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

Improving Water Quality Standards and Assessment Approaches for the Upper Mississippi River: UMR Clean Water Act Biological Assessment Implementation Guidance

Midwest Biodiversity Institute
Center for Applied Bioassessment & Biocriteria
P.O. Box 21561
Columbus, OH 43221-0561
Chris O. Yoder, Principal Investigator
mbi@mwbinst.com

Improving Water Quality Standards and Assessment Approaches for the Upper Mississippi River: UMR Clean Water Act Biological Assessment Implementation Guidance

by

Chris O. Yoder
Robert J. Miltner
Vickie L. Gordon
Center for Applied Bioassessment and Biocriteria
Midwest Biodiversity Institute
P.O. Box 21561
Columbus, OH 43221-0561

and

Edward T. Rankin
Ohio University
Voinovich School for Leadership and Public Affairs
The Ridges, Building 22
Athens, OH 45701

and

Nathaniel B. Kale
David K. Hokanson
Upper Mississippi River Basin Association
415 Hamm Building, 408 St. Peter Street
St. Paul, Minnesota 55102

September 1, 2011

This project was led by the Upper Mississippi River Basin Association's Water Quality Task Force. Illinois, Iowa, Minnesota, Missouri, and Wisconsin jointly funded the project through Section 604(b) of the Clean Water Act, using appropriations from the American Recovery and Reinvestment Act of 2009.

Table of Contents

Table of Contents (continued)

List of Figures

List of Figures (continued)

List of Tables

List of Tables (continued)

List of Acronyms

BCG	Biological Condition Gradient
CWA	Clean Water Act
EMAP - GRE	Environmental Monitoring and Assessment Program - Great River Ecosystems
EMP	Environmental Management Program
EWH	Exceptional Warmwater Habitat
IA DNR	Iowa Department of Natural Resources
IBI	Index of Biotic Integrity
IL DNR	Illinois Department of Natural Resources
IL EPA	Illinois Environmental Protection Agency
LRW	Limited Resources Waters
LTRMP	USACE EMP Long Term Resource Monitoring Program
MBI	Midwest Biodiversity Institute
MCES	Twin Cities Metropolitan Council Environmental Services
MN DNR	Minnesota Department of Natural Resources
MN PCA	Minnesota Pollution Control Agency
MO DNR	Missouri Department of Natural Resources
MO DoC	Missouri Department of Conservation
MWH	Modified Warmwater Habitat
NASQAN	National Stream Accounting Network
NRSA	National Rivers and Streams Assessment
SAV	Submersed Aquatic Vegetation
TALU	Tiered Aquatic Life Uses
UMESC	Upper Midwest Environmental Sciences Center
UMR	Upper Mississippi River
UMRBA	Upper Mississippi River Basin Association
UMRCC	Upper Mississippi River Conservation Committee
UMRS	Upper Mississippi River System
USACE	United States Army Corps of Engineers
US EPA	United States Environmental Protection Agency
USGS	United States Geological Survey
WI DNR	Wisconsin Department of Natural Resources
WQEC	(Upper Mississippi River Basin Association) Water Quality Executive Committee
WQS	Water Quality Standards
WQTF	(Upper Mississippi River Basin Association) Water Quality Task Force
WWH	Warmwater Habitat

Acknowledgements

This project was made possible by support from the Upper Mississippi River Basin Association (UMRBA) and is the result of their support and vision for the Upper Mississippi River. Dave Hokanson (UMRBA) served as the principal technical and project contact for the UMRBA. Members of the Upper Mississippi River Water Quality Task Force (WQTF) provided helpful comments and feedback on prior drafts of this report. We also acknowledge the input from the various other agencies and organizations that participated in the working sessions in 2010-11. A list of project participants follows:

Name	Organization	Name	Organization
Butch Atwood	Illinois DNR	**Ken Lubinski**	USGS, UMESC
Ann Holtrop	Illinois DNR	**Shannan Garretson**	Iowa Environmental Council
Gregg Good	Illinois EPA	**Susan Heathcote**	Iowa Environmental Council
Matt Short	Illinois EPA	**Kim Wright**	Midwest Environmental Advocates
Roy Smogor	Illinois EPA	**Gretchen Benjamin**	The Nature Conservancy
Rochelle Cardinale	Iowa DNR	**Chris Yoder**	Midwest Biodiversity Institute
John Olson	Iowa DNR	**Peg Donnelly**	US EPA, Region 5 IPA to UMRBA
Tom Wilton	Iowa DNR	**Dave Hokanson**	UMRBA
Megan Moore	Minnesota DNR	**Nat Kale**	UMRBA
Tim Schlagenhaft	Minnesota DNR	**Barb Naramore**	UMRBA
Will Bouchard	Minnesota PCA		
Marvin Hora	Minnesota PCA		
Shannon Lotthammer	Minnesota PCA		
Mohsen Dkhili	Missouri DNR		
Jim Baumann	Wisconsin DNR		
Terry Dukerschein	Wisconsin DNR		
John Sullivan	Wisconsin DNR		
Bill Franz	US EPA, Region 5		
Ed Hammer	US EPA, Region 5		
Larry Shepard	US EPA, Region 7		
Ted Angradi	US EPA, ORD		
David Bolgrien	US EPA, ORD		
Ken Barr	USACE, Rock Island District		
Karen Hagerty	USACE, Rock Island District		
Marvin Hubbell	USACE, Rock Island District		
Chuck Spitzack	USACE, Rock Island District		
Mike Coffey	USFWS		
Jon Duyvejonck	USFWS		
Barry Johnson	USGS, UMESC		

Executive Summary

The Upper Mississippi River Basin Association (UMBRA) Water Quality Task Force (WQTF) and Water Quality Executive Committee (WQEC) seek to improve implementation of the states' Clean Water Act (CWA) programs on the Upper Mississippi River (UMR). In order to pursue better use of biological assessments, the WQTF initiated a project in 2009 to develop a Clean Water Act Biological Assessment Implementation Guidance Document for the interstate UMR.

The five UMR states are responsible for the implementation of the CWA on their portion of the UMR within the borders of each. Under the CWA, all of the UMR states have assigned one or more aquatic life uses to their respective reach of the UMR. At present the UMR states primarily utilize analyses of chemical data to assess the condition of the UMR as it relates to aquatic life goals and uses. While the states have not previously utilized biological measures in their assessments of the UMR, they are interested in evaluating the potential benefits and challenges associated with integrating biological assessment along with the current chemical and physical assessment. This Guidance Document is intended for the UMR states' use as they consider how to integrate UMR-specific biological assessment approaches into their CWA programs and is primarily focused on the interstate portion of the UMR, but its analyses also address the navigable portion of the Mississippi River that is internal to Minnesota. The document is also focused specifically on the UMR main channel (as opposed to adjacent aquatic habitat strata) and is scoped to primarily evaluate existing methods and indices, rather than developing new ones.

Fish are the most commonly sampled biological assemblage among all of the different UMR programs, with the majority of entities performing some type of fish sampling. Macro-invertebrates are sampled by at least two entities and represent a commonly used second assemblage. Submersed aquatic vegetation is a promising third assemblage that is sampled by two programs. Based on our review of the extant bioassessment programs, fish and macroinvertebrates emerge as the two assemblages that are in a comparative state of "readiness" to support a near-term biological assessment for the entirety of the UMR main channel. Submersed aquatic vegetation (SAV) comprises a realistically applicable third assemblage and has only recently become available in terms of a comparative state of readiness. Mussels, algae, and the remaining assemblage groups are each in various stages of development, testing, and refinement. Each will need to have a readily available assessment mechanism, i.e., a calibrated index that is relevant to the Biological Condition Gradient (BCG), to serve the goals of this project.

A preliminary analysis of biological condition thresholds for the UMR main channel was conducted using data from US EPA's Environmental Monitoring and Assessment Program-Great Rivers Ecosystems (EMAP-GRE). These analyses also provided an opportunity to test the ability of the existing indices – the Great Rivers Fish Index (GRFIn), Great Rivers Macroinvertebrate Index (GRMIn), and Submerged Macrophyte Index (SMI) – to produce meaningful and CWA-relevant assessments of aquatic life condition. Of the readily available approaches, the EMAP-GRE program protocol currently provides the best "fit" for ongoing CWA biological assessment

because it is spatially comprehensive, has contiguous sampling sites allowing for scalability and detection of pollution gradients, it includes two assemblages, and is a level 4 bioassessment program. In addition, two of the leading index candidates – GRFIn and GRMIn – are compatible with the data produced using EMAP-GRE methods. The key component in selecting a preferred approach is the sampling site as the fundamental unit of assessment. The advantages of the EMAP-GRE sample site approach coupled with the lack of a macroinvertebrate assemblage tool within the USACE EMP-Long Term Resource Monitoring Program (LTRMP), were influential in the identification of EMAP-GRE as a preferred protocol.

