tiered criteria and the associated water quality standard components will result in improved management of Minnesota's aquatic resources.

Table 11. Draft biological criteria for Exceptional, General, and Modified Uses (Abbreviations: RR = high gradient, GP = low gradient).

Class #	Class Name	Exceptional Use	General Use	Modified Use
		Fish		
1	Southern Rivers	71	49	NA
2	Southern Streams	66	50	35
3	Southern Headwaters	74	55	33
4	Northern Rivers	67	38	NA
5	Northern Streams	61	47	35
6	Northern Headwaters	68	42	23
7	Low Gradient Streams	70	42	15
10	Southern Coldwater	82	50	NA
11	Northern Coldwater	60	35	NA
		Macroinvertebrates		
1	Northern Forest Rivers	77	49	NA
2	Prairie Forest Rivers	63	31	NA
3	Northern Forest Streams RR	82	53	NA
4	Northern Forest Streams GP	76	51	37
5	Southern Streams RR	62	37	24
6	Southern Forest Streams GP	66	43	30
7	Prairie Streams GP	69	41	22
8	Northern Coldwater	52	32	NA
9	Southern Coldwater	72	43	NA

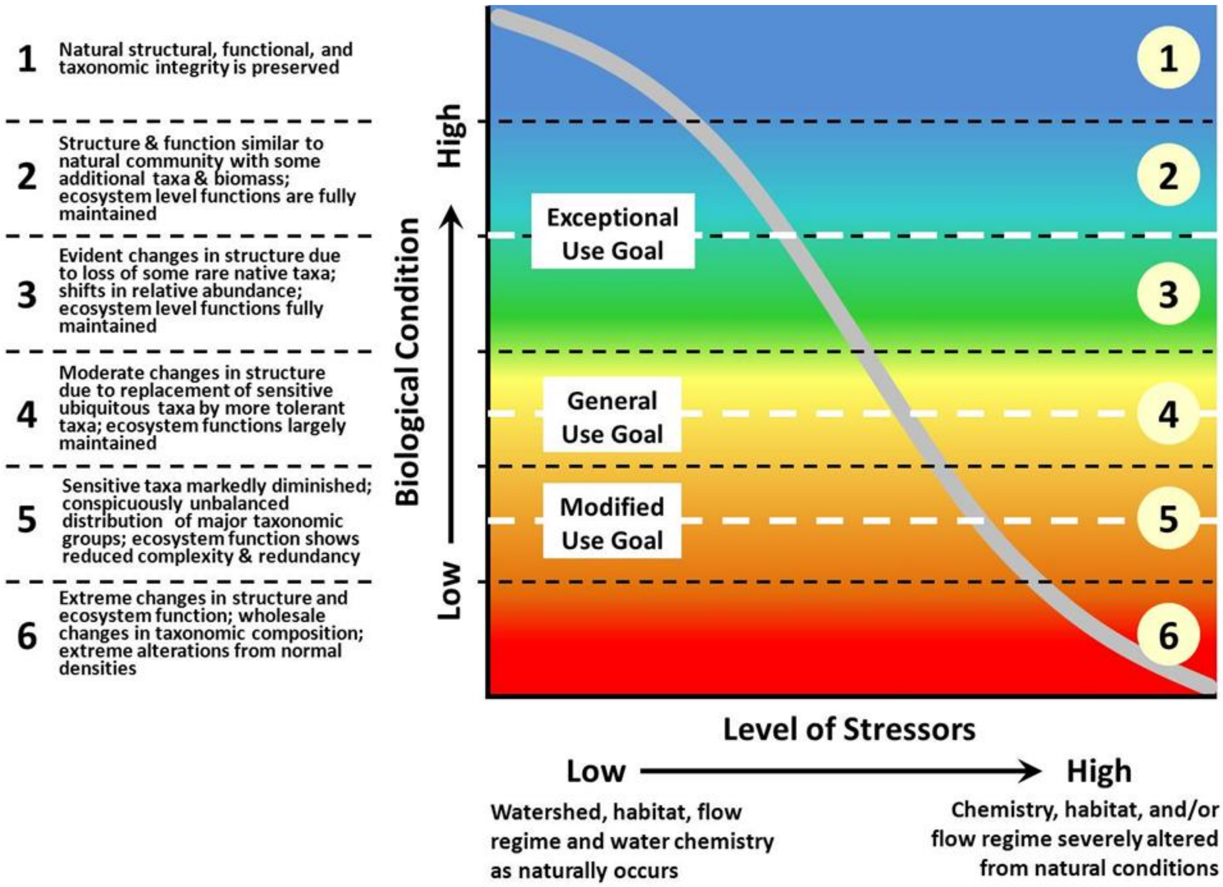

Figure 18. Biological Condition Gradient illustrating the location of draft biocriteria for protection of Minnesota's tiered aquatic life use goals.

Literature cited

Bailey P., J. Enblom, S. Hanson, P. Renard & K. Schmidt (1993) A fish community analysis of the Minnesota River Basin. Minnesota Pollution Control Agency, St. Paul, MN.

