Environ Monit Assess (2016) 188:196
DOI 10.1007/s10661-016-5181-y

A novel approach for the development of tiered use biological criteria for rivers and streams in an ecologically diverse landscape

R. William Bouchard Jr. · Scott Niemela · John A. Genet · Chris O. Yoder · John Sandberg · Joel W. Chirhart · Mike Feist · Benjamin Lundeen · Dan Helwig

Received: 16 October 2015 / Accepted: 10 February 2016
© Springer International Publishing Switzerland 2016

Abstract Water resource protection goals for aquatic life are often general and can result in under protection of some high quality water bodies and unattainable expectations for other water bodies. More refined aquatic life goals known as tiered aquatic life uses (TALUs) provide a framework to designate uses by setting protective goals for high quality water bodies and establishing attainable goals for water bodies altered by legally authorized legacy activities (e.g., channelization). Development of biological criteria or biocriteria typically requires identification of a set of least- or minimally-impacted reference sites that are used to establish a baseline from which goals are derived. Under a more refined system of stream types and aquatic life use goals, an adequate set of reference sites is needed to account for the natural variability of aquatic communities (e.g., landscape differences, thermal regime, and stream size). To develop sufficient datasets, Minnesota employed a reference condition approach in combination with an approach based on characterizing a stream's response to anthropogenic disturbance through development of a Biological Condition Gradient (BCG). These two approaches allowed for the creation of ecologically meaningful and consistent biocriteria within a more refined stream typology and solved issues related to small sample sizes and poor representation of minimally- or least-disturbed conditions for some stream types. Implementation of TALU biocriteria for Minnesota streams and rivers will result in consistent and protective goals that address fundamental differences among waters in terms of their potential for restoration.

Keywords Tiered aquatic life uses · Biological condition gradient · Clean Water Act · Biological integrity · Aquatic ecosystems

Electronic supplementary material The online version of this article (doi:10.1007/s10661-016-5181-y) contains supplementary material, which is available to authorized users.

R. W. Bouchard Jr., (✉) · S. Niemela · J. A. Genet · J. Sandberg · J. W. Chirhart · M. Feist · B. Lundeen · D. Helwig
Environmental Analysis and Outcomes Division, Minnesota Pollution Control Agency, 520 Lafayette Road North, Saint Paul, MN 55155-4194, USA
e-mail: Will.Bouchard@state.mn.us

C. O. Yoder
Midwest Biodiversity Institute & Center for Applied Bioassessment & Biocriteria, P.O. Box 21561, Columbus, OH 43221-0561, USA

Introduction

The objective of the Clean Water Act (CWA) is to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters" (United States [U.S.] Code title 33, section 1251 [a]). Although this statement is central to the CWA, interpreting this language and putting it into practice require a structured process. Following adoption of the CWA, debate began regarding how to define and measure "biological integrity" (see Ballentine and Guarraia 1977; Gakstatter et al. 1981). Frey (1977) defined it as "the capability of supporting and maintaining a balanced,

Published online: 27 February 2016

integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region." This and similar definitions (see Karr and Dudley 1981; Karr and Chu 1999) are well accepted and provide more descriptive language to support the objective of the CWA. The CWA also provides an interim goal for U.S. waters that do not meet this objective: "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]). Using these narrative goals and descriptions as a foundation, states, tribes, and territories have developed water quality standards (WQS) to protect aquatic life uses.

Minnesota has established a narrative in rule that describes the minimum restoration goal for aquatic life in waters of the state: "…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" (Minnesota Rule chapter 7050.0150 subpart 3). This narrative can be considered equivalent to the CWA Interim Goal. Chemical, physical, and biological criteria are established to define the conditions that are protective. Biological criteria or biocriteria are integral to the protection and restoration of aquatic life uses because these criteria directly measure the status of aquatic life and can detect the impacts of multiple stressors over time (Yoder 1995; Karr and Yoder 2004). In addition, it is the U.S. Environmental Protection Agency's (USEPA) policy that states integrate biological assessment into WQS programs (USEPA 1990, 2011).

The reference condition (RC) approach is the most commonly used method for determining appropriate biological thresholds. A regional RC approach involves selection of minimally- and least-impacted reference sites from homogenous regions and water body types using an independent measure of potential stressors (Hughes 1995). These populations of sites characterize biological assemblages that attain aquatic life use goals for specific water body types (Hughes et al. 1986; Gibson et al. 1996; Reynoldson et al. 1997; Pont et al. 2006; Stoddard et al. 2006; Hawkins et al. 2010). In the USA, many state monitoring programs use Indices of Biotic Integrity (IBI) to measure the biological condition (Karr et al. 1986; DeShon 1995; Whittier et al. 2007). The RC approach calculates IBI scores from a reference site dataset, and a percentile of the IBI scores, such as the 25th or 10th, is chosen to represent the RC. Excluding the lower quartile or decile eliminates the impact of outliers and creates a safety factor as the procedure for selecting reference sites can result in the selection of some sites that are not minimally or least impacted (Yoder and Rankin 1995). The resulting statistic is the threshold (i.e., biocriterion) used to assess if a biological assemblage is attaining the designated aquatic life use goal for that stream type.

The Biological Condition Gradient (BCG) maps biological information onto a conceptual and theoretical model using the empirical experience of aquatic ecologists (Davies and Jackson 2006). Development of the BCG is dependent on an extensive knowledge of the autecology of aquatic taxa, and how anthropogenic stressors effect aquatic organisms from the organismal to community level (e.g., spawning, reproductive success, and feeding; MPCA 2014a). The BCG model partitions a cumulative stressor gradient into six ordinal categories that describe the biological condition expected for each level (Davies and Jackson 2006; Gerritsen et al. 2013; USEPA 2016). It can be used to describe where biological assemblages fall along a gradient from the natural, unimpacted state to highly altered and degraded. As such, the BCG can provide a more complete range compared to the RC as it is less reliant on prevailing land use conditions.

Numerous models have been developed to assess the condition of biota in freshwaters (e.g., IBIs, multimetric predictive models, River InVertebrate Prediction And Classification System [RIVPACS], AUStralian RIVer Assessment System [AUSRIVAS], and Benthic Assessment of SedimenT [BEAST]; Karr et al. 1986; Wright et al. 1998; Smith et al. 1999; Hawkins et al. 2000a; Reynoldson et al. 2000; Whittier et al. 2007; Pont et al. 2009; Moya et al. 2011). These models use some form of a RC to set expectations, but these goals are often constrained by the extant conditions of the reference sites. For example, in heavily disturbed regions, there may be few minimally- or least-disturbed sites (Gibson et al. 1996) which can result in an incomplete disturbance gradient. The BCG provides a common biological condition scale based on the estimated natural state of biological assemblages whether or not that condition presently exists (MPCA 2014a). As a result, the BCG can be used to understand regional differences in the biological condition and to establish analogous goals across stream types and regions. This is important because in Minnesota natural factors that influence biological assemblages vary regionally in a pattern that co-varies with land use disturbance which challenges our ability to separate the relative influences of each and therefore set meaningful expectations (MPCA 2014a).

The BCG model has proven to be useful for developing biocriteria in the USA (USEPA 2011; USEPA 2016). In fact the BCG, or at least the theory upon which the BCG is based, has often informed the RC approach when developing biocriteria for tiered uses. For example, similar conceptual models have been used in Ohio and Maine to develop biological goals (Davies and Jackson 2006; USEPA 2011). Ohio used the RC approach to set biocriteria and a BCG-like approach to test and refine thresholds (State of Ohio 2010). For example, using this approach, Ohio determined that the 25th percentile of the RC was not protective of aquatic life uses for the Huron-Erie Lake Plain (HELP) ecoregion and instead used the 90th percentile of all sites in this ecoregion to establish biocriteria (Ohio EPA 1987, 1989; Yoder and Rankin 1995). Maine used a BCG model to develop biocriteria using a probability-based multivariate analysis (Courtemanch et al. 1989).

The Minnesota Pollution Control Agency (MPCA) routinely samples fish and macroinvertebrate assemblages in streams and rivers[1] to assess attainment of aquatic life use goals. Multiple assemblages are monitored because each may respond to different stressors at different spatial and temporal scales (USEPA 2013) and therefore provide a more complete assessment of biological condition that is more likely to detect aquatic life impairment when it exists (Yoder and Rankin 1995). Minnesota's current aquatic life use framework uses a one-size-fits-all approach (i.e., there is a single biocriterion used to measure attainment for each stream type IBI model). However, Minnesota is in the process of developing a tiered aquatic life uses (TALU) framework that divides the current aquatic life use goal for streams into three tiers based on the biological potential of each stream. As part of this effort, Minnesota has developed tiered use biocriteria for exceptional, general, and modified uses for Minnesota streams using RC and BCG statistics. Candidate biocriteria were calculated from minimally-/least-disturbed reference site datasets, "modified" reference site datasets, and BCG models for each of the fish and macroinvertebrate stream types. The objective of this work was to develop statewide biocriteria for Minnesota's draft TALUs that are consistently protective of different stream types and compliant with Minnesota and U.S. rules.

[1] "Streams and rivers" are hereafter referred to collectively as "streams"

Methods

Datasets used to develop biocriteria

Macroinvertebrate and fish data used to develop biocriteria were the result of stream surveys in Minnesota from 1996 through 2011 from 3009 sampling sites. The dataset included biological (i.e., fish [2886 sites] and macroinvertebrates [2455 sites]), chemical, physical, and land use data (MPCA 2002). Sampling was limited to perennial streams and streams that were wetted for a sufficient time period to permit rapid recolonization. As a result, the IBI models are only applicable to site samples that meet these criteria. The length of each sample site was 35 times the stream width with a minimum length of 150 m and maximum length of 500 m (Lyons 1992; Meador et al. 1993).

Fish sampling was conducted during daylight hours within the summer index period of mid-June through mid-September (MPCA 2009). All habitat types within an established site were sampled in the approximate proportion that they occur. An effort was made to collect and identify all fish ≥25 mm in total length. Fish were collected with electrofishing, using one of four methods: (1) backpack electrofisher in small headwater streams, (2) towed stream electrofisher in larger wadeable streams, (3) mini-boom electrofisher in small, non-wadeable rivers, and (4) boat-mounted boom electrofisher in large, non-wadeable rivers. Fish were identified in the field to species; enumerated; the maximum and minimum size for each species was measured; and any deformities, erosions, lesions, or tumors were counted. Vouchers of fish species were retained and sent to the University of Minnesota Bell Museum (St. Paul, MN) for identification confirmation.

For macroinvertebrates, a multi-habitat method was used to obtain a sample representative of the macroinvertebrate assemblage within a sampling site during the summer index period of August through October (MPCA 2004). The habitats sampled include hard bottom (riffle/cobble/boulder), aquatic macrophytes (submerged/emergent vegetation), undercut banks (undercut banks/overhanging vegetation), wood (snags/rootwads), and leaf packs. Twenty D-frame dip net (500-μm mesh) sweeps were divided equally among the dominant habitats present in the site and composited. Each sweep covers approximately 0.09 m^2 of substrate for a total area of approximately 1.9 m^2 sampled. Macroinvertebrate samples were sub-sampled to a minimum of 300 organisms followed by a large and rare pick (Courtemanch 1996; Vinson and Hawkins 1996). Identifications were made to the genus level or higher (e.g., family) with Chironomidae identified to genus.

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 2014b, c). 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 2014c). An IBI model was developed for each stream type (i.e., 18 total IBI models; MPCA 2014b, c) using the approaches described by Whittier et al. (2007). To calculate 90 % confidence limits for IBI scores, the residual error term from an analysis of variance (ANOVA) was used to estimate variation due to measurement error (Fore et al. 1994) for each IBI class. The datasets used to estimate confidence limits included replicate samples collected from sites on the same day (macroinvertebrates) or the same year within the sample index period (fish). These datasets were not restricted to only RC sites due to small numbers of these sites in some stream types. The 90 % confidence limits are used in biological assessments to determine the confidence in an IBI score as it relates to the biocriterion. For example, an IBI score that falls above the confidence limit is very likely to be attaining aquatic life use goals. In contrast, there is less certainty in an assessment when an IBI score falls within the confidence interval. In this case, a comprehensive assessment (e.g., review of available biological, chemical, habitat, and landscape measures) is more important to support a determination of attainment or non-attainment of the applicable biocriterion. 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.