A major emphasis of the project is to identify "potential impairment thresholds for the UMR main channel in determining the attainment of aquatic life uses," as identified in the project's work plan. To that end, we conducted a preliminary assessment of the derivation of biological thresholds for review by the UMRBA Water Quality Task Force (WQTF) in January 2011. While this step was not anticipated in the original project work plan, it proved crucial in making progress in the development of the draft guidance, and also responded to the discussion at the project's second work session regarding an initial CWA biological assessment of the UMR. A stand alone report entitled *"Preliminary Analysis of Biological Assessment Thresholds for Determining Aquatic Life Use Attainment Status in the Upper Mississippi River Mainstem"* was produced that details the analyses conducted and preliminary results of bioassessment thresholds. As such, this report contains options available to the WQTF for an initial CWA bioassessment of the UMR main channel.

Any attempt to develop a threshold for biological measures in response to the intent of CWA Section 101[a][2] for the ". . . protection and propagation of fish, shellfish, and wildlife" is necessarily seen as having an inherent level of subjectivity and reliance on best professional judgment. However, a well organized and developed empirical process can aid in setting such thresholds by providing a systematic and explicit approach for threshold selection. We took two different approaches to reach this point in the project regarding the empirical approach. The statistical derivation of numeric thresholds or biocriteria was attempted using a number of different combinations of biological indices that included both "externally" and "internally" derived thresholds. External approaches involved the analysis of data outside of the UMR and in this case included a Regional EMAP project on selected large tributary rivers of the Upper Mississippi and Ohio Rivers. This approach was further subdivided between a selection of REMAP rivers that represent conditions as close as possible to the UMR and using the closest high quality member of that collection (lower St. Croix River). Additionally, indices developed by GRE for the Missouri River were applied to the open river UMR as an additional point of comparison. Internal approaches included using the data from the UMR itself and extracting thresholds based on different sectionings of the GRE stressor gradient and using the results from sites that represent the "best performing" in the UMR main channel. In addition, to augment the empirical approach, we undertook an initial biological condition gradient (BCG) analysis for the UMR. As such the BCG is used here as an independent method for evaluating the ecological meaning of quantitative thresholds derived by empirical means. The BCG therefore provides a rational and consistent means for helping determine appropriate aquatic life uses for the purpose of setting biological impairment thresholds.

Current non-biological approaches to UMR CWA assessment have resulted in a minority of reaches identified with aquatic life use impairments (e.g., 4 of 13 interstate assessment reaches had aquatic life use impairment in the 2008 assessment cycle). All of the UMR-derived threshold analyses produced significantly greater non-attainment for the UMR main channel as a whole (less in some reaches, more in others) than the current non-biological 303[d] list. As such, the application of a biological assessment to the UMR main channel will comprise a significant change with regard to aquatic life use support under any of the UMR-derived approaches examined. Additionally, multiple methods of analysis, using both internally and externally-derived approaches produced closely similar values for a baseline CWA attainment threshold.

Developing a sustained UMR CWA assessment program based on the principles outlined herein brings the focus on providing a measurement framework that can assess current conditions, but also detects changes in increments of condition and serves as a feedback to the various management programs that are working to restore and maintain the biological quality of the UMR. While the development of thresholds is a critical component of this framework, it is a result of the quality and characteristics of the overall monitoring and assessment program that will eventually be applied to the UMR. Developing a comprehensive strategy that actually leads to the execution of this type of monitoring and assessment program is an essential next step.

As the states move forward utilizing the recommendations made in the report, numerous challenges remain. These challenges include identifying an appropriate entity or entities to conduct monitoring, gathering financial resources to support monitoring, managing data, coordinating assessment methodologies, and addressing the policy implications of adopting biological assessment. However, this project has demonstrated that a UMR CWA assessment incorporating biology is feasible given readily available tools. As such, the states are encouraged to continue their efforts by utilizing the information provided in this guidance to:

1. Develop a UMR-wide CWA monitoring strategy that the follows the principles outlined herein.
2. Utilize a modification of the EMAP-GRE design as the baseline spatial sampling design, i.e., execute an intensive, longitudinal “pollution survey” design.
3. Examine programmatic and organizational options for implementing such a strategy outlining the costs of each and the technical pros and cons.
4. Use the biological assemblage, biological index, and biocriteria threshold recommendations included herein as the basis for an initial biological assessment of the UMR main channel and future assessments based on a new monitoring strategy.
5. Develop and utilize a data management system that is easy to use, easy to access, and which delivers sampling data and transformed data in a portable and relational format.

Introduction

The Upper Mississippi River Basin Association (UMRBA) is a regional interstate organization formed by the governors of its respective member states (Illinois, Iowa, Minnesota, Missouri, and Wisconsin) to coordinate the states' programs and to work with federal agencies that have river responsibilities. UMRBA is involved with programs related to ecosystem protection and restoration, water quality, spills, floodplain management, flood control, water supply, and commercial navigation. To this end, UMBRA supports two work groups, the Water Quality Task Force (WQTF) and the Water Quality Executive Committee (WQEC), which exist as forums of consultation and interaction among the five member states and U.S. EPA Regions 5 and 7 regarding water quality issues.

The UMBRA WQTF and WQEC seek to improve implementation of the states' Clean Water Act (CWA) programs on the Upper Mississippi River (UMR). Specific outcomes from their efforts to date have included adoption of common CWA assessment reaches, enhanced collaboration with ecosystem restoration programs, numerous reports on UMR water quality issues and, most recently, efforts to examine aquatic life use designations and nutrient monitoring, occurrence, and impacts. Moreover, the WQTF and WQEC recognize that current efforts need to move beyond simple coordination to the development of new, more effective and shared assessment tools including the better use of biological assessments.

In order to pursue better use of biological assessments, the WQTF initiated a project in 2009 to develop a Clean Water Act Biological Assessment Implementation Guidance Document for the interstate UMR. UMRBA contracted with the Midwest Biodiversity Institute, Inc. (MBI) to provide its technical and professional services in the development of the guidance document. The guidance document, as presented herein, reflects the collaborative project work between the WQTF, MBI staff, and UMRBA staff over the course of the project.

Chapter 1: Development of UMR CWA Biological Assessment Guidance

Background

The Upper Mississippi River (UMR) offers an abundant and unique ecosystem. The flora and fauna native to the river, in turn, support numerous recreational and economic pursuits on the local, regional, and national level. Over the past century, its aquatic life has been subjected to stressors in the form of landscape alterations, changes in the physical structure of the channel, and a variety of anthropogenic pollution. Together these have altered water quality and habitat in and around the river. One of the regulatory tools available to manage these stressors on aquatic life in the UMR is the Clean Water Act (CWA) and an effective approach to identifying problems at the local, reach, and interstate scale under the CWA is essential in protecting aquatic life.

The five UMR states are responsible for the implementation of the CWA on their portion of the UMR. Under the CWA, all of the UMR states have assigned one or more aquatic life uses to their respective reach of the UMR. At present, the UMR states primarily utilize analyses of chemical data to assess the condition of the UMR as it relates to aquatic life goals and uses (UMRBA 2004; Sullivan et al. 2002). These chemical parameters, such as dissolved oxygen or pollutant concentrations, serve as "surrogates" that only indirectly reflect the integrity of the biological communities. Chemical data-focused assessments are generally performed by comparing observed parameter concentration values to established pollutant criteria thresholds, leading to a "pass-fail" judgment – either the waterbody supports the aquatic life use, or it does not. While the technical inadequacies of using chemical-specific monitoring data to assess aquatic life use attainment are well known (Yoder and Rankin 1998; Karr and Yoder 2004), it is the most readily available to the states of the ambient chemical, physical, and biological data that has been collected in the UMR.

The states, with the encouragement of the U.S. Environmental Protection Agency (EPA), have generally moved more towards using biology in their assessments of smaller river and stream systems, which allows for a more direct measurement and assessment of an aquatic life use, as well as increments of quality between "passing" and "failing" provided the bioassessment method has that capacity. Such biologically-based assessments consist of monitoring representative biotic assemblages, such as fish or macroinvertebrates, and a methodology for interpreting these observations that reports the overall condition of aquatic life in the waterbody. This generally consists of a numeric index that is based on representative attributes of each assemblage that are tailored to the waterbody of application.