Berthouex P. & I. Hau. (1991) Difficulties related to using extreme percentiles for water quality regulations. *Research Journal of the Water Pollution Control Federation* 63: 873-879.

Chirhart J. (2003) Development of a macroinvertebrate index of biological integrity for rivers and streams of the St. Croix River Basin in Minnesota. St. Paul, MN.

Davies S. P. & S. K. Jackson. (2006) The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. *Ecological Applications* 16: 1251-1266.

Genet J. & J. Chirart (2004) Development of a macroinvertebrate Index of biological Integrity (MIBI) for rivers and streams of the Upper Mississipi river basin. Minnesota Pollution Control Agency, St. Paul, MN.

Gerritsen J., L. Zheng, E. Leppo & C. O. Yoder (2013) Calibration of the biological condition gradient for streams of Minnesota. prepared for the Minnesota Pollution Control Agency, St. Paul, MN.

Gibson G. R., M. Barbour, J. B. Stribling, J. Gerritsen & J. R. Karr (1996) Biological criteria: technical guidance for streams and rivers - revised edition. EPA 822-B-96-001. U.S. Environmental Protection Agency, Washington, D.C.

Gorman O. T. & J. R. Karr. (1978) Habitat structure and stream fish communities. *Ecology* 59: 507-515.

Hawkins C. P., R. H. Norris, J. N. Hogue & J. W. Feminella. (2000) Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* 10: 1456-1477.

Hughes R. M., D. P. Larsen & J. M. Omernik. (1986) Regional reference sites: a method for assessing stream potentials. *Environmental Management* 10: 629-635.

Karr J., K. Fausch, P. Angermeier, P. Yant & I. Schlosser. (1986) Assessing biological integrity in running waters: A method and its rationale. *Illinois Natural History Survey Special Publication* 5: 23.

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

Koenker R. (2009) quantreg: Quantile Regression. R package.

Krumrie J., S. Maeder, B. Lundeen & S. Niemela. (2013) Altered Watercourse determination methodology. Report prepared by the Minnesota Geospatial Information Office for the Minnesota Pollution Control Agency, St. Paul, MN.

MBI (2011) Draft - Identification of predictive habitat attributes for Minnesota streams to support Tiered Aquatic Life Uses. Midwet Biodiversity Institute, Columbus, Ohio.

MPCA (2002) Physical Habitat and Water Chemistry Assessment Protocol for Wadeable Stream Monitoring Sites.

MPCA (2004) Invertebrate Sampling Procedures. Minnesota Pollution Control Agency, St. Paul, MN.

MPCA (2009) Fish community sampling protocol for stream monitoring sites.

MPCA (2012) Guidance Manual for Assessing the Quality of Minnesota Surface Waters for Determination of Impairment: 305(b) Report and 303(d) List. Minnesota Pollution Control Agency, St. Paul, MN.

MPCA (2014a) Development of macroinvertebrate indices of biological integrity (MIBI) for Minnesota streams. Minnesota Pollution Control Agency, St. Paul, MN.

MPCA (2014b) Development of fish indices of biological integrity (FIBI) for Minnesota rivers and streams. Minnesota Pollution Control Agency, St. Paul, MN.

MPCA (2014c) Development of a Human Disturbance Score (HDS) for Minnesota streams. Minnesota Pollution Control Agency, St. Paul, MN.

Niemela S. & M. Feist (2000) Index of biotic integrity (IBI) guidance for coolwater rivers and streams of the St. Croix River Basin in Minnesota. Minnesota Pollution Control Agency, St. Paul, MN.

Niemela S. & M. D. Feist (2002) Index of biological integrity (IBI) guidance for coolwater rivers and streams of the Upper Mississippi River Basin. Minnesota Pollution Control Agency, Biological Monitoring Program, St. Paul, MN.

Niemela S. L., P. E, T. P. Simon, R. M. Goldstein & P. A. Bailey (1999) Development of an index of biotic integrity for the species-depauperate Lake Agassiz Plain ecoregion, North Dakota and Minnesota. In: *Assessing the Sustainability and Biological Integrity of Water Resources using Fish Communities* (ed T. P. Simon) pp. 339-365. CRC Press, Boca Raton, FL.

Ohio EPA (1987) Biological criteria for the protection of aquatic life: Volume I: The role of biological data in water quality assessment. Ohio EPA division of Water Quality Planning and Assessment, Columbus, OH.

Ohio EPA (1989) Biological Criteria for the Protection of Aquatic Life: Vol. III. Standardized Biological Field Sampling and Laboratory Methods for Assessing Fish and Macroinvertebrate Communities. Ohio EPA division of Water Quality Planning and Assessment, Columbus, OH.

Omernik J. (2002) Level III ecoregions of the continental United States. Environmental Protection Agency.

R Development Core Team (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Schlosser I. J. (1987) A conceptual framework for fish communities in small warmwater streams. *Community and evolutionary ecology of North American stream fishes*: 17-26.

Stoddard J., D. Larsen, C. Hawkins, R. Johnson & R. Norris. (2006) Setting expectations for the ecological condition of streams: The concept of reference condition. *Ecological Applications* 16: 1267-1276.