Development of Minnesota's reference condition

Human disturbance score

To select the sites that are used to define the reference condition, a process for identifying sites in minimally disturbed watersheds is required (Hughes 1995). This includes using watershed and reach-scale attributes (e.g., urbanization, row crop agriculture, riparian condition, mines, feedlots, and channelization) that are positively or negatively associated with biological stressors. To accomplish this, the MPCA has developed an index to estimate the potential anthropogenic stress resulting from human activity upstream of and within the sampling site called the human disturbance score (HDS) (MPCA 2016). The HDS includes a series of primary and secondary measures of human activity at the watershed and sampling site scales. The HDS consists of eight primary metrics (Table 1). Before scoring, primary metrics were transformed to reduce skewness (Table 1). Each metric was scored proportionally by rescaling the transformed values on a scale of 0–10. The metrics percent impervious surface and number of point sources had values outside three times the interquartile range (i.e., outliers). For these two metrics, three times the interquartile range was used to set the upper and lower bounds for rescaling the transformed metric values. Secondary score adjustments were made for four additional measures (Table 1). The secondary score adjustments are submetrics for four of the primary metrics and were used to refine the watershed score. All four of the secondary score adjustments subtracted a point if a site's score was above the interquartile range for the metric. Additionally, a point was added if the number of road crossings per squared kilometer in a site's watershed was below the interquartile range for this metric. Score adjustments were also made for three factors were a single point was subtracted (Table 1) if a site was within or adjacent to an urban area, if it was adjacent to a

Table 1 Metrics used for the calculation of Minnesota's human disturbance (HDS) score

Human disturbance score metric	Scale	Primary metric or adjustment	Transformation	Maximum score
Number of animal units per km^2	Watershed	Primary	Log	10
Percent agricultural land use	Watershed	Primary	Arcsine	10
Number of point sources per km^2	Watershed	Primary	Square root	10
Percent impervious surface	Watershed	Primary	None	10
Percent channelized stream per stream km	Watershed	Primary	Square root	10
Degree channelized at site	Site	Primary	None	10
Percent disturbed riparian habitat in 15 and 30 m buffers	Watershed	Primary	Arcsine	10
Condition of riparian zone	Site	Primary	None	10
Number of feedlots per km^2	Watershed	Secondary	–	−1
Percent agricultural land use on >3 % slope	Watershed	Secondary	–	−1
Number of road crossings per km^2	watershed	secondary	–	−1 or +1
Percent agricultural land use in 100 m buffer	Watershed	Secondary	–	−1
Feedlot adjacent to site	Site (proximity)	Adjustment	–	−1
Point source adjacent to site	Site (proximity)	adjustment	–	−1
Urban land use adjacent to site	Site (proximity)	Adjustment	–	−1
		Maximum		81

feedlot or a feedlot was immediately upstream of the site (only streams <130 km^2), or if a continuously discharging point source was <8 km upstream of the site. The scores for each metric and the adjustments are summed resulting in a maximum score of 81. A higher HDS score indicates a site's watershed, and stream channel is more natural and therefore there is a lower potential for anthropogenic stress (Table 1).

Selection of reference condition sites

The HDS was used as the primary filter to generate datasets of least- and minimally-disturbed sites. Sites where HDS ≥61 (i.e., the upper 25 % of the HDS distribution) were first selected as potential RC sites then additional filters were used to remove sites with obvious stressor sources in close proximity to the sampling site. Sites were removed if the sample site was in close proximity to point sources, feedlots, or urbanization. Similarly, sites that were ≥50 % channelized as determined by a site visit, lidar, and aerial photography were also removed. The 50 % channelization threshold was selected to remove sites that lacked habitat to support natural channel biological assemblages. In general, streams with some channelization in these datasets had short channel alterations that were the result of road crossings and were not representative of the broad-scale channel condition. Regardless, 98–100 % sites in these datasets had channels that were 100 % natural. To test the effectiveness of the reference selection process, biological scores were compared between reference and non-reference sites using box plots. IBI scores from reference and non-reference site populations were compared for each stream type using a Mann–Whitney rank sum test in SigmaPlot ver. 12 (Systat Software 2011) because most datasets were not normally distributed.

Selection of reference sites for channelized streams

A set of "reference" channelized streams were required to develop a RC for channelized systems (i.e., modified RC). The term "reference" in this context is used to denote streams that have been channelized, but that have been managed in a manner that have promoted stream attributes that positively influence biological communities (e.g., sufficient riparian buffers and other best management practices [BMPs]). Only sampling sites that were ≥50 % channelized as determined from aerial photography, lidar, and site visits were used. The majority (93 %) of the channelized reference sites were channelized throughout the entire sampling reach. Two riparian buffer metrics from the HDS were used to select candidate modified RC sites. First, at the watershed scale, sites with <80 % of row crop in the upstream riparian habitat

(100 m) were selected. Second, sites with ≤80 % of the sampling site riparian (average of 15 and 30 m buffers) disturbed by human activity (e.g., crop, turf grass, roads, and active pasture) were selected. The 80 % disturbance threshold was selected for the sampling site riparian metric to approximate the permanent 5-m strips of perennial vegetation along ditches required by Minnesota's Drainage Law (Minnesota Statute 103E.021). The selection of this buffer width aligns the modified RC criteria with existing state requirements and selects sites with more stable banks and reduced erosion. In addition to these criteria, sites in close proximity to point sources, feedlots, or urbanization were excluded. To remove sites organically enriched, sites with dissolved oxygen concentrations <4 or >12 mg L^{-1} were excluded. Differences between the IBI scores for populations of channelized reference and non-reference sites were tested using a Mann–Whitney rank sum test run in SigmaPlot ver. 12 (Systat Software 2011).

Minnesota's biological condition gradient models

BCG models for Minnesota stream types were developed as part of two separate efforts for warm and cold water streams. Development of Minnesota's warm water BCG models involved participation from aquatic ecologists from MPCA and the Minnesota Department of Natural Resources (Gerritsen et al. 2013). Warm water fish and macroinvertebrate BCG models were developed for each of the seven warm water stream types. BCG models for cold water stream types were developed with aquatic ecologists from the upper midwest of the USA

Table 2 Taxa attributes used to characterize the BCG

Attribute	Description
I. Historically documented, sensitive, long-lived, or regionally endemic taxa	Taxa known to have been supported according to historical, museum, or archeological records, or taxa with restricted distribution (occurring only in a locale as opposed to a region), often due to unique life history requirements (e.g., blue sucker, American eel, crystal darter, *Goera*, *Apatania*).
II. Highly sensitive (typically uncommon) taxa	Taxa that are highly sensitive to pollution or anthropogenic disturbance. Tend to occur in low numbers, and many taxa are specialists for habitats and food type. These are the first to disappear with disturbance or pollution (e.g., brook trout, northern brook lamprey, weed shiner, *Stempellina, Ephemerella, Paraleuctra, Ophiogomphus, Parapsyche, Diplectrona, Lepidostoma, Dolophilodes, Rhyacophila*).
III. Intermediate sensitive and common taxa	Common taxa that are ubiquitous and abundant in relatively undisturbed conditions but are sensitive to anthropogenic disturbance/pollution. They have a broader range of tolerance than attribute II taxa and can be found at reduced density and richness in moderately disturbed sites (e.g., mottled sculpin, northern redbelly dace, stonecat, *Diamesa*, *Hexatoma*, *Plauditus*, *Isoperla*, *Boyeria*, *Amphinemura*, *Pycnopsyche*, *Brachycentrus*).
IV. Taxa of intermediate tolerance	Ubiquitous and common taxa that can be found under almost any conditions, from undisturbed to highly stressed sites. They are broadly tolerant but often decline under extreme conditions (e.g., blackside darter, spotfin shiner, tadpole madtom *Simulium*, *Tricorythodes*, *Dubiraphia*, *Orthocladius*, *Limonia*, *Perlesta*, *Heptagenia*, *Libellula*, *Hydropsyche*, *Planorbella*).
V. Tolerant taxa	Taxa that typically are uncommon and of low abundance in undisturbed conditions but that increase in abundance in disturbed sites. They are opportunistic species able to exploit resources in disturbed sites (e.g., creek chub, sand shiner, brook stickleback, *Polypedilum*, *Cricotopus, Helobdella*, *Trichocorixa, Pseudocloeon, Enallagma, Caecidotea, Physidae*).
Va. Highly tolerant taxa	Similar to attribute V, but these species especially thrive in the most disturbed sites and are the last species to survive at highly degraded sites (e.g., fathead minnow, green sunfish, central mudminnow).
VI. Sensitive nonnative or intentionally introduced species	Any species not native to the ecosystem that is generally sensitive to pollution or anthropogenic disturbance (e.g., brown trout, rainbow trout). Although these species may be sensitive their presence may still be disruptive to native taxa.
VIa. Tolerant nonnative or intentionally introduced species	Any species not native to the ecosystem that is generally tolerant to pollution or anthropogenic disturbance (e.g., common carp, goldfish, *Corbicula*, zebra mussel). These species are often an indicator of stress and are often disruptive to native taxa.

and included participants 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 (Gerritsen and Stamp 2013). This effort included two cold water stream types each for fish and macroinvertebrates.

Conceptually, the development of BCG models is similar, but requires regional calibration to account for differences in communities and their response to regionally-specific stressor gradients (Gerritsen and Stamp 2013). BCG models were developed for 18 stream types that paralleled the IBI model stream typology. The development of all these models required participation from aquatic ecologists knowledgeable with fish and macroinvertebrates in the region. The first step was to assign taxa to BCG attributes (Table 2) based on relationships of these taxa to the HDS and professional judgement by the panel of biologists (Gerritsen and Stamp 2013). The next step was for the panel to review fish and macroinvertebrate samples from stream sites and to place them into one of six BCG levels (Table 3). This was accomplished by reviewing

Table 3 Descriptions of biological condition gradient levels (Modified from Davies and Jackson 2006).

BCG Level	Description
Level 1: Natural or native condition.	Native structural, functional, and taxonomic integrity is preserved; ecosystem function is preserved within the range of natural variability. Level 1 represents biological conditions as they existed (or may still exist) in the absence of measurable effects of stressors.
Level 2: Minimal changes in structure of the biotic community and minimal changes in ecosystem function.	Virtually all native taxa are maintained with some changes in biomass and/or abundance; ecosystem functions are fully maintained within the range of natural variability. Level 2 represents the earliest changes in densities, species composition, and biomass that occur as a result of slight elevation in stressors (such as increased temperature regime or nutrient enrichment).
Level 3: Evident changes in structure of the biotic community and minimal changes in ecosystem function.	Evident changes in structure due to loss of some highly sensitive native taxa; shifts in relative abundance of taxa but sensitive-ubiquitous taxa are common and abundant; ecosystem functions are fully maintained through redundant attributes of the system. Level 3 represents readily observable changes that, for example, can occur in response to organic enrichment or increased temperature.
Level 4: Moderate changes in structure of the biotic community with minimal changes in ecosystem function.	Moderate changes in structure due to replacement of some intermediate-sensitive taxa by more tolerant taxa, but reproducing populations of some sensitive taxa are maintained; overall balanced distribution of all expected major groups; ecosystem functions largely maintained through redundant attributes.
Level 5: Major changes in structure of the biotic community and moderate changes in ecosystem function.	Sensitive taxa are markedly diminished; conspicuously unbalanced distribution of major groups from those expected; organism condition shows signs of physiological stress; ecosystem function shows reduced complexity and redundancy; increased build-up or export of unused materials. Changes in ecosystem function (as indicated by marked changes in food-web structure and guilds) are critical in distinguishing between levels 4 and 5.
Level 6: Severe changes in structure of the biotic community and major loss of ecosystem function.	Extreme changes in structure; wholesale changes in taxonomic composition; extreme alterations from normal densities and distributions; organism condition is often poor; ecosystem functions are severely altered. Level 6 systems are taxonomically depauperate (low diversity and/or reduced number of organisms) compared to the other levels.

worksheets that included the list of taxa, the abundance of each taxon, and the BCG attribute of each taxon in the biological sample. The panelists were not provided with water body names, geographic location, or HDS scores to avoid biases. The process included a discussion of each sample to determine which characteristics of the community (e.g., number of sensitive taxa, dominance of a single taxon, and proportion of tolerant individuals) were informing each panelist's BCG decision. Depending on the stream type, 29–75 samples were assigned to BCG levels based on the panel's majority decision during the calibration workshops. Using BCG assignments and the responses of the panel, fuzzy set models (Zadeh, 1965, 2008; Gerritsen and Stamp 2013) were developed to replicate the panel's decision for samples that had not been reviewed. To independently test the BCG models, the panel of biologists was reconvened to assign BCG levels to a set of samples not used in the calibration of the models. These samples were used to refine the final BCG models. The final model replicated the panel's majority decision BCG assignment for 76–98 % of samples with all mismatches within 1 level of the panel's majority decision (Gerritsen and Stamp 2013). BCG and IBI scores were compared for each stream type using a Spearman's Rank-Order Correlation in SigmaPlot ver. 12 (Systat Software 2011).

Sample size sufficiency analysis

Small datasets can result in thresholds that do not reflect the true population of stream sites. Ninety percent confidence intervals (CIs) were estimated for the quantiles used to calculate candidate thresholds (i.e., 25th, 50th, and 75th) from the IBI scores of RC and BCG datasets. The 90 % confidence intervals were estimated for the quantiles by generating 1000 random samples with replacement at five sample increments up to the total sample size using the "sample" and "quantile" functions in R (R Development Core Team 2015). The CIs were plotted against the sample size and fit with quantile regression splines at the 10th, 25th, 50th, 75th, and 90th percentiles in R (R Development Core Team 2015; rq in quantreg package [Koenker 2009]; and bs in splines package). The widths of the CIs and relationship between CI width and sample size were used to determine the dataset size sufficient to accurately estimate threshold statistics used for setting biocriteria.