While biological indicators and assessment techniques are now routinely applied by the states to assess their inland rivers, streams, wetlands, and lakes, such indicators and the supporting technical infrastructure have historically been lacking for the UMR on a comprehensive scale. This is not to say that some promising approaches do not already exist, but achieving the uniform usage of sufficiently rigorous and relevant bioassessments is in need of further development, including review for technical adequacy, rigor, and practicality of wider application throughout the UMR. In addition, some of the important policy driving issues also need to be examined and include the adequacy of current state water quality standards with specific reference to designated aquatic life uses.

Guidance Purpose and Scope

While the states have not previously utilized biological measures in their assessments of the UMR, they are interested in evaluating the potential benefits and challenges associated with integrating biological assessment along with the current chemical and physical assessment. The states have acted on this interest by developing a project – of which this guidance is the principal product – to accomplish the following:

1) Identify the scientific and regulatory issues associated with applying biological assessment under the CWA on the UMR;
2) Review available biological assessment protocols and indices to gauge their suitability for supporting biological assessment of the aquatic life use in the *main channel* of the UMR; and
3) Provide guidance to the state CWA programs regarding the implementation of biological assessment on the UMR.

The Guidance Document is intended for the UMR states' use as they consider how to integrate UMR-specific biological assessment approaches into their CWA programs. This Guidance Document is primarily focused on the interstate portion of the UMR[1], but its analyses also address the navigable portion of the Mississippi River that is internal to Minnesota.

The Guidance Document summarizes the underlying research conducted by MBI in accordance with the project work plan, reflects the discussions at three working sessions for the project, as well as additional conversations with the WQTF, and describes options for developing a biological assessment of the UMR main channel using existing information. Specifically, this document includes:

1. An overview of current UMR monitoring and assessment programs;
2. A discussion and description of biological assessment concepts and applications;
3. Identification, description, and evaluation of currently available UMR biological assessment protocols and indices;
4. A preliminary examination of and recommendations for biological assessment thresholds;
5. A discussion of the implications of adopting biological assessment approaches on the UMR; and
6. A discussion of implementation considerations and next steps.

[1] The interstate UMR is defined as the Mississippi River mainstem from its confluence with the St. Croix River to its confluence with the Ohio River, including interstate border reaches in Minnesota, Wisconsin, Illinois, Iowa, and Missouri.

Potential Other Uses

As is apparent in the discussion above, this guidance document is structured specifically to address CWA program needs. However, a May 2009 workshop entitled "Examining Biological Indicators for the Upper Mississippi River: Applications in Clean Water Act & Ecosystem Restoration Programs" (UMRBA 2009) revealed a wide range of issues, interests, and understanding among all of the state and federal regulatory (i.e., CWA programs) and natural resource agencies, all of which should benefit from many of the goals and outputs of this project. Additionally, the perspectives of natural resource managers, river scientists, other river experts and stakeholders were explicitly included in the guidance document development process, both to benefit the quality of the guidance and to maintain its relevance for these groups.

Chapter 2: UMR Monitoring and Assessment Programs Overview

Overview of UMR Monitoring and Assessment Programs

This project's development of UMR CWA bioassessment guidance takes place in the context of an existing scientific, regulatory, and institutional setting – i.e., monitoring and assessment on the UMR is certainly not a "blank slate." As such, understanding the current structure of UMR monitoring and assessment – in both a scientific and institutional sense – is critical in evaluating existing and in-development biological assessment protocols. Additionally, examining existing efforts enhances the potential for synergies between programs and helps minimize the possibility for duplication of effort. Accordingly, an extensive survey of relevant reports and publications applicable to the UMR was completed as part of this project's scoping phase. This included a mix of open requests, online library searches, canvassing various websites, and WQTF input.

As displayed in Table 1, our results show that the UMR has been and continues to be the subject of extensive monitoring and research at many times, by many entities, and at various levels of government. Our review revealed that no fewer than 14 major entities/programs either are or have recently have been engaged in UMR monitoring and assessment. These programs have collected data for various purposes and under differing mandates.

Table 1: UMR Monitoring and Assessment Programs.[2]

Level	Agency	Program	Monitoring*	Assemblages**
Federal	U.S. EPA	EMAP-GRE***	Bio, Chem, Phys	Fish, Mac, Ppt, Pkt, SAV
		NRSA	Bio, Chem, Phys	Fish, Mac, Ppt, Pkt
	USGS	NASQAN	Chem, Physical	None
	USACE	EMP-LTRMP	Bio, Chem, Phys	Fish, Pkt, SAV
		Program-Specific	Chem, Physical	None
State	Illinois	IL EPA (CWA)	Chem, Phys	None
		IL DNR	Bio	Fish
	Iowa	IA DNR (CWA)	None	N/A
	Minnesota	MN PCA (CWA)	Bio, Chem, Phys	Fish, Mac
		MN DNR	Bio	Fish
	Missouri	MO DNR (CWA)	None	N/A
		MO DoC	Bio	Fish
	Wisconsin	WI DNR (CWA)	Bio, Chem, Phys	Fish
Local	Twin Cities Metropolitan Council	MCES	Bio, Chem, Phys	Mac

***Bio: Biological communities; Chem: Chemical parameters; Phys: Physical conditions.**
****Mac: Macroinvertebrates; Ppt: Periphyton; Pkt: Phytoplankton; SAV: Submersed Aquatic Vegetation.**
*****EMAP-GRE monitoring limited to 2004-2006, it is not an ongoing program.**

[2] See Background and Scoping Report (Yoder et al. 2010) for detailed information about these programs.

In addition to the summary presented in Table 1, the following observations can be made regarding UMR monitoring and assessment programs:

- U.S. EPA (NRSA, EMAP-GRE) and USGS (NASQAN) monitoring programs operate nationally, and include anywhere from a few (NRSA) to many (EMAP-GRE) monitoring stations on the UMR. NRSA is designed to be a periodic snapshot (once every 5 years) at a national scale, while EMAP-GRE, as a one-time research and development program, does not have an ongoing monitoring presence. NASQAN does not have a biological monitoring component.
- With an exclusive focus on the UMR, the USACE Environmental Management Program's LTRMP has generated much more data about the river since the late 1980s than any other monitoring program listed here. Data collection is performed by six field stations, five of which are located on the UMR mainstem. Each UMR state contains one mainstem field station. The field stations are primarily funded through LTRMP, and are operated collaboratively by the USGS and a state agency. The USGS' Upper Midwest Environmental Services Center (UMESC) provides data management and scientific analysis in support of LTRMP. Of note, the LTRMP field stations also collected data for EMAP-GRE program.
- USACE also collects data, including biological data, for projects it is involved in along the UMR. These project-specific data are not unified by a single overarching program or methodology.
- While Illinois, Minnesota, and Missouri all collect biological data, none have established a protocol specifically for assessing the condition of the UMR based on that data. Wisconsin DNR has developed and employs a fish index of biotic integrity (IBI) on the UMR as part of a statewide assessment of its large rivers. This is a separate program from the Wisconsin DNR component of the LTRMP. Iowa DNR does not sample in the UMR, and bases all CWA assessments on data obtained from other programs or states. Therefore, with the exception of the Wisconsin DNR's large rivers program (and LTRMP-supported field stations), no state agency independently collects biological data in a manner that supports the goals of this project.
- The Twin Cities Metropolitan Council operates a monitoring program that includes the collection of macroinvertebrate data mostly on the non-interstate UMR - only one site is on the interstate portion of the UMR. No other significant local sampling protocols came to the attention of MBI or UMRBA during the research conducted for this project.

UMR programs with a biological component are discussed in greater detail in Chapter 4 of this document, where they are considered as potential candidates for UMR CWA assessment. Federal, state, and local program reports and field manuals applicable to biological monitoring and assessment in the UMR are listed in the appendices to this project's Background and Scoping Report (Yoder et al. 2010). This includes all of the aforementioned entities including EMAP-GRE, LTRMP, NRSA, the UMR states, and the local Twin Cities Metropolitan Council. Relevant peer reviewed literature was also documented in the Background and Scoping Report (Yoder et al. 2010).

Current UMR CWA Assessment Approaches

Within the context of the current monitoring and assessment on the UMR, the states are required by the CWA to assess the condition of their portion of the UMR. Currently, UMR CWA assessments are conducted independently by each of the states, i.e., there is no unified CWA assessment of the River. Furthermore, the current assessments of aquatic life use status by the states in the UMR are based primarily on chemical and physical water quality data.