Systat Software (2011) SigmaPlot for Windows. Systat Software. Inc., Chicago, Illinois.

USEPA (1990) Biological Criteria: National program guidance for surface waters. EPA-440/5-90-004. U.S. Environmental Protection Agency, Washington, D.C.

USEPA (2005) Use of biological information to better define designated aquatic life uses in state and tribal water quality standards: tiered aquatic life uses. EPA-822-R-05-001. U.S. Environmental Protection Agency, Washington, D.C.

USEPA (2011) A primer on using biological assessment to support water quality management. EPA 810-R-11-01. Office of Science and Technology, Office of Water, Washington, DC.

USEPA (2013) Biological assessment program review: Assessing level of technical rigor to support water quality management. EPA 820-R-13-001. Office of Science and Technology, Washington, DC.

Vermont Depatment of Environmental Conservation (2004) Biocriteria for Fish and Macroinvertebrate Assemblages in Vermont Wadeable Streams and Rivers - Development Phase. Vermont Department of

Environmental Conservation, Water Quality Division, Biomonitoring and Aquatic Studies Section, Waterbury, VT.

Vinson M. & C. Hawkins. (1998) Biodiversity of stream insects: Variation at Local, Basin, and Regional Scales. *Annual Review of Entomology* 43: 271-293.

Whittier T., R. Hughes, J. Stoddard, G. Lomnicky, D. Peck & A. Herlihy. (2007) A Structured Approach for Developing Indices of Biotic Integrity: Three Examples from Streams and Rivers in the Western USA. *Transactions of the American Fisheries Society* 136: 718-735.

Yarbro L. A., E. J. Kuenzler, P. J. Mulholland & R. P. Sniffen. (1984) Effects of stream channelization on exports of nitrogen and phosphorus from North Carolina coastal plain watersheds. *Environmental Management* 8: 151-160.

Yoder C. (1995) Policy issues and management applications of biological criteria. In: *Biological assessment and criteria: Tools for water resource planning and decision making* eds W. S. Davis & T. Simon) pp. 327-344. Lewis Publishers, Boca Raton, FL.

Yoder C. O. (2012) Framework and implementation recommendations for tiered aquatic life uses: Minnesota rivers and streams. Center for Applied Bioassessment and Biocriteria, Midwest Biodiversity Institute, Columbus, OH.

Yoder C. O. & E. T. Rankin (1995) Biological criteria program development and implementation in Ohio. In: *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* eds W. S. Davis & T. P. Simon) pp. 109-144. Lewis Publishers, Boca Raton, FL.

Appendices

Sample size sufficiency for developing biocriteria

Biocriteria are affected by the population of sampling sites that are used to calculate thresholds. There are two main attributes of these datasets that can lead to error in the calculation of biocriteria. First, the sample size can be too small to be sufficient to effectively calculate statistics from the dataset that accurately characterize Aquatic Life Use goals while minimizing error. Second, the population of sites can be biased toward good or poor condition, which needs to be understood and accounted for in order to accurately identify biological thresholds. In reality, many datasets will have this second bias especially when datasets are subdivided to stream classes to account for natural variation. However, the use of tools such as the BCG and an understanding of these systems permit their use in the development of protective and attainable biocriteria that minimize the error associated with condition biases. These stream class condition biases are addressed by the process used by Minnesota to develop biocriteria. This section deals with the first issue: Sample size sufficiency for developing accurate biocriteria.

Independent of sample size, there are approaches that can be used to minimize error with the calculation biocriteria. Specifically, the statistics used to calculate biocriteria can affect error when determining biocriteria. For the development of biological criteria in Minnesota, quantiles such as the median and 25th percentile were used because these are more robust measures than statistics such as the mean which can be strongly affected by outliers. In addition, these are not extreme percentiles which can also improve the accuracy in the estimate of the statistic (Berthouex & Hau 1991). In some cases, the 10th percentile of the Reference Condition was used for stream classes where there is greater confidence that the reference sites are of high quality (i.e., reference) and error is of less concern.

Regardless of the statistic used to determine biocriteria, small sample sizes can introduce errors to these calculations. An examination of estimated standard errors for the statistics used to develop biocriteria in Minnesota was performed to determine optimal sample sizes. Using bootstrap resampling in R (R Development Core Team 2009) the standard errors for the 25th, 50th, and 75th percentiles were estimated for each stream class for BCG Levels 2-5 using natural channel reaches only and for BCG Levels 4-5 for both natural and channelized reaches. This analysis was not performed for BCG Levels 1 and 6 due to the small number of samples in these levels. The 10th and 25th percentiles were estimated for each stream class for the reference and modified reference sites (see Sections 6 and 10) Standard error was plotted against sample size and a locally weighted scatterplot smoothing (LOESS) regression was fit to the data using the “loess” function in R (R Development Core Team 2009). Most of these relationships identified a strong negative relationship between sample size and standard error (Figures 19-21). These relationships tended to be heteroscedastic with more variability in the standard error at low sample sizes. In most of these plots there is an apparent threshold response. For example, with the 25th and 75th percentiles of BCG 3 there was a lower threshold reached at a sample size of ~60 samples (Figure 19). Most lower thresholds for BCG levels using datasets consisting of only natural channel streams were present at a sample size of ~60-100 samples. For BCG Levels 4 and 5 using natural and channelized streams these lower thresholds were reached at ~100-200 samples (Figure 20). Lower thresholds were less apparent for the Reference Condition dataset, but appear to be present at ~100-150 samples (Figure 21). The Modified Reference Condition dataset had the lowest SE among the BCG and Reference datasets with a lower threshold of 40-50 samples. Standard errors for each dataset and stream class are provided in Table 12.