Development of candidate biocriteria

Candidate biocriteria thresholds to protect tiered aquatic life use goals were derived from different BCG levels and contrasted with thresholds established using the RC approach. The 25th, 50th, and 75th percentiles 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 be difficult to estimate precisely (Berthouex and Hau 1991). Candidate RC- and BCG-derived thresholds were compared for each stream type to determine the relationships between these approaches and to identify protective biocriteria. The 90 % CI widths and sample sizes were also used to determine the confidence in the candidate biological thresholds.

General use

Biological experts involved with the development of the BCG have indicated that general use aquatic life goals (i.e., CWA Interim Goal) are represented somewhere within BCG levels 3 and 4 (Davies and Jackson 2006; Gerritsen et al. 2013). To develop candidate BCG thresholds, the 25th percentile of IBI scores from BCG level 3 samples and the 25th, 50th, and 75th percentile of IBI scores from BCG level 4 samples from natural channel sites were calculated for each stream type and contrasted with the 25th percentile of IBI scores for RC sites. The 25th percentile of IBI scores for the RC sites was selected as this statistic was used by Ohio to derive biocriteria for most ecoregions (Yoder and Rankin 1995).

Exceptional use

Based on the narrative descriptors of the BCG levels (Davies and Jackson 2006; Gerritsen et al. 2013), biological assemblages that fall within BCG Levels 1, 2, or 3 could be considered exceptional. Several statistics from the BCG and RC were calculated from natural channel sites to identify candidate thresholds for measuring attainment of an exceptional use. The 50th and 75th percentiles of IBI scores from BCG level 3 datasets and the 25th percentile of IBI scores from BCG level 2 datasets were calculated. No statistics were calculated for BCG level 1 as these samples were rare to absent in all stream types. The candidate thresholds based on the BCG were contrasted with the 75th percentile of IBI scores for RC sites. This statistic and the underlying approach are consistent with methods used by Ohio to calculate the exceptional use biocriteria (Yoder and Rankin 1995).

Modified use

Based on narrative descriptions from the BCG (Table 3; Davies and Jackson 2006; Gerritsen et al. 2013), modified use biocriteria could fall in the lower end of BCG level 4 or in BCG level 5. The 25th, 50th, and 75th percentiles of IBI scores from BCG level 5 datasets and the 25th percentile of BCG level 4 datasets were calculated and contrasted with the 25th percentile of IBI scores from the modified RC sites. The BCG datasets included both channelized and natural sample sites whereas the modified RC dataset included only channelized sample sites.

Results

Sample size sufficiency

Confidence intervals were wider and more variable below 25–35 samples (Fig. 1) indicating that at a minimum 25–35 samples should be used to estimate statistics used for developing biocriteria. There were differences between the CI widths for BCG and RC derived statistics with the BCG CI width estimates lower at smaller sample sizes and with more similarity in this estimate among stream types (Fig. 1).

BCG

The IBI and BCG scores for both assemblages were significantly correlated ($p<0.0001$) for all stream types (Figs. 2 and 3). For both assemblages and across stream types, there was a descending stepwise relationship between IBI and BCG scores with the 1st quartile of a BCG level often falling between the 2nd and 3rd quartile of the lower, adjacent BCG level. This pattern did not hold for BCG level 6 in some stream types and BCG level 5 for one macroinvertebrate stream type, but these datasets had small sample sizes. Not every stream type had samples representing all six BCG levels. BCG level 1 samples were absent or uncommon in southern stream types for both assemblages, and BCG level 2 samples were absent in southern macroinvertebrate stream types. In contrast, several northern stream types had few BCG level 5 or 6 samples (Figs. 2 and 3). For the datasets used to calculate candidate biocriteria, sample sizes across stream types for most BCG levels were ≥35 samples. Notable exceptions were BCG level 2 and BCG level 5 for several stream types (Tables 4 and 5).

Minimally- and least-disturbed reference condition datasets

The degree of anthropogenic stressors as estimated by the HDS differed between stream types (Figs. 4 and 5). Overall, large river and cold water stream types had smaller sample sizes which reflect the abundance of these stream types in Minnesota. Northern stream types tended to have lower predicted stressors levels (i.e., better HDS scores) compared to southern or plains stream types which tended to have higher predicted stressors with few sites with HDS scores ≥61 (Table 4, Fig. 6). Consequently, southern warm water reference datasets for both assemblages consisted of 25 or fewer sites (Table 4).

A comparison of IBI score distributions from populations of RC and non-RC sites provided an additional assessment of whether or not the method for selecting reference sites was effective. For most northern stream types, the 2nd quartile of IBI scores for RC sites was above or overlapped the 3rd quartile of non-RC sites. IBI scores for RC sites were greater than non-RC sites for most southern stream types, but the 2nd quartile for RC sites often overlapped the 2nd quartile of non-RC sites (Figs. 2 and 3). IBI scores from RC and non-RC populations did not differ significantly ($\alpha=0.05$) for three fish (three southern) and three macroinvertebrate (two southern and one northern) stream types (Figs. 2 and 3).

Modified reference datasets

Nine stream types had sample sizes of <35 sites and were not included in this analysis. For most stream types, the 2nd quartile of the modified RC sites overlapped the 3rd quartile of the non-RC sites. Modified RC sites statistically ($\alpha=0.05$) had higher IBI scores compared to modified non-RC sites for most stream types (Figs. 2 and 3)

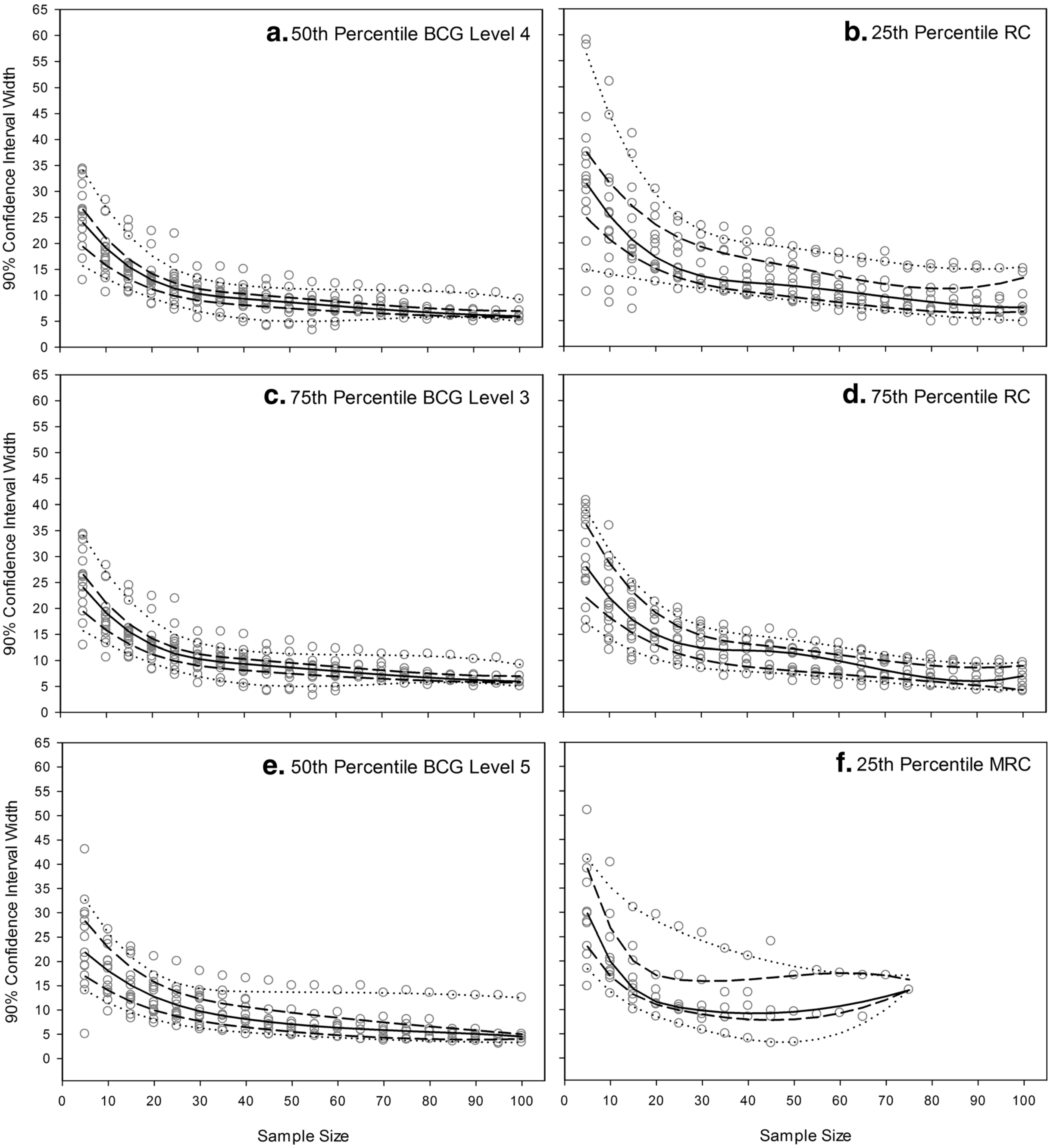

Fig. 1 Relationship between sample size and 90 % confidence interval width based on different sample sizes for randomly selected samples (with replacement). The fits are quantile LOESS fits at the 50th (*solid line*), 25th/75th (*dashed lines*), and 10th/90th (*dotted lines*) quantiles. Abbreviations: *BCG* biological condition gradient, *RC* reference condition, *MRC* modified reference condition

indicating that criteria used to distinguish between RC and non-RC channelized streams were effective. Significant differences were present between modified RC and non-RC samples for eight of the nine stream types. The macroinvertebrate high gradient southern streams type was the only stream type where IBI scores for modified RC and non-RC sites were not significantly different, most likely because of the low power of the test associated with a small sample size.

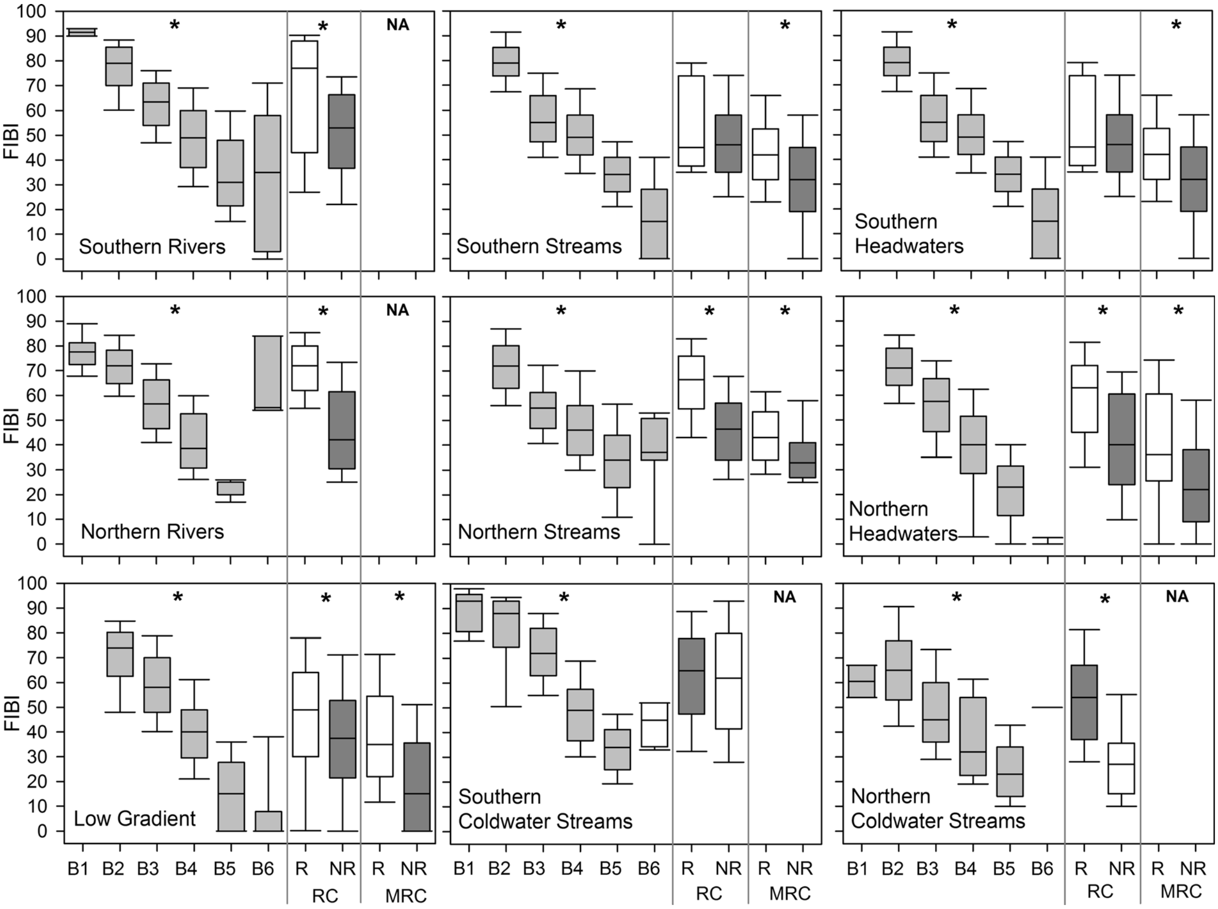

Fig. 2 Frequency distributions of scores for fish indices of biotic integrity from datasets used to develop candidate biological criteria for the protection of aquatic life use goals. Symbols: *upper and lower bounds of box* indicates 75th and 25th percentiles, *middle bar in box* indicates 50th percentile, *upper and lower whisker caps* indicates 90th and 10th percentiles, *asterisk* indicates significant difference at the $\alpha = 0.05$ level based on Spearman's Rank-Order Correlation (biological condition gradient datasets) or a Mann–Whitney rank sum test (reference condition datasets). Abbreviations: *FIBI* fish index of biotic integrity, *B1 through B6* biological condition gradient levels 1 through 6, *R* reference, *NR* non-reference, *RC* reference condition, *MRC* modified reference condition, *NA* not assessed