The states do, however, use a common set of assessment reaches to organize the interstate UMR into segments for Clean Water Act 305[b] water quality assessments and 303[d] listings and reporting (Figure 1). These reaches were defined by the Upper Mississippi River Conservation Committee Water Quality Technical Section in its March 2002 Upper Mississippi River Water Quality Assessment (Sullivan et al. 2002). This framework was adopted by the states via an MOU facilitated by UMRBA in September 2003, and as such comprises "minimum assessment reaches" for the UMR mainstem.

Currently, a minority of UMR assessment reaches (4 of 13 in the 2008 assessment cycle) are listed as having an impairment of the CWA aquatic life designated use (Table 2). In conducting their assessments of aquatic life use, the states have been able to incorporate some, but not all, of the chemical/physical data collected by the various programs described previously in this report. As a result of the independent application of their own assessment protocols and a focus on physical and chemical data, CWA impairment listings for the UMR are predominated by impairments related to toxic pollutants and "legacy" contaminants and are often inconsistent between bordering states (see Appendix A) **.** In addition, the states are limited by the lack of a methodology to translate biological data into an assessment of the status of aquatic life uses. It is in light of the limitations of the current approaches that the states sought to pursue the potential application of biological approaches in UMR aquatic life use assessment.

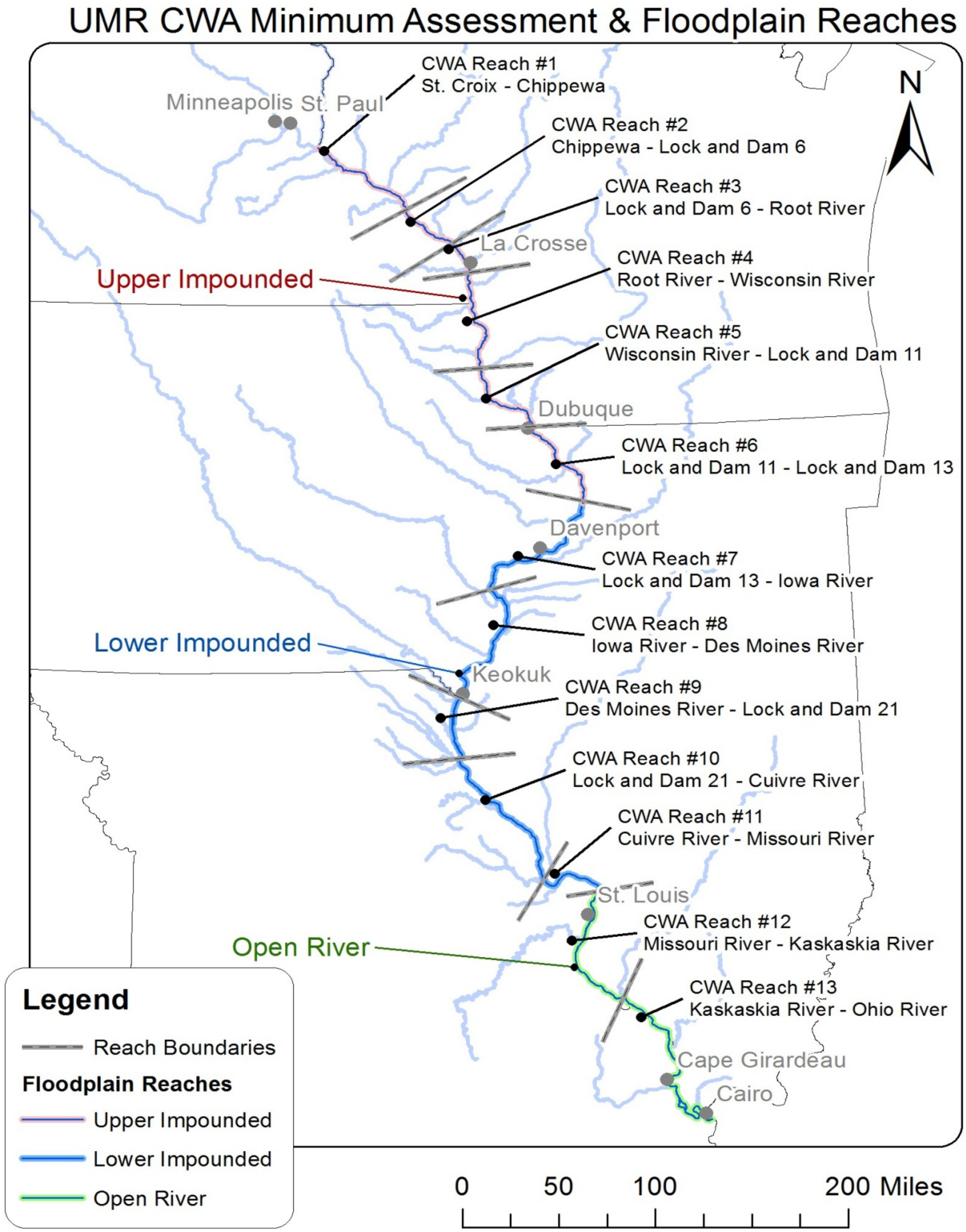


Figure 1. UMR CWA minimum assessment reaches. Note that assessment reaches 1-6 comprise the "upper impounded" floodplain reach, assessment reaches 7-11 comprise the "lower impounded" floodplain reach, and assessment reaches 12-13 comprise the "open river" floodplain reach.

Table 2. Attainment of UMR Aquatic Life Designated Use As Reflected in States' 2008 303(d) Impairment Listings as Submitted to U.S. EPA.

<table>
<tr><th>State</th><th>Aquatic Life Use Attained? (If no, cause of impairment)</th><th>UMR Minimum Assessment Reach</th><th>Aquatic Life Use Attained? (If no, cause of impairment)</th><th>State</th></tr>
<tr><td rowspan="4">MN</td><td>No
(Turbidity)</td><td>Reach1
St. Croix River to Chippewa River</td><td>No
(Suspended Solids)</td><td rowspan="7">WI</td></tr>
<tr><td>Yes</td><td>Reach 2
Chippewa River to Lock & Dam 6</td><td>Yes</td></tr>
<tr><td>Yes</td><td>Reach 3
Lock & Dam 6 to Root River</td><td>Yes</td></tr>
<tr><td>Yes</td><td rowspan="2">Reach 4
Root River to Wisconsin River</td><td rowspan="2">Yes</td></tr>
<tr><td rowspan="6">IA</td><td>Yes</td></tr>
<tr><td>Yes</td><td>Reach 5
Wisconsin River to Lock & Dam 11</td><td>Yes</td></tr>
<tr><td rowspan="2">Yes</td><td rowspan="2">Reach 6
Lock & Dam 11 to Lock & Dam 13</td><td>Yes</td></tr>
<tr><td>Yes</td><td rowspan="8">IL</td></tr>
<tr><td>No
(Aluminum, localized nutrients)</td><td>Reach 7
Lock & Dam 13 to Iowa River</td><td>Yes</td></tr>
<tr><td>No
(Aluminum)</td><td>Reach 8
Iowa River to Des Moines River</td><td>Yes</td></tr>
<tr><td rowspan="5">MO</td><td>Yes</td><td>Reach 9
Des Moines River to Lock & Dam 21</td><td>Yes</td></tr>
<tr><td>Yes</td><td>Reach 10
Lock & Dam 21 to Cuivre River</td><td>Yes</td></tr>
<tr><td>Yes</td><td>Reach 11
Cuivre River to Missouri River</td><td>Yes</td></tr>
<tr><td>No
(Localized lead and zinc)</td><td>Reach 12
Missouri River to Kaskaskia River</td><td>Yes</td></tr>
<tr><td>Yes</td><td>Reach 13
Kaskaskia River to Ohio River</td><td>Yes</td></tr>
</table>

Chapter 3: Biological Assessment Concepts and Application

This chapter provides a brief summary of important biological assessment concepts relevant for the Upper Mississippi River. However, it is not intended to be complete description of these concepts. The documents referenced herein should be consulted for more complete information.

Biologically-Based Approaches

Awareness of biological response was central to early efforts to protect water quality. In fact, an early 20th century approach utilized in Europe, the Saprobien system, recognized three facts ignored by the later emphasis on chemical criteria: 1) not all water bodies are the same (large or small streams; cold, cool, or warmwater streams), 2) differences can exist within the same water body, and 3) the effects of human activities inevitably leads to a continuum, or gradient of biological condition (Karr and Yoder 2004).