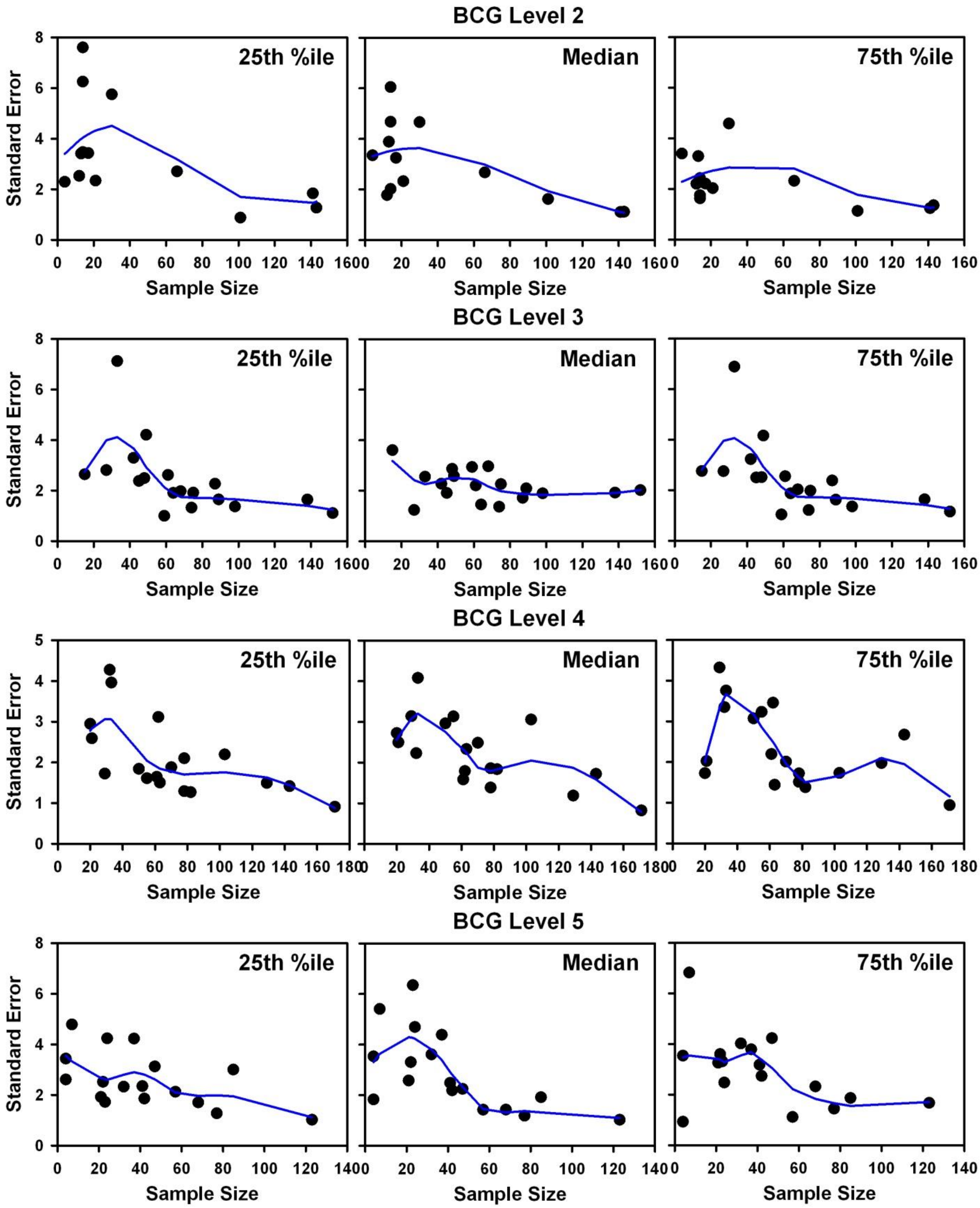

Figure 19. Relationships between sample size and estimates of standard error for statistics used to develop biocriteria from the Biological Condition Gradient. BCG datasets include samples only sites with natural channels. Fitted lines are LOESS regression fits (BCG3, BCG4, BCG5: α=0.75, degree=2; BCG2: α=1, degree=2).