Comparison of BCG and RC derived thresholds

General use

The general use RC threshold (i.e., 25th percentile of reference sites) was most similar to the median of BCG level 4. BCG level 4 medians for nine of the 18 stream types were higher than the RC threshold, and there were only two stream types where this difference was >10 points. Of the nine stream types where the median of BCG level 4 was lower, eight had a difference of ≤8 points. The remaining stream type, Northern Rivers (fish), had a median of BCG level 4 that was 25 points less than the RC threshold. In contrast, the 25th percentile of BCG level 4 was lower than the RC threshold for 16 of 18 stream types and thresholds for ten stream types were at least 10 points lower. The majority of thresholds for the 75th percentile of BCG level 4 and the 25th percentile of BCG level 3 were higher than the RC threshold (12 and 16 stream types, respectively) with differences of >10 IBI points common (Table 6). Three stream types for BCG level 4 datasets had small sample sizes ($n < 35$) including Northern Rivers and Northern Cold Water fish stream types and Northern Forest Rivers for macroinvertebrates. Two of these stream types (i.e., Northern Cold Water fish and Northern Forest Rivers macroinvertebrate stream types) had median BCG level 4 values within 2–3 points of the RC threshold, indicating this threshold was appropriate

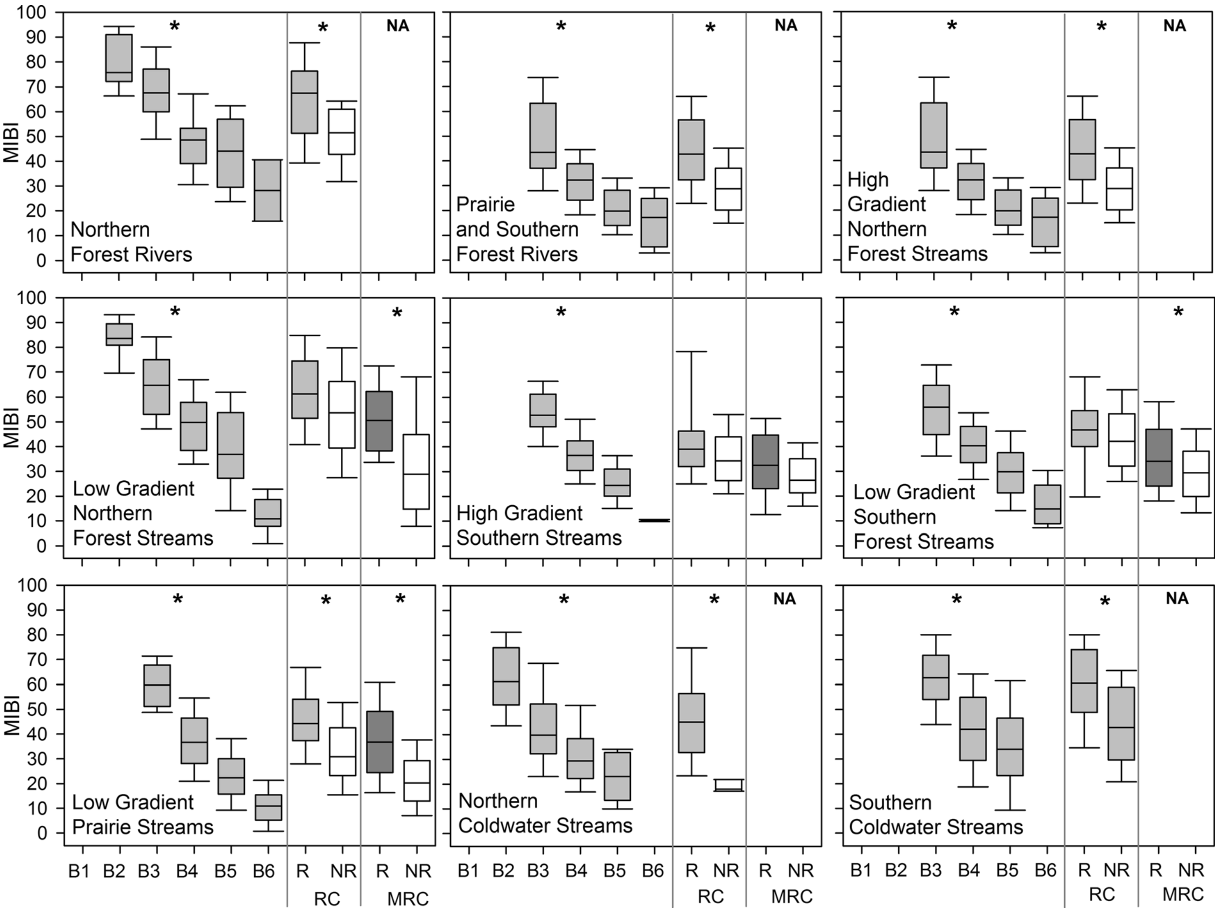

Fig. 3 Frequency distributions of scores for macroinvertebrate indices of biotic integrity from datasets used to develop candidate biological criteria for the protection of aquatic life use goals. Symbols: *upper and lower bounds of box* indicates 75th and 25th percentiles, *middle bar in box* indicates 50th percentile, *upper and lower whisker caps* indicates 90th and 10th percentiles, *asterisk* indicates significant difference at the $\alpha = 0.05$ level based on Spearman's Rank-Order Correlation (biological condition gradient datasets) or a Mann–Whitney rank sum test (reference condition datasets). Abbreviations: *MIBI* macroinvertebrate index of biotic integrity, *B1 through B6* biological condition gradient levels 1 through 6, *R* reference, *NR* non-reference, *RC* reference condition, *MRC* modified reference condition, *NA* not assessed

despite a smaller sample size (Table 4; $n = 21$–28). The RC threshold for Northern Rivers for fish was most similar to the 75th percentile of BCG level 3 or the 25th percentile of BCG level 2 (Fig. 2).

Exceptional use

The 75th percentile of BCG level 3 was most similar to the exceptional RC threshold (i.e., 75th percentile of reference sites; Table 7). Eleven stream types had a difference between the 75th percentile of BCG level 3 and the exceptional RC threshold of ≤8 points. Four fish and three macroinvertebrate stream types had a difference of ≥10 points between the BCG and RC statistics (Table 7). For nine stream types, the 75th percentile of BCG level 3 was higher than the RC threshold. The RC threshold was ≥10 points higher for two northern fish stream types and one southern fish stream type. The RC threshold (IBI = 80) for the Northern Rivers fish stream type was nearly equivalent to the 75th percentile of BCG level 1 (Fig. 2), indicating that most BCG level 1 sites would not meet this threshold. Similarly, the RC threshold (IBI score of 76) for the Northern Streams fish stream type is between the 50th and 75th percentiles of BCG level 2 (Fig. 2) indicating that the majority of BCG level 2 stream sites in this stream type would not meet this threshold. In addition, the RC threshold for the Southern Rivers fish stream type was equivalent to the 90th percentile of BCG level 2 (Fig. 2). The RC dataset for the Southern Rivers for fish included data from only

Table 4 Sample sizes for datasets used to develop general and exceptional use candidate biocriteria

Stream type	Reference condition	BCG4	BCG3	BCG2
Fish				
Southern rivers	18	82	61	17
Southern streams	8	74	102	22
Southern headwaters	15	183	49	7
Northern rivers	116	28	47	106
Northern streams	186	155	54	149
Northern headwaters	215	37	127	120
Low gradient streams	111	67	52	17
Southern cold waters	61	43	101	14
Northern cold waters	196	21	118	71
Macroinvertebrates				
Northern forest rivers	83	28	56	15
Prairie and southern forest rivers	9	88	28	0
Northern forest streams-high gradient	162	41	159	11
Northern forest streams-low gradient	210	71	135	14
Southern streams-high gradient	15	225	57	0
Southern forest streams-low gradient	25	88	57	0
Prairie streams-low gradient	13	91	12	0
Northern cold waters	185	63	56	55
Southern cold waters	60	118	73	0

Biological condition gradient (BCG) and reference condition datasets include only sites with natural channels. For BCG datasets, the number following "BCG" corresponds to the BCG level to which the sites in that dataset belong (i.e., BCG4 = BCG level 4 sites)

three rivers which raises concern regarding the applicability of criteria derived from this dataset to the entire population.

The 50th percentile of BCG level 3 for 14 stream types was less than the 75th percentile of the RC. The 25th percentile of BCG level 2 could not be calculated for five macroinvertebrate stream types because of the lack of BCG level 2 samples, and only five stream types for both fish and macroinvertebrates had BCG level 2 sample sizes ≥35 (Table 4). Of the 13 stream types with BCG level 2 samples, the 25th percentile of BCG level 2 was less than the RC threshold for ten stream types indicating that the BCG level 2 threshold was similar to the 75th percentile of BCG level 3.

Modified use

Modified RC datasets had only six stream types with ≥35 samples (Table 5). Nine stream types had <15 samples, the large river and cold water stream types for both assemblages and the high gradient northern forest streams for macroinvertebrates. These low sample sizes are a reflection of small numbers of channelized streams in these stream types and make calculation of modified use biocriteria technically difficult. Furthermore, the scarcity of channelized streams in these groups indicated that a modified use is not needed, nor appropriate. As a result, these nine stream types are not included in further analysis of modified use biocriteria.

The median of BCG level 5 was similar to the modified RC with the difference between these statistics ranging from 1 to 7 points (Table 8). In contrast, the 25th percentile of IBI scores for BCG level 5 was less than the modified RC for all stream types (Table 8). Four stream types had a difference of >10 points including two that had a difference of >20 points. For all stream types, the 75th percentile of BCG level 5 and the 25th percentile of BCG level 4 were greater than the modified RC thresholds. These included 2–4 stream types with >10 point difference between these statistics.

Discussion

Development of the reference condition and modified reference condition

By using measures of potential anthropogenic disturbance at local and watershed scales, the HDS adequately distinguished reference from non-reference sites. Among northern and coldwater stream types, the RC approach identified least- and minimally-disturbed conditions which could be used to support development of biocriteria. However, southern stream classes were characterized by relatively high levels of anthropogenic disturbance, which allowed for the identification of least-disturbed conditions at best. The greater predicted stress (Figs. 4 and 5) and the observed lack of significant differences in IBI scores between RC and non-RC sites (Figs. 2 and 3) in most of the southern stream types are the result of the uniformity and poor overall condition of streams in these groups. Minimal differences in IBI scores between RC and non-RC sites and small sample

Table 5 Sample sizes for datasets used to develop modified use candidate biocriteria

Stream type	Modified reference	BCG5	BCG4
Fish			
Southern rivers	12	41	95
Southern streams	33	214	114
Southern headwaters	19	100	371
Northern rivers	3	7	30
Northern streams	53	78	201
Northern headwaters	77	217	73
Low gradient streams	41	83	114
Southern cold waters	1	62	50
Northern cold waters	6	35	25
Macroinvertebrates			
Northern forest rivers	0	6	25
Prairie and southern forest rivers	9	25	91
Northern forest streams-high gradient	3	3	42
Northern forest streams-low gradient	49	43	99
Southern streams-high gradient	18	132	217
Southern forest streams-low gradient	39	157	112
Prairie streams-low gradient	68	284	183
Northern cold waters	7	10	59
Southern cold waters	1	35	118

Biological condition gradient (BCG) datasets included both natural channel and channelized stream sites and modified reference condition datasets include only channelized streams

sizes for southern stream types indicated that the RC approach alone may not be appropriate in this region.