Chemical criteria, as is employed currently for CWA assessments on the UMR, produce a bivariate assessment hierarchy, variously termed impaired or unimpaired, in compliance or not. Simple to understand and easy to use, this dichotomy is neither sufficiently accurate nor robust enough to address the issues that challenge 21st century water quality management (Karr and Chu 1999). While it was useful to detect and regulate the very visible and gross point source pollution problems that were the impetus for the 1972 CWA amendments, the steady-state limitations and assumptions that are inherent to this approach are challenged by the dynamics of nonpoint sources and non-toxic impacts such as nutrients, habitat, and alien species.

In aquatic life use assessment, biologically-based approaches also have the advantages of integrating multiple stressors (Figure 2) and measuring data closer to the endpoint of concern than chemical and physical criteria (Figure 3). Because of this proximity, the biological component can be considered the "gold standard" in CWA aquatic life use assessment.

Bioassessment and the Biological Condition Gradient

Recent advances in biological assessment provide a strong conceptual and technical basis to move beyond the prevailing pass-fail paradigm that is used by most states to fulfill their CWA reporting obligations. Modern analytical tools, such as multimetric biological indices, enhance our ability to measure quality in a manner that communicates the severity and extent of impairment beyond simple statements of attainment and non-attainment.

The Biological Condition Gradient (BCG; Davies and Jackson 2006) describes the continuum of possible conditions of a water body, ranging from exceptional (pristine/undisturbed) to very poor (modified/degraded). When sufficiently integrated within the concepts of the BCG, bioassessment fosters the recognition and more complete interpretation of patterns in biological data, including the relationships among chemical, physical, and biological stressors and their influences on the five major features of water resources (Figure 2) and along the BCG (Figure 4).

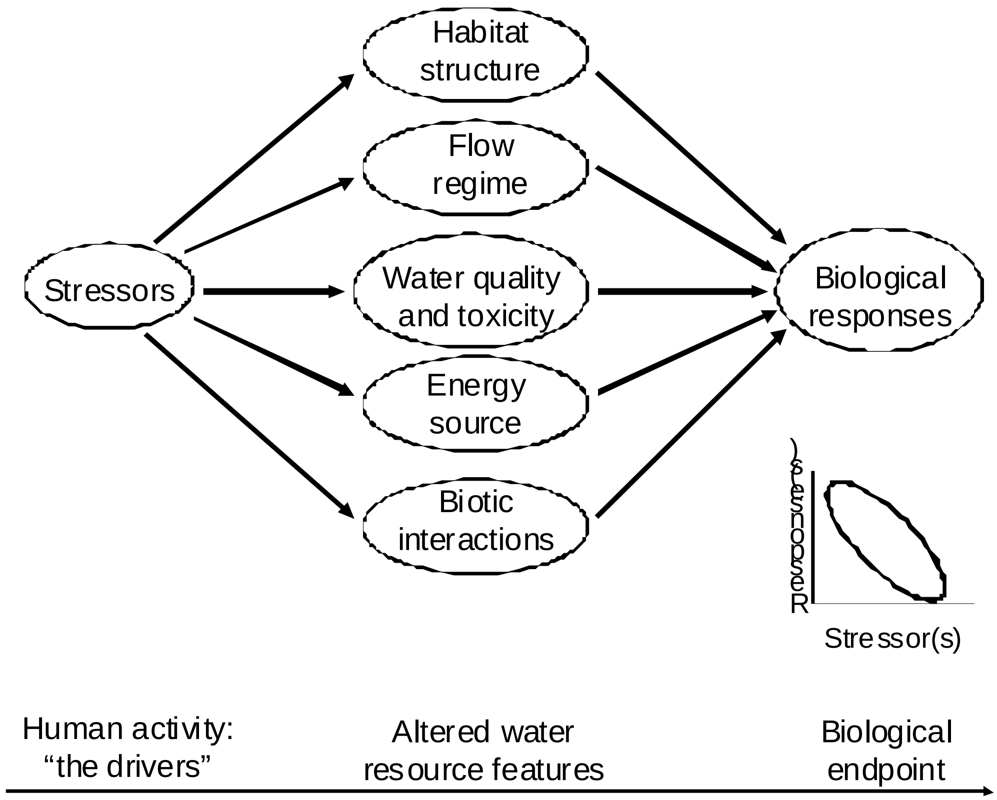


Figure 2. *Linkages from human activity (the stressors or drivers of system change) through the five major water resource features, to the biological responses producing ambient condition and in response to alterations of one or more of these factors, i.e., the biological endpoints are of primary interest in biological assessment programs. This model illustrates the multiple causes of water resource changes associated with human activities. The insert illustrates the relationship between stressor dose and the gradient of biological responses that signal a good biological metric (after Karr and Yoder 2004).*

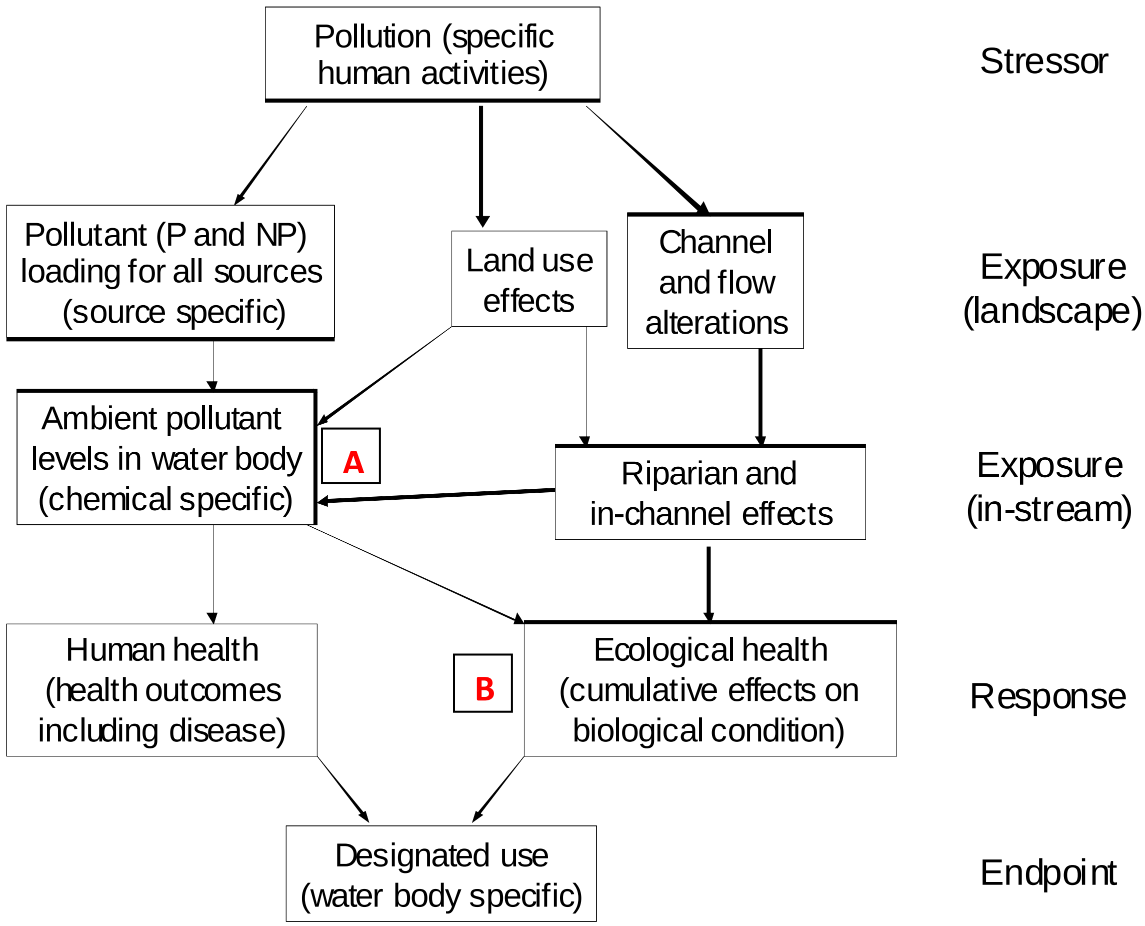


Figure 3. *Position of the criterion (stressor, exposure, or response), illustrating the relationships among human activities, specific types of criteria, and designated uses that define the endpoint of interest to society (modified from NRC 2001). Pollution includes any impact to the chemical, physical, and/or biological integrity of a waterbody, whereas pollutants are specifically defined parameters in the CWA (after Karr and Yoder 2004). Typical CWA approaches have focused on ambient pollutant levels (at A), which occupy a position more removed from the actual waterbody condition than biological approaches, which measure the condition at B.*

Combining the BCG with pattern recognition as a biotic assemblage changes through both space and time enables the refinement of biological thresholds and endpoints within the context of aquatic life uses. Biological condition gradients make it easier to communicate assessment results in numbers and words to the public and policymakers and in legal and regulatory proceedings, where confidence intervals or hypothesis testing are required. The BCG framework also helps to document impacts of restoration efforts. If the framework is sufficiently developed with adequate technical rigor, five to six non-overlapping categories of biological condition should be distinguishable in a variety of assemblages, including fish (Fore et al. 1994) and invertebrates (Doberstein et al. 2000).