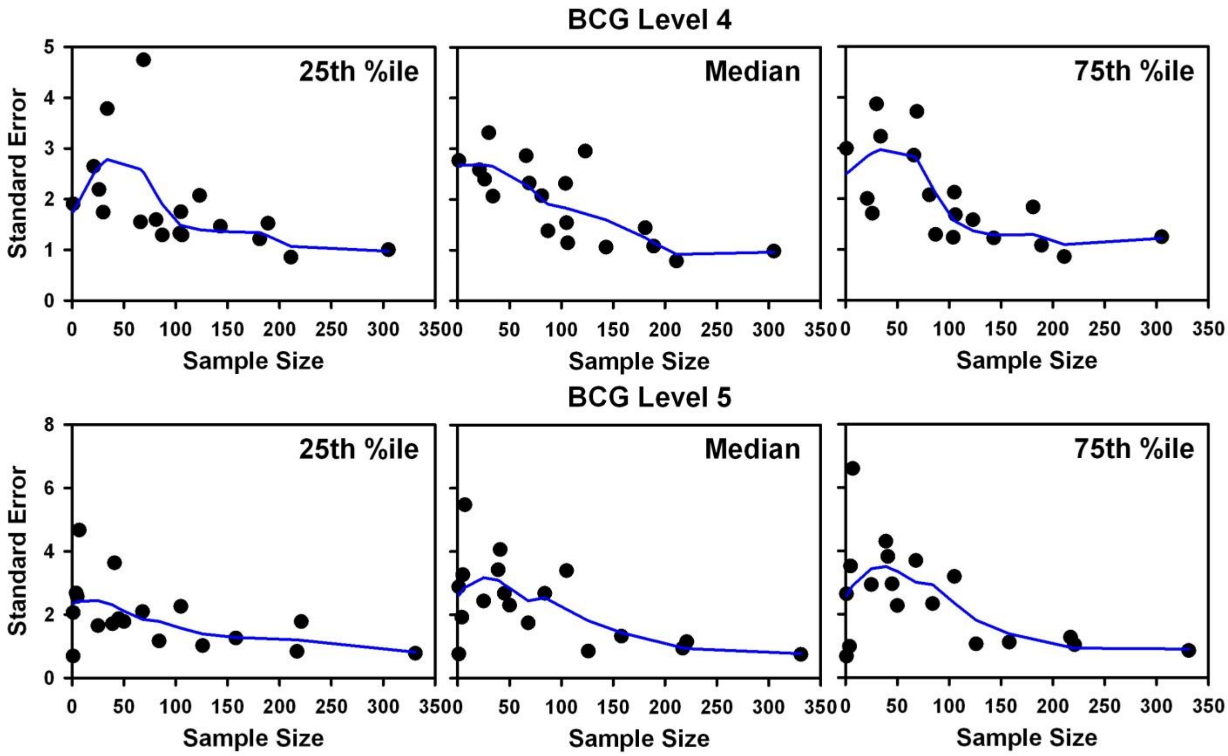

Figure 20. Relationships between sample size and estimates of standard error for statistics used to develop Modified Use biocriteria from the Biological Condition Gradient. BCG datasets include samples both natural channel and channelized sites. Fitted lines are LOESS regression fits (BCG3, BCG4, BCG5: α=0.75, degree=2; BCG2: α=1, degree=2).