Although many stream types were characterized by low numbers of modified RC sites, the site selection process was effective for identifying better performing sites within the population (Figs. 2 and 3). However, some stream types had few modified RC sites because either very few channelized streams met modified RC criteria or because channelized streams were uncommon within the stream type (e.g., large rivers and cold water streams). In these stream types, the development of modified use biocriteria may not be necessary or would require additional information to improve the confidence of the thresholds derived from these data. For example, the BCG could be used to provide consistency among dissimilar regions and/or stream types by identifying where each reference site distribution fell along an ecologically-based gradient of disturbance.

Development of biocriteria

In heterogeneous regions, stream classification types are needed to minimize the effects of natural variation. This requires sufficient environmental monitoring data to develop the typology and to calculate precise statistics from these populations. Many water body typologies used to develop biological monitoring tools are based on broad-scale geographic classifications (e.g., watershed and ecoregion). However, the classification strength of such frameworks may be weak and could introduce errors in biological assessments (Hawkins et al. 2000b; Van Sickle and Hughes 2000; Waite et al. 2000). The typology approach used in Minnesota uses both broad-scale (i.e., geographic location) and reach-scale (i.e., gradient, presence or rocky riffles/runs, thermal regime, and stream size) measures to determine the placement of a sampling site into a stream type. The inclusion of both broad-scale and reach-scale determinants of stream type minimizes natural differences between sites with a stream type and sets appropriate expectations for biological assemblage composition (Hawkins et al. 2000b).

The implementation of a water body typology in a biological assessment framework can be either discrete or continuous. Advantages of a continuous typology are that it recognizes the gradation between water body

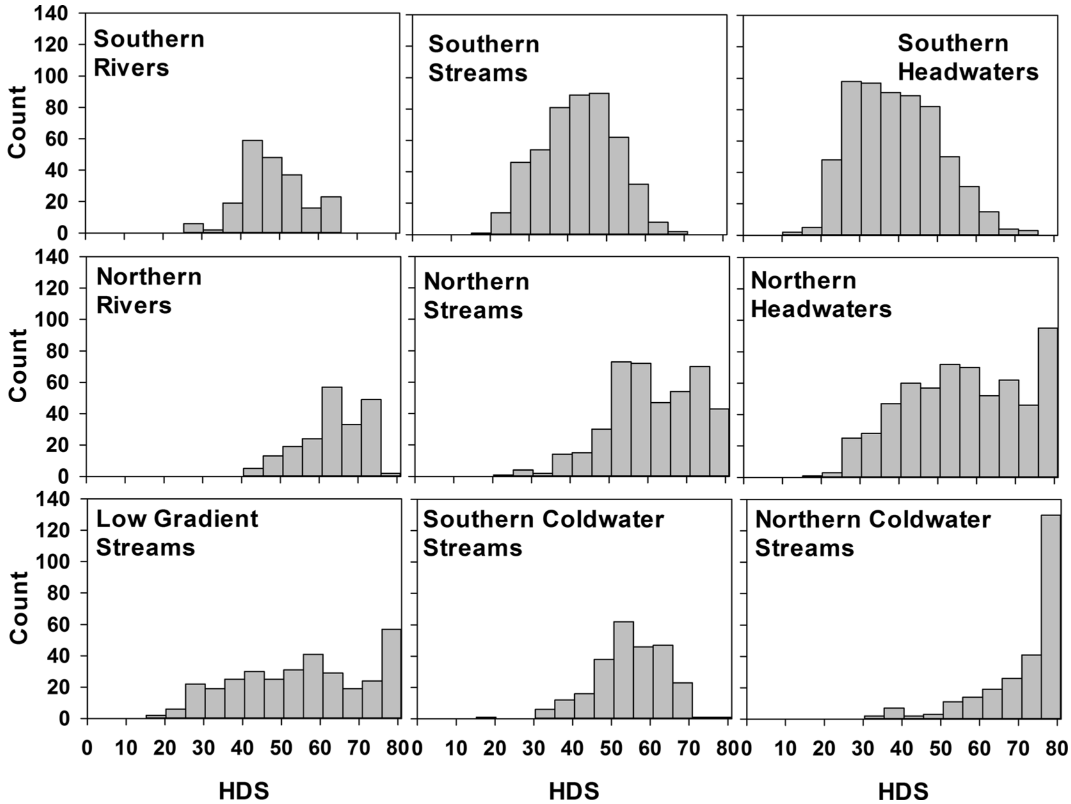

Fig. 4 Histograms of human disturbance scores (HDS) for fish sites by stream type

types and can be more easily applied over larger geographic regions. The latter advantage can be further used to develop an assessment framework that can be applied across watersheds and political boundaries thereby standardizing assessment outcomes (Pont et al. 2009). Advantages of a discrete typology are that the assessment models can be more specific to individual water body types and may be more sensitive subtle anthropogenic impacts (Pont et al. 2009). For example, rather than relying on a core set of metrics or biological measures, IBIs developed for each stream type can consist of a unique suite of metrics that are demonstrated to be effective measures of condition for those specific biological communities. A disadvantage of a discrete typology is that larger datasets may be needed to support the model development for each stream type. In addition, comparison of biological condition across types may be difficult in heterogeneous regions and could result in thresholds among stream types that vary greatly along a biological gradient of disturbance and have little relationship to the CWA Interim Goal. However, biological goals across stream types, watersheds, and political boundaries can be standardized using BCG models.

The stream typology is largely influenced by the physical features of streams and their watersheds, and these physical factors are also associated with prevailing land use patterns. In Minnesota, stream types exhibited different forms and degrees of disturbance and the number of sites and overall condition of the reference datasets were diminished for some stream types where anthropogenic stressors were ubiquitous. For example, reference sites in the heavy agricultural plains regions (i.e., southern and western) of Minnesota (Fig. 6) were of relatively poorer quality than reference sites in the more forested, northeastern region of the state. Consequently, reference sites in the plains regions were more degraded and few if any could be considered "least disturbed". The opposite was true for the northern stream types where disturbance levels were generally low, and the reference condition was near natural conditions.

Our analysis indicated that the sample size needed to calculate sufficiently precise statistics for each stream type ranged from 25 to 35 samples. This is similar to the maximum estimate of 35–40 samples per ecoregion provided by Yoder and Rankin (1995). However, most

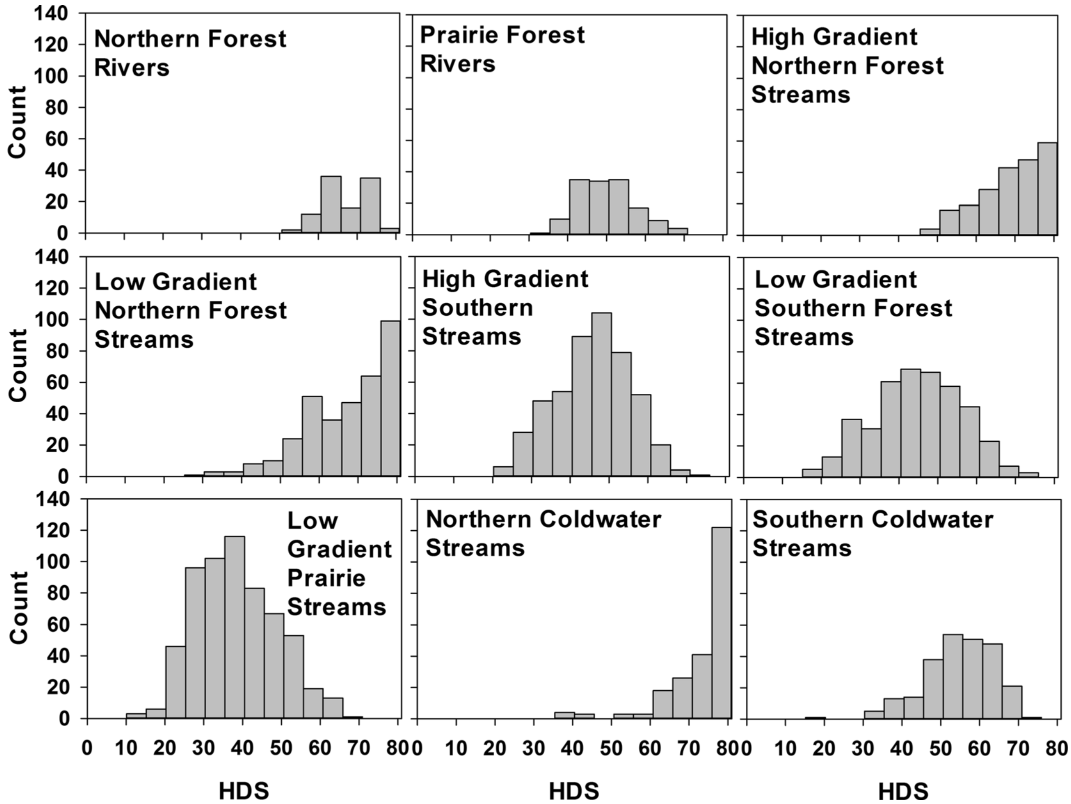

Fig. 5 Histograms of human disturbance scores (HDS) for macroinvertebrate sites by stream type

southern stream types had <35 reference sites, and the overall condition of these reference sites was less than that of the northern classes. This indicated that for these stream types, the RC approach alone was insufficient to set protective goals or to ensure consistency among stream types. In heavily impacted southern streams, a higher percentile of the reference sites or alternatively an "all sites" approach (e.g., 90th percentile of all sites; Yoder and Rankin 1995) could have been used to select biocriteria protective of aquatic life use goals. However, Minnesota chose not to focus solely on a RC approach to avoid making an a priori judgment that a known percentage of sites are impaired for aquatic life. Instead, the BCG was used to anchor the biocriteria in each stream type along a known ecological endpoint and to provide both consistency among stream types and alignment with Minnesota's aquatic life use goal narratives. Stream types with large reference site sample sizes (i.e., northern stream types) established the association between the BCG and the RC. From this analysis, the protective BCG-level threshold could be applied to stream types that lacked robust RC datasets. The final proposed biocriteria for all stream types were based on statistics calculated from the BCG to maintain uniformity in biological expectations across stream types and tiered uses. Despite challenges with using the RC for some stream types, the BCG and RC approaches generated similar candidate biocriteria for many stream types thereby providing greater confidence in the proposed biocriteria.

Aligning the BCG to Minnesota's water quality rules

The general use is the minimum threshold for attainment and it is the principle restoration goal when it is determined that a water body is in non-attainment (MPCA 2014a). Therefore, it is a trigger for management actions such as use attainability analyses (UAAs) and total maximum daily loads (TMDLs). The BCG level 4 narrative is pertinent to Minnesota's general use goal for aquatic life. The ecosystem in BCG level 4 water bodies have a structure and function that is largely intact although the species composition may be moderately altered from the natural condition (MPCA 2014a). BCG level 4 assemblages are described as having an "overall balanced distribution of all expected major

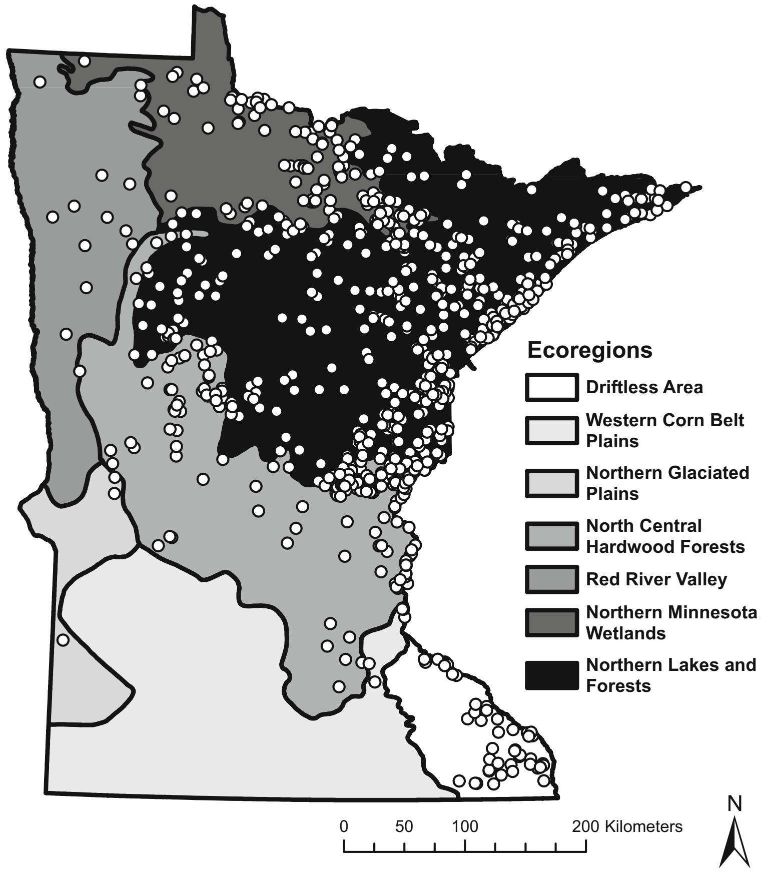

Fig. 6 Location of least- and minimally-disturbed reference condition stream sites showing level 3 ecoregions (White and Omernik 2007)

groups" with "ecosystem functions largely maintained through redundant attributes" (Davies and Jackson 2006; see Table 3). These assemblages retain their function in the ecosystem through redundancy in species composition. In other words, some intolerant or sensitive species may be replaced by more tolerant species that occupy similar ecological roles. Narrative language in Minnesota rule describing compliance with aquatic life use goals states that "…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…" (Minnesota Rule chapter 7050.0150 subpart 3). Based on conceptual similarities between BCG level 4 and language for the protection of aquatic life in Minnesota rule, many BCG level 4 assemblages are therefore consistent with Minnesota's narrative aquatic life use goals and the CWA Interim Goal.