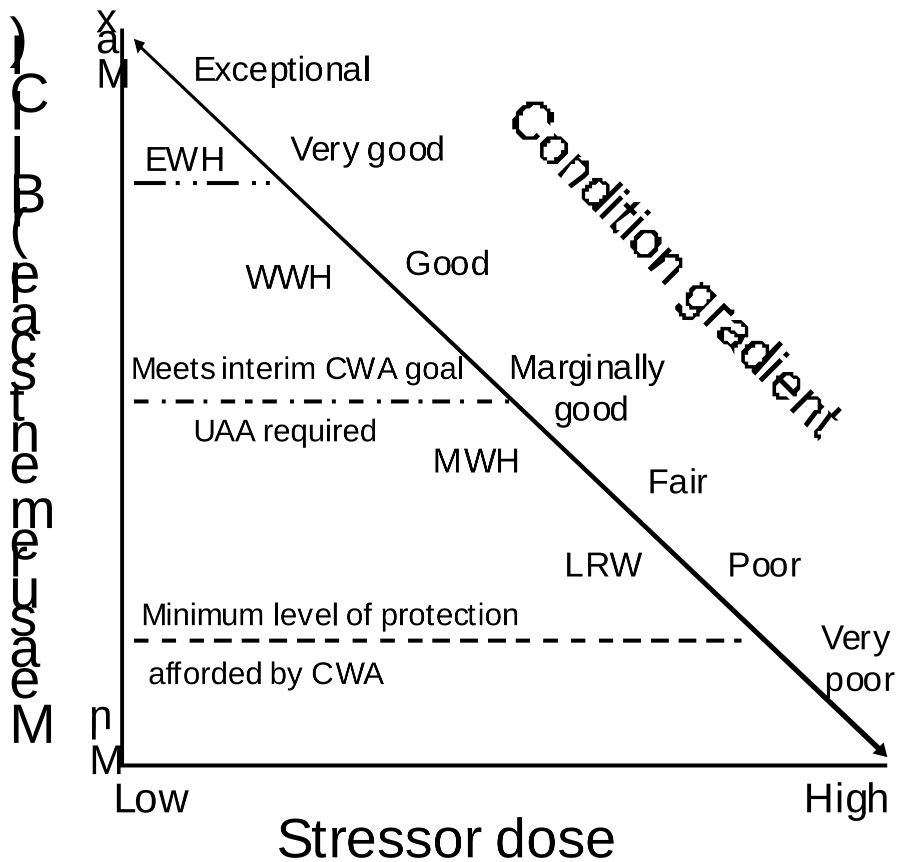


Figure 4. *Relationship between stressor dose and biological measurement scale, such as index of biotic integrity (IBI) or invertebrate community index (ICI), showing level of biological condition (exceptional to very poor) and associated aquatic life designated uses; (exceptional warmwater habitat (EWH), warmwater habitat (WWH), modified warmwater habitat (MWH), and limited resources waters (LRW)) as defined by Ohio EPA and as codified in the Ohio WQS (after Karr and Yoder 2004).*

Adequate Monitoring & Assessment and Multiple Management Needs

Fully realizing the benefits of a biological assessment also requires an understanding of the multiple uses of the information in the management of water resources. A fundamental tenet of adequate monitoring and assessment (Yoder 1998) is that the same set of core resources, methods, standards, data, and information should support multiple water quality and water resource management needs (Figure 5). It also requires a commitment to monitoring program maintenance and upkeep (i.e., maintenance of adequate resources, facilities, and professionalism) over the long term. Sustaining such an effort for the UMR is a major implementation “next step” that will need to be addressed.

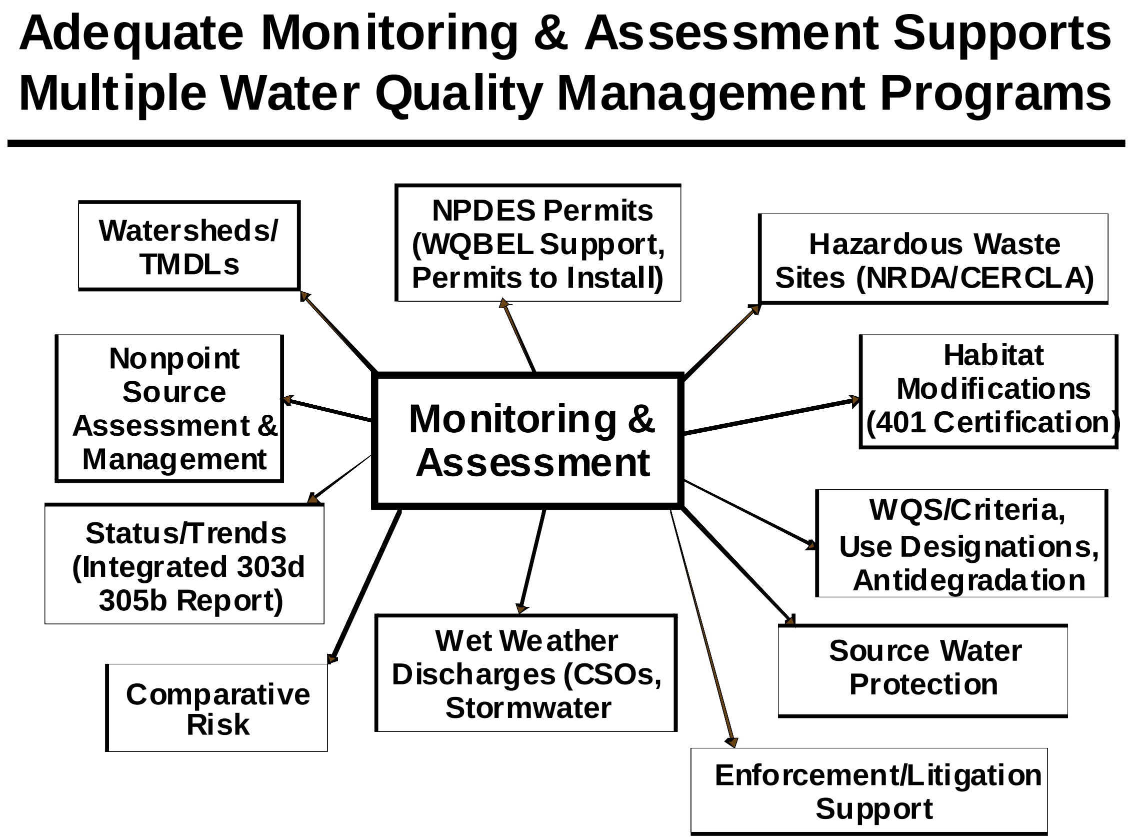


Figure 5. *Adequate monitoring and assessment should be capable of supporting multiple program support needs with the same core base of indicators, parameters, and designs.*

Professionalism includes the qualifications of the monitoring and assessment personnel and their ability to carry out all tasks, including data analysis and the sequencing and interpretation of multiple indicators. Some of the indicators require specialized expertise in terms of data collection, field observations, laboratory methods, taxonomic practice, and data analysis and interpretation skills. Thus the professional qualifications of the personnel who execute and manage a systemwide program is a pivotal issue.

The spatial aspects of a monitoring and assessment design is equally important and for a major mainstem river a longitudinal pollution survey design seems the most useful in terms of fulfilling the current CWA assessment objective, but also for providing maximum utility for future uses of such data and their attendant analyses. As will be discussed later in this report, In terms of currently available approaches we have two major programs to consider (EMAP-GRE and LTRMP) and the EMAP-GRE design seems to fit these needs the best.

Critical Technical Elements of Bioassessment Programs

In cases where biological assessment is being undertaken, there is a need to systemically and objectively evaluate the bioassessment program(s). The *Critical Technical Elements of Bioassessment Programs* was developed with U.S. EPA's support to fulfill this need (Yoder and Barbour 2009). While this methodology has mostly been applied in the context of a statewide program, it can be used to evaluate the technical capacity of a particular bioassessment methodology or protocol. Given the implications for using bioassessment on the UMR beyond the minimal status reporting required by the CWA, a critical technical elements review of the two systemwide programs (LTRMP and EMAP-GRE) was conducted as part of this project, and will be further discussed in Chapter 4. The five WQTF member states have each also been the subject of prior critical technical elements evaluations, but these focused primarily on their inland streams and rivers, not the UMR mainstem.