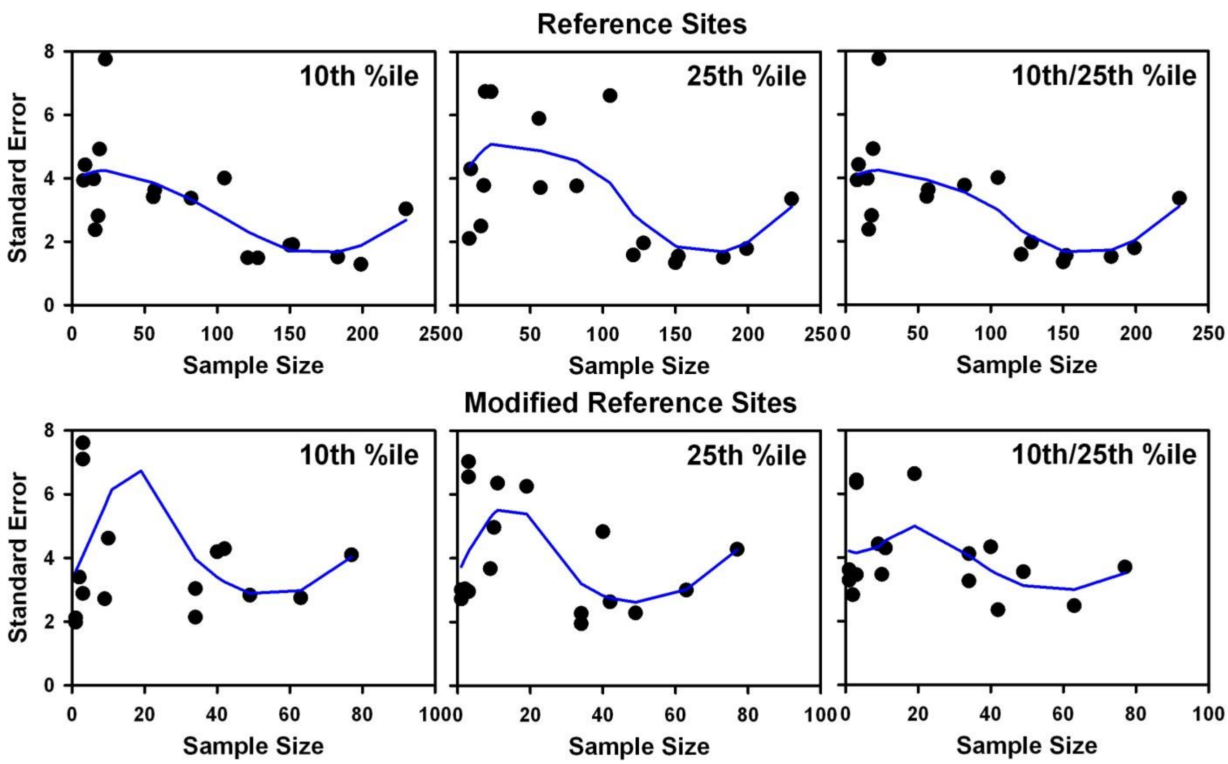

Figure 21. Relationships between sample size and estimates of standard error for statistics used to develop biocriteria from the Reference Condition. Reference dataset includes only data from natural channel sites and the Modified Reference dataset includes both natural channel and channelized sites. Fitted lines are LOESS regression fits (α=0.75, degree=2)

Table 12. Bootstrapped (n=1000) standard errors for quantiles used in biocriteria development. (Abbreviations: SR = Southern Rivers, SS = Southern Streams, SH = Southern Headwaters, NR = Northern Rivers, NS = Northern Streams, NH = Northern Headwaters, LG = Low Gradient Streams, SC = Southern Coldwater, NC = Northern Coldwater, NFR = Northern Forest Rivers, PFR = Prairie Forest Rivers, NFRR = High Gradient Northern Forest Streams, NFGP = Low Gradient Northern Forest Streams, SSRR = High Gradient Southern Streams, SFGP = Low Gradient Southern Forest Streams, PSGP = Low Gradient Prairie Streams.)

Fish										
Dataset	**Statistic**	**SR**	**SS**	**SH**	**NR**	**NS**	**NH**	**LG**	**SC**	**NC**
BCG2	**25th %ile**	3.43	2.34	2.29	0.88	1.84	1.27	6.25	7.60	5.75
	Median	3.25	2.33	3.35	1.62	1.11	1.12	6.04	4.67	4.66
	75th %ile	2.21	2.03	3.39	1.13	1.24	1.35	2.43	1.62	4.58
BCG3	**25th %ile**	0.99	2.27	2.80	3.29	2.37	1.36	2.49	1.64	1.97
	Median	2.94	1.72	1.24	2.27	1.91	1.90	2.86	2.09	2.97
	75th %ile	1.04	2.38	2.75	3.23	2.50	1.35	2.51	1.62	2.03
BCG4	**25th %ile**	1.88	1.64	1.49	1.72	1.41	3.96	3.11	1.61	1.84
	Median	2.49	1.58	1.19	3.14	1.72	4.08	1.79	3.13	2.96
	75th %ile	2.01	2.19	1.97	4.32	2.67	3.75	3.46	3.23	3.07
BCG5	**25th %ile**	2.32	1.02	4.23	2.61	3.12	3.00	1.72	1.85	2.35
	Median	3.61	1.03	4.39	1.83	2.25	1.92	6.35	2.19	2.49
	75th %ile	4.02	1.67	3.79	0.92	4.23	1.86	3.31	2.73	3.18
RCA	**10th %ile**	6.73	2.11	2.50	1.58	1.51	3.35	6.61	3.71	1.55
	25th %ile	7.75	3.93	2.37	1.49	1.51	3.03	4.00	3.62	1.90
	10th/25th %ile	7.75	3.93	2.37	1.58	1.51	3.35	4.00	3.62	1.55
MRC	**10th %ile**	8.50	4.29	8.96	3.39	2.13	4.09	4.19	1.98	7.61
	25th %ile	6.35	2.62	6.25	3.03	1.94	4.27	4.83	2.72	7.03
	50^{th} %ile	4.29	2.36	6.63	2.83	3.26	3.70	4.33	3.28	6.35
Macroinvertebrates										
Dataset	**Statistic**	**NFR**	**PFR**	**NFRR**	**NFGP**	**SSRR**	**SFGP**	**PSGP**	**NC**	**SC**
BCG2	**25th %ile**	3.41		2.52	3.47				2.70	
	Median	3.88		1.78	2.02				2.67	
	75th %ile	3.30		2.20	1.75				2.32	
BCG3	**25th %ile**	1.91	7.12	1.10	1.64	1.32	4.20	2.64	2.61	1.91
	Median	1.45	2.55	2.02	1.92	1.37	2.58	3.60	2.21	2.26
	75th %ile	1.87	6.89	1.15	1.63	1.21	4.16	2.76	2.55	1.98
BCG4	**25th %ile**	2.59	1.29	4.27	1.50	0.91	1.26	2.10	2.94	2.19
	Median	2.50	1.38	2.23	2.33	0.82	1.84	1.86	2.72	3.06
	75th %ile	2.02	1.51	3.35	1.44	0.94	1.38	1.72	1.72	1.73
BCG5	**25th %ile**	4.78	1.92	3.44	4.23	1.27	2.12	1.70		2.51
	Median	5.40	2.58	3.53	4.69	1.19	1.42	1.43		3.30
	75th %ile	6.82	3.26	3.53	2.47	1.44	1.12	2.32		3.60
RCA	**10th %ile**	3.77	4.30	1.34	1.78	3.78	6.74	8.31	1.96	5.89
	25th %ile	3.37	4.42	1.88	1.28	2.81	4.92	3.98	1.49	3.41
	10th/25th %ile	3.77	4.42	1.34	1.78	2.81	4.92	3.98	1.96	3.41
MRC	**10th %ile**		2.71	7.10	2.83	4.62	3.03	2.74	2.88	2.12
	25th %ile		3.67	6.55	2.28	4.96	2.27	3.00	2.95	3.01
	50^{th} %ile		4.42	6.43	3.55	3.47	4.12	2.49	3.46	3.61