Exceptional use water bodies support biological assemblages with community structure and ecosystem functions that are minimally altered from the natural condition. It is a preservation use, meaning that it applies to waters that already demonstrate an exceptional condition as determined by measurement of the biology. The narrative for BCG level 2 matches goals for these assemblages: "Structure & function similar to natural community with some additional taxa and biomass; ecosystem level functions are fully maintained" (Davies and Jackson 2006; Table 3). BCG level 3 may also apply to exceptional uses: "Evident changes in structure due to loss of some rare native taxa; shifts in relative abundance; ecosystem level functions fully maintained" (Davies and Jackson 2006; Table 3). BCG levels 2 and 3 are similar in that ecosystem functions (e.g., primary and secondary production, respiration, nutrient cycling, and decomposition) are intact, but differ in that some sensitive native species (BCG attribute levels I, II, and III; Table 3) may be extirpated at BCG level 3 while all native species are present at BCG level 2 (MPCA 2014a).

In Minnesota, many thousands of kilometers of rivers and streams have been channelized and straightened to

Table 6 Candidate general use biocriteria with a comparison of biological condition gradient (BCG) and reference condition derived criteria

Stream type	Reference condition	BCG4 25th %ile	BCG4 50th %ile	BCG4 75th %ile	BCG3 25th %ile
Fish					
Southern rivers	43	37 (−6)	49 (6)	61 (18)	54 (11)
Southern streams	38	41 (4)	50 (13)	58 (21)	47 (10)
Southern headwaters	46	45 (−1)	55 (9)	62 (16)	64 (18)
Northern rivers	62	30 (−32)	38 (−25)	52 (−11)	46 (−16)
Northern streams	55	37 (−18)	47 (−8)	59 (4)	47 (−8)
Northern headwaters	45	31 (−14)	42 (−3)	56 (11)	48 (3)
Low gradient streams	30	31 (1)	42 (12)	54 (24)	49 (19)
Southern cold water	48	37 (−11)	50 (3)	59 (12)	63 (16)
Northern Cold Water	37	26 (−12)	35 (−2)	55 (18)	37 (0)
Macroinvertebrates					
Northern Forest Rivers	51	39 (−12)	49 (−3)	53 (2)	60 (9)
Prairie and Southern Forest Rivers	32	24 (−8)	31 (−1)	38 (6)	37 (5)
Northern Forest Streams-High Gradient	58	43 (−14)	53 (−5)	58 (0)	58 (0)
Northern forest streams-low gradient	52	39 (−12)	51 (0)	58 (7)	56 (4)
Southern streams-high gradient	32	31 (−1)	37 (5)	44 (11)	49 (17)
Southern forest streams-low gradient	40	38 (−3)	43 (3)	50 (10)	48 (8)
Prairie streams-low gradient	37	33 (−5)	41 (4)	47 (10)	58 (21)
Northern cold waters	33	22 (−11)	32 (−1)	42 (9)	32 (−1)
Southern cold waters	49	30 (−19)	43 (−6)	56 (7)	54 (5)

Numbers in *parentheses* are the difference between BCG criteria and reference condition criteria

%ile percentile

promote agricultural drainage under Minnesota Drainage Law (Minnesota Statute 103E). The biological assemblages in these water bodies are limited by poor physical habitat structure that precludes support of the baseline CWA goal use. As such some cannot meet Minnesota's general use goals for aquatic life. The negative effects of degraded physical habitat structure on biological assemblages are well documented (Gorman and Karr 1978; Karr et al. 1986; Schlosser 1987; Rankin 1995; Paulsen et al. 2008; Schinegger et al. 2012). Despite limitations to biological assemblages associated with poor physical habitat structure in channelized streams, these watersheds can be managed (e.g., buffer maintenance and BMPs) such that these water bodies can meet some goal below general use and not regarded as incapable of supporting aquatic life or providing benefits other than drainage (MPCA 2014a).

The narrative for BCG Level 5 is "Sensitive taxa markedly diminished; conspicuously unbalanced distribution of major taxonomic groups; ecosystem function shows reduced complexity and redundancy" (Davies and Jackson 2006; Table 3). Although biological assemblages with distinctly reduced taxonomic complexity and ecosystem function are not consistent with general use goals, it describes the consequence of physically altered streams. Biological assemblages in channelized streams are often comprised of a greater proportion of tolerant species dominated by omnivores and generalists and may have increased biomass because of greater productivity (Yoder and Rankin 1995). In addition, to supporting a lower proportion of intolerant and sensitive species, channelized streams also often have altered functions (e.g., nutrient assimilation, primary and secondary production, and sediment transport) compared to unaltered streams (Yarbro et al. 1984). Despite the limitations imposed by poor physical habitat structure, physically altered streams can partially support beneficial aquatic life and the goals for these water bodies should be consistent with what is attainable with appropriate landscape and riparian management. This can also

Table 7 Candidate exceptional use biocriteria with a comparison of biological condition gradient (BCG) and reference condition derived thresholds

Stream type	Reference condition	BCG3 50th %ile	BCG3 75th %ile	BCG2 25th %ile
Fish				
Southern rivers	88	64 (−24)	71 (−17)	70 (−18)
Southern streams	74	55 (−19)	66 (−8)	74 (0)
Southern headwaters	62	69 (7)	74 (12)	73 (11)
Northern rivers	80	57 (−23)	67 (−13)	65 (−15)
Northern streams	76	55 (−21)	61 (−15)	63 (−13)
Northern headwaters	72	60 (−12)	68 (−4)	64 (−8)
Low gradient streams	64	58 (−6)	70 (6)	59 (−6)
Southern cold waters	78	72 (−6)	82 (4)	75 (−4)
Northern cold waters	67	47 (−21)	60 (−7)	54 (−13)
Macroinvertebrates				
Northern forest rivers	76	68 (−9)	77 (1)	72 (−4)
Prairie and southern forest rivers	57	44 (−13)	63 (7)	–
Northern forest streams-high gradient	83	70 (−13)	82 (−1)	81 (−2)
Northern forest streams-low gradient	74	67 (−8)	76 (2)	81 (6)
Southern streams-high gradient	46	54 (8)	62 (16)	–
Southern forest streams-low gradient	55	58 (3)	66 (11)	–
Prairie streams-low gradient	54	61 (7)	69 (15)	–
Northern cold waters	57	40 (−17)	52 (−4)	52 (−5)
Southern cold waters	74	63 (−11)	72 (−2)	–

Numbers in *parentheses* are the difference between BCG criteria and reference condition thresholds

%ile percentile

be interpreted to mean that modified use streams with biological assemblages falling into the lower end of BCG level 5 or BCG level 6 would not be in attainment of aquatic life use goals for a modified use.

Contrasting BCG and RC derived biocriteria

A comparison of the BCG approach with the more traditional RC approach indicated that the two methods are generally complementary. For the general use, the median of BCG level 4 was most similar to the RC (Table 6, Fig. 7a–b). Because the median of BCG level 4 was generally equivalent to the 25th percentile (Fig. 7a, b) of the RC and had narrower 90 % CI widths (i.e., greater statistical precision; see Fig. 1a, b), it was determined to be a better statistic for defining aquatic life use goals. As a result, streams in regions with overall heavy disturbance and a lack of reference analogs will not have lower restoration goals than streams in regions with overall higher condition. Although empirically driven by the comparison between the BCG and RC, the use of the median of BCG level 4 recognizes that regional biologists participating in BCG development indicated that Minnesota's general use goal was somewhere in BCG level 4 (Fig. 8). The use of the median of BCG level 4 also protects against variability within the BCG level and provides a degree of safety).

A limitation with the RC approach was apparent in the Northern River fish stream type where the threshold was 25 points greater than the BCG level 4 threshold. The distribution of sites used to determine the RC in northern rivers was heavily skewed towards BCG level 2 and 3 conditions as indicated by a preponderance of sites in BCG levels 2 and 3 and relatively good HDS scores. A biocriterion based on the 25 % of RC would have placed aquatic life use goals for Northern Rivers at the 67th percentile of BCG level 3 and would require that these rivers meet much higher goals than other streams in the state. The RC

Table 8 Candidate modified use biocriteria with a comparison of biological condition gradient (BCG) and modified reference condition (MRC) derived thresholds

Stream Type	MRC 25th %ile	BCG5 25th %ile	BCG5 50th %ile	BCG5 75th %ile	BCG4 25th %ile
Fish					
Southern rivers	50	21 (−29)	28 (−22)	44 (−7)	37 (−13)
Southern streams	32	27 (−5)	35 (3)	41 (9)	42 (10)
Southern headwaters	38	18 (−21)	33 (−6)	49 (11)	41 (3)
Northern rivers	47	20 (−27)	25 (−22)	25 (−22)	31 (−16)
Northern streams	34	26 (−8)	35 (1)	45 (11)	36 (2)
Northern headwaters	26	12 (−14)	23 (−3)	32 (6)	29 (3)
Low gradient streams	22	0 (−22)	15 (−7)	28 (6)	29 (7)
Southern cold waters	–	25 (−)	34 (−)	41 (−)	37 (−)
Northern cold waters	26	14 (−12)	23 (−3)	34 (8)	23 (−4)
Macroinvertebrates					
Northern forest rivers	–	29 (−)	44 (−)	57 (−)	39 (−)
Prairie and southern forest rivers	18	14 (−4)	21 (3)	28 (10)	24 (6)
Northern forest streams-high gradient	29	21 (−7)	51 (22)	61 (33)	45 (16)
Northern forest streams-low gradient	38	27 (−12)	37 (−1)	54 (16)	39 (0)
Southern streams-high gradient	23	20 (−3)	25 (1)	31 (8)	30 (7)
Southern forest streams-low gradient	24	22 (−2)	30 (6)	38 (14)	34 (10)
Prairie streams-low gradient	24	16 (−9)	22 (−2)	30 (6)	28 (4)
Northern cold waters	14	13 (0)	23 (9)	33 (19)	22 (8)
Southern cold waters	–	23 (−)	34 (−)	47 (−)	29 (−)

Numbers in *parentheses* are the difference between BCG criteria and MRC criteria

%ile percentile

thresholds were determined to not be appropriate for this stream type because many streams near natural condition would not meet these goals. An alternative approach could use different percentiles of the reference site dataset (e.g., 5th or 10th percentile) to recognize the high condition of these waters and the low proportion of poorly performing reference sites.

For the exceptional use, a comparison between candidate biocriteria developed using the RC and those using statistics from the BCG demonstrated that the 75th percentile of BCG level 3 was most similar to the 75th percentile of the RC (Table 7; Fig. 7c, d). The available example for developing biocriteria for an exceptional use in Minnesota is the state of Ohio. In Ohio, biocriteria for exceptional use waters were calculated as the 75th percentile of IBI scores from all reference sites using a statewide database (Yoder and Rankin 1995). This approach was not possible in Minnesota because IBI models for each stream type were calibrated using independent datasets of sites with different levels of overall disturbance. Because IBI scores were not equivalent among stream types, it necessitated the independent development of exceptional use biocriteria for each stream type. A limitation of this approach is that for stream types where conditions are predominantly poor there may be insufficient minimally-disturbed sites to determine exceptional use goals. However, the BCG models provided a tool to establish exceptional use biocriteria for stream types with small RC site sample sizes (Fig. 8).