Although a variety of technical approaches and methods are used throughout the U.S., it is the purpose of the critical elements process to evaluate and reveal the overall level of rigor of a state or tribal (or, in this case, systemwide) program. While it is recognized that different technical approaches can achieve similar levels of rigor, it is likely that a review of the critical technical elements will reveal some differences between such different approaches. Such differences can significantly affect the accuracy and comprehensiveness of biologically based assessments of resource condition and status.

The critical technical elements process is a product of U.S. EPA's national biocriteria program (Yoder and Barbour 2009). The overall goal of the process is to determine the technical rigor of a state or tribal bioassessment program to support multiple tasks that are essential to developing and implementing a "TALU (Tiered Aquatic Life Uses) based approach"[3]. In 2000 U.S. EPA convened a process that produced a general framework and detailed technical and implementation guidance to states and tribes for using biological data to achieve two objectives: 1) refine designated aquatic life uses based on numeric biological criteria, and 2) integrate criteria and uses within a monitoring and assessment program that is designed to support multiple water quality management program needs (U.S. EPA 2005). This process also revealed a need to review and evaluate the technical approaches being employed by states and tribes, and establish a baseline for determining what types of improvements these approaches would need in order to attain the two objectives. Hence the critical technical elements process was developed and tested as a pilot with selected EPA Regions and their states. Since 2004 a total of 22 state and one tribal programs have been evaluated, 12 of these on multiple occasions. The evaluation of the EMAP-GRE and LTRMP included in this project constitute the first systemwide application of this process.

The critical elements consist of 13 technical attributes of a biological assessment program, grouped into three distinct areas: program design, methods, and data interpretation (Table 3). The result of the critical elements evaluation is a determination of the overall level of rigor with level 4 being the most desirable and effective for supporting the multiple management issues that are common to state or tribal water quality management programs. Levels 1-3 are suitable for fewer support tasks with level 2 being amenable for general status reporting only. Written

[3] The TALU based approach includes tiered aquatic life uses based on numeric biological criteria and implementation via an adequate monitoring and assessment program that includes biological, chemical, and physical measures, parameters, indicators and a process for stressor identification.

feedback is provided to each state or tribe via a technical memorandum and a critical elements "checklist" that describes the status of each technical element and what is needed to improve those that are below the highest possible score. Ongoing developmental efforts within the state or tribe are especially highlighted as to their potential to affect the current status of each element and the overall level of rigor within an anticipated time frame.

Table 3. The 13 critical technical elements with the scoring ranges for each element (after Yoder and Barbour 2009).

			LOWEST	→	→	HIGHEST
Design	1.	Index Period	1.5	2.5	3.5	4.5
	2.	Spatial coverage	1.5	2.5	3.5	4.5
	3.	Natural Classification	2	3	4	5
	4.	Criteria for reference sites	1	2	3	4
	5.	Reference conditions	2	3	4	5
Methods	6.	Taxonomic Resolution	2	3	4	5
	7.	Sample collection	2	3	4	5
	8.	Sample processing	2	3	4	5
	9.	Data Management	2	3	4	5
Assessment	10.	Ecological attributes	1.5	2.5	3.5	4.5
	11.	Biological endpoints	1	2	3	4
	12.	Diagnostic capability	1	2	3	4
	13.	Professional review	1.5	2.5	3.5	4.5
		Total Score	**21**	**34**	**47**	**60**

Desired Characteristics of UMR Biological Assessment

In light of the desired elements of biological assessment described in the preceding sections, biological assessment applied to the Upper Mississippi River for CWA purposes should have the following characteristics:

- Incorporates systematic monitoring and assessment that includes biological measures as the primary response indicator – this strengthens the ability to connect stressor, exposure, and response gradients (Figure 2).
- Provides a more complete basis on which to frame water quality standards (WQS) that relate more directly to the stressors that management programs are concerned about (Figure 3) and utilizes the biological condition gradient (Figure 4) as tool to communicate the relationships between stressors and biological condition.
- Incorporates biological data which are closer to the endpoint outcome of concern than are chemical criteria (Figure 3) – thus constituting a powerful complement to current chemical and physical approaches.

- Uses a scalable, site-based approach that allows for assessment at multiple scales including site, reach, and river-wide and along the continuum of the UMR from upriver to downriver.
- Is supported by an assessment program shown to demonstrate level 4 through the critical technical elements evaluation, which includes a commitment to program maintenance, upkeep, and professionalism.
- Additionally, with an emphasis on biological endpoints for determining status, it becomes easier to discern and visualize the biological benefits of management and restoration actions.

Chapter 4: Identification of Readily Available UMR Biological Protocols and Data

A central goal of this project is to review existing biological monitoring programs to determine their value in a future, biology-driven UMR aquatic life use assessment. Table 1 in Chapter 2 includes a summary of all current or recent monitoring efforts on the UMR, including those with biological components. However, not all of the programs in Chapter 2 include the following key elements for a CWA bioassessment program with system-wide applicability:

1. Collects biological data on the UMR main channel;
2. Utilizes methodologies that produce sufficiently rigorous assessments in order to support multiple management programs;
3. Has results applicable to the entire UMR main channel, or a large portion thereof; and,
4. Forms the basis in terms of methods, logistics, indicators, and design for a sustained bioassessment of the UMR main channel.

The inherent need for a system-wide approach to fulfill the primary goal of a system-wide bioassessment assessment informs us about which programs can be considered as viable candidates for establishing a UMR-wide bioassessment program. As such we highlight here the two leading programs – EMAP-GRE and LTRMP – identified by and further described in this project's Background and Scoping Report (Yoder et al. 2010), the results of the critical technical elements evaluations of these programs, and then provide a description of and "readiness" of their respective assemblage assessment tools in light of the four project criteria listed above.

Leading UMR Biological Assessment Programs

U.S. EPA EMAP-GRE

The Environmental Monitoring and Assessment Program (EMAP) is a research program run by U.S. EPA's Office of Research and Development to develop the tools necessary to monitor and assess the status and trends of national ecological resources. EMAP collected field data from 1990 to 2006 (2004 to 2006 only for the UMR). EMAP's goal was to develop the scientific understanding for translating environmental monitoring data from multiple spatial and temporal scales into assessments of current ecological condition and forecasts of future risks to aquatic resources. EMAP aimed to advance the science of ecological monitoring and ecological risk assessment, guide national monitoring with improved scientific understanding of ecosystem integrity and dynamics, and demonstrate multi-agency monitoring through large regional projects. EMAP developed indicators to monitor the condition of ecological resources. EMAP also investigated designs that addressed the acquisition, aggregation, and analysis of multi-scale and multi-tier data.

The EMAP-GRE program dataset includes information collected from the UMR in 2004-2006. Sampling was conducted in the main channel and during a summer-early fall seasonal index period over three years (2004-6). There is a longitudinal aspect available in the EMAP-GRE data that is very important to the goals of this project as the entire UMR was assessed, using a stratified random sampling approach, with multiple sampling sites in each navigational pool of the interstate and non-interstate UMR from St. Anthony Falls, MN downstream to the confluence with the Ohio River. A total of 123 discrete sampling sites were sampled following standardized protocols and the result is a fairly complete coverage of the entire main channel UMR.

Biological data was collected for multiple assemblages including fish, macroinvertebrates, algae (periphyton and phytoplankton), submersed aquatic vegetation, and zooplankton. Water quality parameters were also collected at each sampling location and included chlorophyll-conductivity, secchi depth, total suspended solids, total nitrogen, total phosphorus, turbidity, and water temperature. Together, these comprise important components of an adequate monitoring and assessment approach.

USACE EMP-LTRM Program (LTRMP)

The Long Term Resource Monitoring Program (LTRMP) was authorized by Congress in 1986 as part of the U.S. Army Corps of Engineers Upper Mississippi River System Environmental Management Program (EMP). The program was designed primarily to support natural resource management needs by developing a better understanding of the UMR ecosystem and monitoring long term ecological changes. Key indicators have included fish, macroinvertebrates, aquatic vegetation, and water quality. The USGS Upper Midwest Environmental Sciences Center (UMESC) participates in the implementation of LTRMP. USGS has executed cooperative agreements with each of the five UMR states whereby the state natural resource agencies provide the staffing of the LTRMP field stations. There are five such field stations on the UMR at Lake City, MN; Onalaska, WI; Bellevue, IA; Brighton, IL; and Jackson, MO. From these field stations, biological and water quality monitoring is accomplished in four navigation pools (Pools 4, 8, 13, and 26) and in the Open River (OR) reach. The current monitoring design combines fixed site sampling, at approximately 120 main channel, backwater, and tributary sites, with stratified random sampling across entire pools. As such, the LTRMP takes a representative approach by sampling selected pools and reaches in a very intensive manner.