An additional assessment was performed to examine how quantile statistics change over time as additional biological samples are added to the dataset each year. The median of IBI scores for BCG Levels 4 and 5, the 75th percentile of BCG Level 3, the 10th/25th percentiles of Reference sites, and the 25th percentile of Modified Reference sites were calculated for each stream class for the years from 1996 through 2010. Changes in the number of samples from year-to-year are not modeled and represent the true accumulation of sampling visits by year for each class over this period. Descriptions of these datasets can be found in Sections 6, 8.3, 9.2, 10.1.1, and 10.1.2 for the Reference Condition, BCG Level 3, BCG Level 4, the Modified Reference Condition, and BCG Level 5, respectively. Year-to-year increases in stream class datasets ranged from 0 to 116 sites. The quantile statistics described above were plotted against the total sample size and a 95th percentile quantile regression was fit to the data. Additive quantile regression smoothing (AQRS) “rqss” in “quantreg” package; Koenker 2009) was performed in the program R ver. 2.10.0 (R Development Core Team 2009). This method is similar to linear quantile regression, but instead of fitting a single line to the data, this approach fits a regression line to subsets of the data (see Figure 23). As a result, AQRS was used to identify changepoints along the outside of the data wedge. The 95th percentile (τ = 0.95) was used to make a conservative determination of the sample size at which error in the estimated statistic is minimized. The AQRS approach required the selection of a lambda (λ) value which determines the amount of smoothing. Values of λ were selected to minimize the number of breakpoints (1-2 breakpoints) and to identify a lower breakpoint where increasing sample size has a minimal effect on estimation of the statistic. After the 95th percentile quantile regression was fitted the lower breakpoint was determined.