For the modified use, the candidate biocriteria calculated from the 50th percentile of BCG level 5 were most comparable to the values calculated from the modified RC (Table 8; Fig. 7e, f). BCG level 5 datasets also had larger sample sizes than most of the modified RC datasets (Table 5). Greater sample sizes and the datasets themselves resulted in greater precision in estimation of

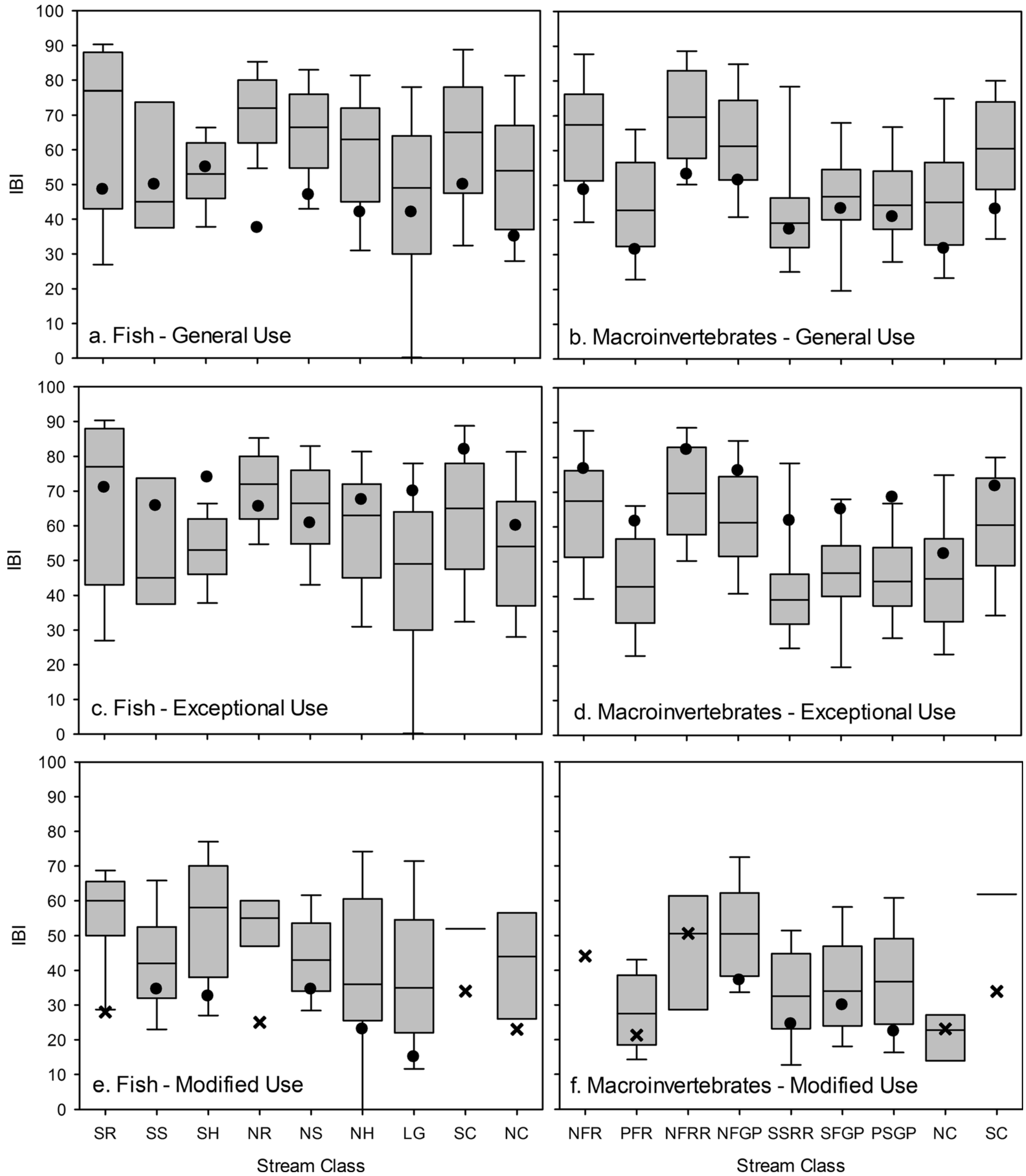

Fig. 7 Frequency distribution of fish (**a**, **c**, **e**) and macroinvertebrate (**b**, **d**, **f**) indicies of biotic integrity scores from reference condition and modified reference condition sites by stream type for the general use (**a**, **b**), exceptional use (**c**, **d**), and modified use (**e**, **f**). Symbols: black circle indicates biocriterion; *letter X* indicates unused threshold; *upper and lower bounds of box* indicates 75th and 25th percentiles; *middle bar in box* indicates 50th percentile; *upper and lower whisker caps* indicates 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 cold waters, *NC* northern cold waters, *NFR* northern forest rivers, *PFR* prairie and southern 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

these thresholds (Fig. 1). The biocriteria for modified use streams do not permit an extreme loss of the function, diversity, or biomass in biological assemblages (Fig. 8). Under a TALU framework, these altered systems will be held to an attainable goal that recognizes the effects of legal habitat alteration, but establishes

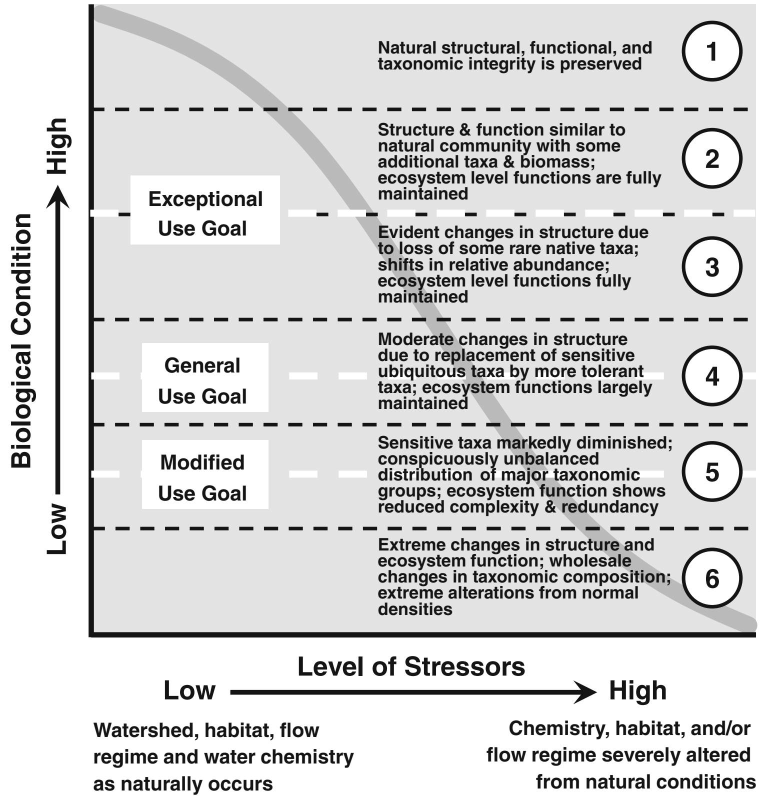

Fig. 8 Conceptual model of the biological condition gradient (BCG) illustrating the location of draft biocriteria for Minnesota's tiered aquatic life uses (adapted from Davies and Jackson 2006)

biological goals that are attainable with appropriate water quality management practices. The CWA permits the removal of a designated use if it is not an existing use (i.e., attained on or after November 28, 1975) and if it can be demonstrated that attainment of that designated use is not feasible (40 CFR Part 131). Reasons for removing a designated use include natural conditions such as naturally occurring pollutants and limiting flow conditions or physical conditions (e.g., insufficient cover, substrate, and depth). Human-caused reasons for removing a designated use include human-caused pollution that cannot be remedied or would cause more harm to correct or hydrologic modifications that cannot be restored or operated in a manner that would restore the designated use. To demonstrate that one more of these reasons apply to a stream, the CWA requires that a UAA is performed. Designation of a modified use therefore requires that a UAA be performed to demonstrate that the biological assemblage cannot be feasibly restored to general use goals and that physical habitat structure is limiting attainment of the minimum restoration goal (Yoder and Rankin 1995). In addition, water bodies with physical habitat structure that is limiting aquatic life, an evaluation is necessary to assess whether or not the state of the habitat is the consequence of legal modifications to the channel and that the habitat is not restorable (MPCA 2015). When these requirements are met, the water body could qualify for a modified use designation. The designated uses in modified use waters will be continually reviewed to determine if a ditch should be designated general use because the biological assemblages meet the general use or it supports habitat that indicates the general use is attainable. It is also important to note that over time, practices for maintaining ditches are likely to change which should lead to ditches with better physical habitat structure and better biological condition. As a result of shifting conditions in these water bodies, it will be necessary to periodically review the modified use biocriteria. These improved management practices should result in an increase in the number of ditches that either meet general use goals or are identified as restorable to the general use.

Table 9 Draft biocriteria for exceptional, general, and modified uses for Minnesota streams

Stream type	Exceptional use	General use	Modified use
Fish			
Southern rivers	71	49	NA
Southern streams	66	50	35
Southern headwaters	74	55	33
Northern rivers	67	38	NA
Northern Streams	61	47	35
Northern headwaters	68	42	23
Low gradient streams	70	42	15
Southern cold waters	82	50	NA
Northern cold waters	60	35	NA
Macroinvertebrates			
Northern forest rivers	77	49	NA
Prairie and southern forest rivers	63	31	NA
Northern forest streams-high gradient	82	53	NA
Northern forest streams-low gradient	76	51	37
Southern streams-high gradient	62	37	24
Southern forest streams-low gradient	66	43	30
Prairie streams-low gradient	69	41	22
Northern cold waters	52	32	NA
Southern cold waters	72	43	NA

Conclusion

Minnesota's proposed TALU framework includes three tiers for the protection of aquatic life: general, exceptional, and modified. To assess attainment of aquatic life use goals, biocriteria are associated with most tiers for each stream type (Table 9; Fig. 8). General use waters harbor "good" assemblages of freshwater organisms as measured by assemblage attributes and indices that represent the minimum CWA restoration goal. Exceptional use waters support biological assemblages with a community structure and ecosystem function that is minimally altered from the natural condition and provide an opportunity to protect high quality waters. Modified use waters have extensive physical alterations that predate the November 28, 1975 "existing use" date in U.S. water quality regulations (40 CFR Part 131) and are subject to a UAA. In addition, biological assemblages must demonstrate non-attainment of the general use biocriteria and be limited by these physical modifications as determined through a UAA before a modified use can be designated. These requirements are part of detailed UAA process (MPCA 2015).

A TALU framework is a structure of designated aquatic life uses that incorporates a hierarchy of use subclasses based on reasonable restoration potential. By refining these uses, good and exceptional quality water bodies can be protected while establishing attainable goals for water bodies degraded by legacy activities such as channelization. In doing so, the resulting management actions can be better tailored to attainable goals thereby improving effectiveness and efficiency of these actions. In a TALU framework, biocriteria are fundamental as they are used to both determine the designated aquatic life use (i.e., exceptional, general, or modified) for a water body and to assess if that designated use is attained. For example, a stream with fish and macroinvertebrate assemblages that meet the exceptional use biocriteria would be designated exceptional because the biological data demonstrates that the use is attainable. The same is true for the general use. As a consequence, channelized streams that meet the general or exceptional uses could not be designated modified use because a higher use is demonstrated to be attainable. Waters that do not meet at least the general use would be subject to a UAA to determine the appropriate and attainable aquatic life use.

In the USA, most states use biological assemblages to assess attainment of aquatic life goals, but at this time only three states have a TALU framework formally adopted into rule (i.e., Maine, Ohio, and Vermont). In Minnesota, extensive datasets of biological, physical, chemical, and landscape information were used to develop tiered biocriteria to support a TALU framework that protects aquatic biota in Minnesota streams. Compared to other states that have adopted a TALU framework, Minnesota's biocriteria development process was novel although comparable tools were used (i.e., BCG and RC). For Minnesota, the BCG provided narrative language which could be applied to the tiered uses to improve outreach to professionals and the general public. The BCG narratives and their associated biocriteria could also be linked to Minnesota's aquatic life use goals to demonstrate that they are consistent with state rules. The implementation of a stream typology consisting of nine fish and macroinvertebrate stream types each created some challenges, but this refined framework was necessary to minimize the effects of natural factors in a diverse landscape. Specifically, this typology resulted in a deficiency of least- and minimally-disturbed sites in the southern or prairie steam types which required greater reliance on BCG models to establish biocriteria. The BCG is a universal scale that was used to synchronize biocriteria for stream types from regions with different levels of stress. As a result, assessments of aquatic assemblages, regardless of prevailing land use or geographic location, will be based on comparable levels of degradation and will ensure that Minnesota's aquatic life use goals are protective and will comply with Minnesota and U.S. rules.

Acknowledgments The authors would like to thank the many staff and interns that performed field work and collected the extensive biological, habitat, and chemical data used in this study. Brenda DeZiel, Jeroen Gerritsen, Ed Hammer, Susan Jackson, Ed Rankin, Bryan Spindler, Kevin Stroom, and Mark Tomasek provided valuable input during the development of this project. We would also like to thank two anonymous reviewers for their helpful comments.

References

Ballentine, R. K., & Guarraia, L. J. (Eds.). (1977). *The integrity of water. Proceedings of a symposium, March 10–12, 1975.* Washington: US Environmental Protection Agency.

Berthouex, P., & Hau, I. (1991). Difficulties related to using extreme percentiles for water quality regulations. *Research Journal of the Water Pollution Control Federation, 63*(6), 873–879.