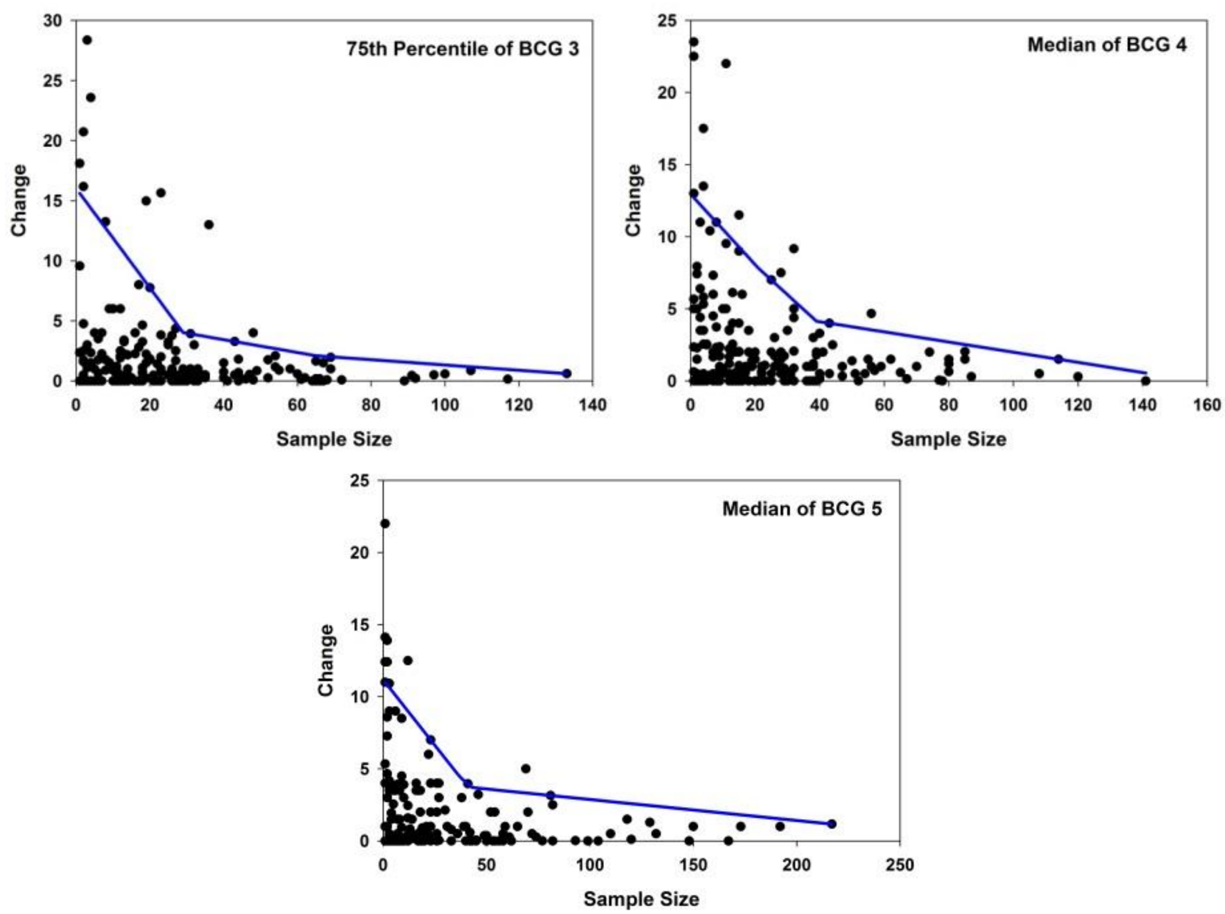

Figure 22. Yearly change in IBI scores for the 75th percentile of BCG Level 3 and the median of BCG Levels 4 and 5 as a function of the sample size.

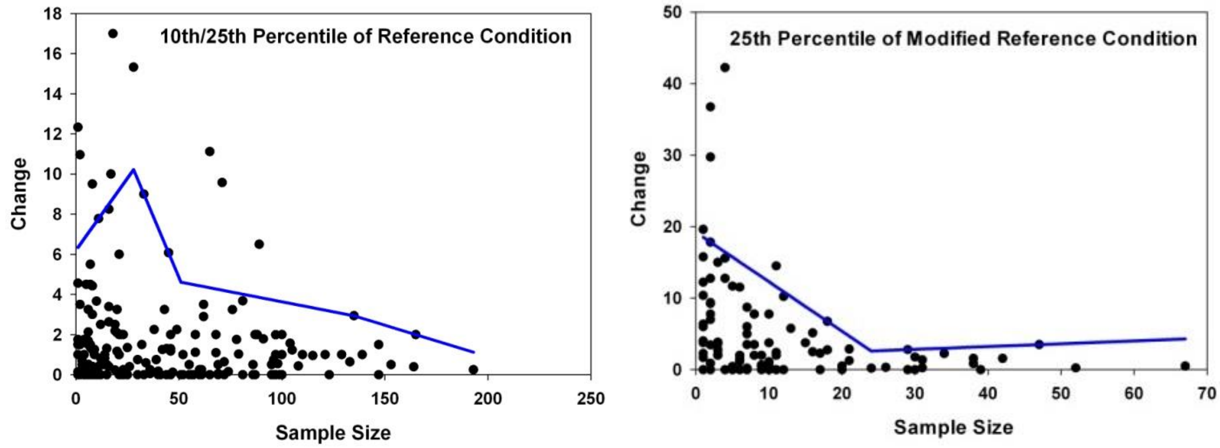

Figure 23. Change in IBI scores for the 10th/25th percentile of Reference Condition and the 25th percentile of the modified Reference Condition as a function of the sample size.

The required size for BCG levels ranged from about 30-80 samples depending on the BCG level and stream class. The Reference Condition required a considerably larger dataset of about 130-150 samples to minimize error. In contrast the Modified Reference Condition dataset required approximately 25-50 samples. Lower break points for the 75th percentile of BCG 3 and the median of BCG Levels 4 and 5 were 29, 39, and 56 samples, respectively (Figure 22). The breakpoint for the 10th/25th percentiles of the reference sites was 51 samples and 24 samples for the 25th percentile of modified reference sites (Figure 23). In general this pattern was similar to the previous analysis of standard errors although the year-to-year change indicated that smaller sample sizes are needed to minimize error. The analyses of the standard error and year-to-year change as a function of dataset sample size indicated that lower sample sizes are needed for accurate estimation of statistics from BCG datasets compared to the Reference Condition dataset.

Although these thresholds are apparent in most of these datasets, estimates of these statistics for some classes had minimal error with a small sample size (see Figures 19-23). However, these small datasets require additional analyses to determine if the statistics estimated are accurate. Therefore the conservative estimates from these analyses can provide a minimum threshold for sample size in the absence of additional analyses to determine the characteristics of datasets used to develop biocriteria. Smaller datasets can be used but error associated with statistics calculated from these datasets should be determined to ensure biocriteria develop from them are an accurate reflection of Minnesota's aquatic life use goals.