Courtemanch, D. L. (1996). Commentary in the subsampling procedures used for rapid bioassessments. *Journal of the North American Benthological Society, 15*(3), 381–385.

Courtemanch, D. L., Davies, S. P., & Laverty, E. B. (1989). Incorporation of biological information in water quality planning. *Environmental Management, 13*(1), 35–41.

Davies, S. P., & Jackson, S. K. (2006). The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. *Ecological Applications, 16*(4), 1251–1266.

DeShon, J. E. (1995). Development and application of the Invertebrate Community Index (ICI). In W. S. Davis & T. P. Simon (Eds.), *Biological assessment and criteria: tools for water resource planning and decision making* (pp. 217–243). Boca Raton: Lewis.

Fore, L., Karr, J., & Conquest, L. (1994). Statistical properties of an index of biological integrity used to evaluate water resources. *Canadian Journal of Fisheries and Aquatic Sciences, 51*(5), 1077–1087.

Frey, D. G. (1977). *Biological integrity of water—an historical approach. In The integrity of water. Proceedings of a symposium. US Environmental Protection Agency. Washington, DC, USA, 1977* (pp. 127–140). Washington: US Environmental Protection Agency.

Gakstatter, J., Gammon, J. R., Hughes, R. M., Ischinger, L. S., Johnson, M., Karr, J., Murphy, T., Murray, T. M., & Stuart, T. (1981). *A recommended approach for determining biological integrity in flowing waters*. Corvallis: U.S. Environmental Protection Agency.

Gerritsen, J., Stamp J. (2013). Calibration of the Biological Condition Gradient (BCG) in Cold and Cool Waters of the Upper Midwest BCG for Fish and Benthic Macroinvertebrate Assemblages: Prepared by Tetra Tech, Inc. for U.S. Environmental Protection Agency. (Avaiable at: https://www.uwsp.edu/cnr-ap/biomonitoring/Documents/pdf/USEPA-BCG-Report-Final-2012.pdf).

Gerritsen, J., Zheng, L., Leppo, E., & Yoder, C. O. (2013). Calibration of the biological condition gradient for streams of Minnesota. St. Paul, MN: prepared for the Minnesota Pollution Control Agency. (Avaiable at: https://www.pca.state.mn.us/sites/default/files/wq-s6-32.pdf).

Gibson, G. R., Barbour, M., Stribling, J. B., Gerritsen, J., & Karr, J. R. (1996). Biological criteria: technical guidance for streams and rivers - revised edition. Washington, D.C.: U.S. Environmental Protection Agency. (Avaiable at: http://www.epa.gov/nscep).

Gorman, O. T., & Karr, J. R. (1978). Habitat structure and stream fish communities. *Ecology, 59*(3), 507–515.

Hawkins, C. P., Norris, R. H., Hogue, J. N., & Feminella, J. W. (2000a). Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications, 10*(5), 1456–1477.

Hawkins, C., Norris, R., Gerritsen, J., Hughes, R., Jackson, S., Johnson, R., & Stevenson, R. (2000b). Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. *Journal of the North American Benthological Society, 19*(3), 541–556.

Hawkins, C. P., Olson, J. R., & Hill, R. A. (2010). The reference condition: predicting benchmarks for ecological and water-quality assessments. *Journal of the North American Benthological Society, 29*(1), 312–343.

Hughes, R. M., Larsen, D. P., & Omernik, J. M. (1986). Regional reference sites: a method for assessing stream potentials. *Environmental Management, 10*(5), 629–635.

Hughes, R. M. (1995). Defining acceptable biological status by comparing with reference conditions. In W. S. Davis & T. P. Simon (Eds.), *Biological assessment and criteria: tools for water resource planning and decision making* (pp. 31–47). Boca Raton: Lewis.

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

Karr, J. R., & Chu, E. W. (1999). *Restoring life in running waters*. Washington: Island Press.

Karr, J. R., & Dudley, D. R. (1981). Ecological perspective on water quality goals. *Environmental Management, 5*(1), 55–68.

Karr, J. R., & Yoder, C. O. (2004). Biological assessment and criteria improve total maximum daily load decision making. *Journal of Environmental Engineering, 130*, 594–604.

Koenker, R. (2009). quantreg: Quantile Regression. R package. (4.44 ed.).

Lyons, J. (1992). The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. *North American Journal of Fisheries Management, 16*, 241–256.

Meador, M.R., Cuffney, T.F., and Gurtz, M.E. (1993). Methods for sampling fish communities as a part of the National Water-Quality Assessment Program (Open-File Report 93–104). U.S. Geological Survey (Available at: http://water.usgs.gov/nawqa/protocols/OFR-93-104/fish1.html)

Moya, N., Hughes, R. M., Domínguez, E., Gibon, F.-M., Goitia, E., & Oberdorff, T. (2011). Macroinvertebrate-based multimetric predictive models for evaluating the human impact on biotic condition of Bolivian streams. *Ecological Indicators, 11*(3), 840–847.

MPCA. (2002). *Physical Habitat and Water Chemistry Assessment Protocol for Wadeable Stream Monitoring Sites (wq-bsm3-01)*. St. Paul: Minnesota Pollution Control Agency (Available at: https://www.pca.state.mn.us/sites/default/files/wq-bsm3-01.pdf).

MPCA. (2004). *Invertebrate Sampling Procedures (wq-bsm3-08)*. St. Paul: Minnesota Pollution Control Agency (Available at: https://www.pca.state.mn.us/sites/default/files/wq-bsm3-08.pdf).

MPCA. (2009). *Fish community sampling protocol for stream monitoring sites (wq-bsm3-03)*. St. Paul: Minnesota Pollution Control Agency (Available at: https://www.pca.state.mn.us/sites/default/files/wq-bsm3-03.pdf).

MPCA. (2014a). *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*. St. Paul: Minnesota Pollution Control Agency, Environmental Analysis and Outcomes Division (Available at: https://www.pca.state.mn.us/sites/default/files/wq-bsm4-02.pdf).

MPCA. (2014b). *Development of fish indices of biological integrity (FIBI) for Minnesota rivers and streams*. St. Paul: Minnesota Pollution Control Agency (Available at: https://www.pca.state.mn.us/sites/default/files/wq-bsm2-03.pdf).

MPCA. (2014c). *Development of macroinvertebrate indices of biological integrity (MIBI) for Minnesota streams*. St. Paul: Minnesota Pollution Control Agency (Available at: https://www.pca.state.mn.us/sites/default/files/wq-bsm4-01.pdf).

MPCA. (2015). *Draft technical guidance for designating aquatic life uses in Minnesota streams and rivers*. St. Paul: Minnesota Pollution Control Agency (Available at: https://www.pca.state.mn.us/sites/default/files/wq-s6-34.pdf).

MPCA. (2016). *Development of a Human Disturbance Score (HDS) for Minnesota streams*. St. Paul: Minnesota Pollution Control Agency (Available at: https://www.pca.state.mn.us/sites/default/files/wq-bsm3-10.pdf). Accessed 17 Feb 2016.

Ohio, E. P. A. (1987). *Biological criteria for the protection of aquatic life: Volume I: the role of biological data in water quality assessment*. Columbus: Ohio EPA division of Water Quality Planning and Assessment (Available at: http://www.epa.state.oh.us/Portals/35/documents/Vol1.pdf).

Ohio, E. P. A. (1989). *Biological Criteria for the Protection of Aquatic Life: Volume III. Standardized Biological Field Sampling and Laboratory Methods for Assessing Fish and Macroinvertebrate Communities*. Columbus: Ohio EPA division of Water Quality Planning and Assessment (Available at: http://www.epa.state.oh.us/portals/35/documents/BioCrit15_Vol3.pdf).

Paulsen, S. G., Mayio, A., Peck, D. V., Stoddard, J. L., Tarquinio, E., Holdsworth, S. M., Sickle, J. V., Yuan, L. L., Hawkins, C. P., & Herlihy, A. T. (2008). Condition of stream ecosystems in the US: an overview of the first national assessment. *Journal of the North American Benthological Society, 27*(4), 812–821.

Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., & Schmutz, S. (2006). Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages. *Journal of Applied Ecology, 43*(1), 70–80.

Pont, D., Hughes, R. M., Whittier, T. R., & Schmutz, S. (2009). A predictive index of biotic integrity model for aquatic-vertebrate assemblages of Western U.S. streams. *Transactions of the American Fisheries Society, 138*, 292–305.

R Development Core Team. (2015). R: A language and environment for statistical computing. (3.1.3 ed.). Vienna, Austria: R Foundation for Statistical Computing.

Rankin, E. T. (1995). Habitat indices in water resource quality assessments. In W. S. Davis & T. P. Simon (Eds.), *Biological assessment and criteria: tools for water resource planning and decision making* (pp. 181–208). Boca Raton: Lewis Publishers.

Reynoldson, T. B., Day, K. E., & Pascoe, T. (2000). The development of the BEAST: a predictive approach for assessing quality in the North American Great Lakes. In J. F. Wright & D. W. Sutcliffe (Eds.), *Assessing the biological quality of fresh waters: RIVPACS and other techniques* (pp. 167–180). Cumbria: Freshwater Biological Association.

Reynoldson, T. B., Norris, R. H., Resh, V. H., Day, K. E., & Rosenberg, D. M. (1997). The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society, 16*(4), 833–852.

Schinegger, R., Trautwein, C., Melcher, A., & Schmutz, S. (2012). Multiple human pressures and their spatial patterns in European running waters. *Water Environment Journal, 26*(2), 261–273.

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

Smith, M. J., Kay, W. R., Edward, D. H. D., Papas, P. J., Richardson, J., Simpson, J. C., Pinder, A. M., Cale, D. J., Horwitz, P. H. J., Davis, J. A., Yung, F. H., Norris, R. H., & Halse, S. A. (1999). AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. *Freshwater Biology, 41*, 269–282.

State of Ohio. (2010). State of Ohio Water Quality Standards, Water use designations and statewide criteria, Chapter 3745–1. (Chapter 3745–1). Cincinnatti, Ohio. (Availabe at: http://www.epa.ohio.gov/portals/35/rules/01_all.pdf).

Stoddard, J., Larsen, D., Hawkins, C., Johnson, R., & Norris, R. (2006). Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecological Applications, 16*(4), 1267–1276.

Systat Software. (2011). *SigmaPlot for Windows* (12th ed.). Chicago: Systat Software. Inc.

USEPA. (1990). *Biological Criteria: National program guidance for surface waters (EPA 440/5-90-004)*. U.S. Environmental Protection Agency: Washington. Avaiable at: http://www.epa.gov/nscep.

USEPA. (2011). *A primer on using biological assessment to support water quality management (EPA 810-R-11-01)*. Washington: Office of Science and Technology, Office of Water. Avaiable at: http://www.epa.gov/nscep.

USEPA. (2013). *Biological assessment program review: assessing level of technical rigor to support water quality management (EPA 820-R-13-001)*. Washington: Office of Science and Technology. Avaiable at: http://www.epa.gov/nscep.

USEPA. (2016). *A Practitioner's Guide to the Biological Condition Gradient: A Framework to Describe Incremental Change in Aquatic Ecosystems*. Washington: U.S. Environmental Protection Agency. Available at: http://www.epa.gov/wqc/biological-assessment-technical-assistance-documents-states-tribes-and-territories. Accessed 17 Feb 2016.

Van Sickle, J., & Hughes, R. M. (2000). Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. *Journal of the North American Benthological Society, 19*(3), 370–384.

Vinson, M. R., & Hawkins, C. P. (1996). Effects of sampling area and subsampling procedure in comparisons of taxa richness among streams. *Journal of the North American Benthological Society, 15*(3), 392–399.

Waite, I. R., Herlihy, A. T., Larsen, D. P., & Klemm, D. J. (2000). Comparing strengths of geographic and nongeographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. *Journal of the North American Benthological Society, 19*(3), 429–411.

White, D., & Omernik, J. M. (2007). *Minnesota level III and IV ecoregions map (scale 1:2,500,000)*. Corvallis: U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Western Ecology Division.

Whittier, T., Hughes, R., Stoddard, J., Lomnicky, G., Peck, D., & Herlihy, A. (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*(3), 718–735.

Wright, J. F., Furse, M. T., & Moss, D. (1998). River classification using invertebrates: RIVPACS applications. *Aquatic Conservation: Marine and Freshwater Ecosystems, 8*, 617–631.

Yarbro, L. A., Kuenzler, E. J., Mulholland, P. J., & Sniffen, R. P. (1984). Effects of stream channelization on exports of nitrogen and phosphorus from North Carolina coastal plain watersheds. *Environmental Management, 8*(2), 151–160.

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

Yoder, C. O., & Rankin, E. T. (1995). Biological criteria program development and implementation in Ohio. In W. S. Davis & T. P. Simon (Eds.), *Biological assessment and criteria: tools for water resource planning and decision making* (pp. 109–144). Boca Raton: Lewis Publishers.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338–353.

Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information Sciences, 178*(13), 2751–2779.