Biological data currently includes fish and submersed aquatic vegetation as routine assemblages and developmental work with additional assemblages such as freshwater mussels. Chemical/physical data are collected for 25 variables, including conductivity, dissolved oxygen, turbidity, temperature, and pH, which are typically measured *in situ*; and parameters such as total suspended solids, volatile suspended solids, nitrogen, phosphorus, and ammonia, among other chemical constituents, which are measured by laboratory analysis.

A brief overview of EMP-GRE and LTRMP programs is provided in Table 4.

Table 4: Overview of EMAP-GRE and LTRMP.		
	EMAP-GRE	***LTRMP***
Monitoring Components	Biological, Chemical, Physical	Biological, Chemical, Physical
Biological Assemblages	Fish, Macroinvertebrate, Periphyton, Phytoplankton, Submersed Aquatic Vegetation	Fish, Macroinvertebrate (through 2004), Submersed Aquatic Vegetation
Sample Site Type	Stratified Random Sampling	Stratified Random Sampling and Fixed Site Sampling
Monitoring Sites	123 stratified random sampling sites	Component Dependent with: - Fish between 160-200 sites per pool, mix of stratified random and fixed sites. - Vegetation with 450 stratified random sites for Pools 4, 8, and 13 only.
Spatial Coverage	Entire Length of UMR, Main Channel (Border) Only	Pools 4, 8, 13, 26, and Open River, All Strata
Sample Site Size (Fish)	1.0 km	0.2 km
Monitoring Duration	2004-2006 (non-ongoing)	1986-present (ongoing)
Samples Collected By	Field station personnel	Field station personnel

Critical Technical Elements Evaluation

Based on our review in the Background and Scoping Report it was concluded that the two programs with the biological assessment protocols and databases that provide the most potential for an initial biological assessment of the UMR main channel are the U.S. EPA EMAP-GRE program and the USACE LTRMP. Our reasoning for this conclusion is that both include the requisite interstate coverage, both include multiple biological assemblages, and both have sufficiently rigorous protocols as identified by a critical technical elements evaluation.

The USACE LTRMP and U.S. EPA EMAP-GRE were evaluated during on-site visits on July 14-15, 2010 and August 15, 2010, respectively. Key program and technical staff were present and day two at LTRMP included observing and participating in field sampling. The LTRMP was determined to be a Level 3+ (91.7%) and EMAP-GRE was determined to be at Level 4 (95.8%), each of which reflects the requisite rigor within each program's protocols and databases for supporting multiple water quality and water resource management needs.

The slightly lower score for LTRMP reflects only minor departures in some of the interpretive elements, which is understandable since it was not initially designed to deliver those functions.

However, the sufficiency of the design and methods mean that the database would be sufficient to support the elevation of those elements with a short-term development effort. EMAP-GRE being already oriented to providing a CWA type of assessment “naturally” scored higher in some of the elements where LTRMP scored lower, primarily due to the efforts devoted to the development of GRFIn and GRMIn and the accompanying data analyses that effort required. Had LTRMP been similarly tasked then the CE score would likely have attained Level 4. In any event, LTRMP has the baseline infrastructure to support a CWA style of assessment, but it would require making some adjustments and supplements to the current monitoring program design.

Whether the WQTF ultimately decides to utilize these programs or adopt a new one entirely, we recommended that any candidate design undergo a critical technical elements evaluation. The critical technical elements matrices for these two programs appear in Appendix B.

Supplementing System-Wide Programs with State Programs

Although our analysis focused primarily on system-wide programs, one potential approach is also to make the fullest possible use of the elements of existing state or locally focused programs. Selected state programs can offer supplemental data to the overall assessment objective. The Wisconsin DNR Non-wadeable Rivers (NWR) sampling program is an example of a potentially applicable statewide program. However, preliminary analyses conducted by Dukerschein et al. (in preparation) suggest that the differences among the LTRMP, EMAP-GRE, and Wisconsin DNR-NWR methodologies are such that care will need to be taken when integrating data collected by these different entities. We suggest that similar analyses be undertaken prior to considering how other programs can be used.

Available Assemblage Assessment Tools

The following is a brief description of the biological assemblages that are presently sampled by the various entities that are conducting or have recently conducted bioassessment in the UMR main channel (see Table 1). By biological assemblage we are referring to distinct taxonomic groups (e.g., fish, macroinvertebrates, submerged aquatic vegetation, algae, mussels, etc.) that are sampled to collect all species and taxa that are present. It excludes biological sampling that is aimed at single species, i.e., those of specific management concern. Methods are critical, as they determine the ability to collect enough taxa and organisms to allow the bioassessment mechanisms to function properly. In turn, the structure of the assessment mechanism (e.g., an IBI type of index) is equally important in determining the power of the assemblage assessment to not only detect impairments, but to incrementally portray condition along the scale of the BCG. Finally, the spatial intensity and coverage is critical to recognizing the magnitude and severity of impairments and in developing the stressor:response relationships.

Fish are the most commonly sampled biological assemblage among all of the different UMR programs, with the majority of entities performing some type of fish sampling. Macro-invertebrates are sampled by at least two entities (EMAP-GRE and MCES) and represent a commonly used second assemblage. Submersed aquatic vegetation is a promising third assemblage that is assessed by LTRMP and by EMAP-GRE. Algae include periphyton (EMAP-GRE) and phytoplankton (LTRMP and EMAP-GRE). Assessment tools for other assemblages, such as freshwater mussels, are in development. While the current thinking is that adequate biological monitoring programs must include at least two assemblages, it will likely be necessary to eventually have at least three assemblages for a resource as large and complex as the UMR.

Fish

The specific methodological characteristics for the six entities that either have or currently sample the fish assemblage on the UMR are detailed in Appendix B. Fish methods are usually the most comparable in terms of sampling gear and the type and resolution of data that are recorded. The differences usually occur in the sampling protocol and the execution of the sampling, and are issues that need to be examined prior to identifying which programs will be the most suitable to meet the goals of this project. Of the programs listed in Table 1, the LTRMP and EMAP-GRE are the geographically most extensive and as a result provide the most promise to at least initially serve as the basis for ongoing, systemic, biologically based CWA assessment of the UMR main channel (acknowledging that EMAP-GRE is not an ongoing program). Each offers system-wide coverage as well as methods that are amenable to the development and application of a fish IBI. This is not to say that the state-specific efforts will not be useful; in fact, they provide the most potential to supplement the system-wide assessment and support potentially more detailed state needs.

Currently available indices that have been applied to the UMR include the Wisconsin DNR Index of Biotic Integrity (IBI: Lyons et al. 2001) and the Great Rivers Fish Index (GRFIn) by the EMAP-GRE program. Of these, the EMAP-GRE index is the most calibrated to a system-wide application, and presently distinguishes between the impounded and open river portions of the UMR. A regional index developed for large river tributaries to the UMR and Ohio River (FACI; Emery et al. 2007) was also available for evaluation and application. Additionally, in response to discussions at the June 2011 work session, indices developed for the Missouri River were also examined to determine their applicability to the open river portion of the UMR.

Macroinvertebrates

Macroinvertebrates are currently assessed by the EMAP-GRE program, the EPA NRSA, and the Twin Cities Metropolitan Council. LTRMP dropped this as an indicator assemblage in 2004. Of these programs, EMAP-GRE is the most geographically extensive and provides the most promise to at least initially serve as the basis for a biologically based assessment of the UMR main channel. Having one program to draw upon also alleviates the need to deal with what can be serious comparability issues that can exist with fundamentally different macroinvertebrate sampling and sample processing protocols. The EMAP-GRE program has the only fully calibrated macroinvertebrate IBI (GRMIn; Angradi et al. 2009) that applies on a system-wide scale; hence, it is the preferred program for this assemblage. A regional index developed for large river tributaries to the UMR and Ohio River (NMACI; Blocksom and Johnson 2009) was also available for evaluation and application.

Submersed Aquatic Vegetation
Submersed aquatic vegetation (SAV) comprises a realistically applicable third assemblage and is currently included by the LTRMP and EMAP-GRE programs. A methodology and assessment process – the Submersed Macrophyte Index (SMI) - has recently been developed (Moore et al. 2011) and is being applied to selected pools of the impounded portion of the UMR. Hence it has recently become available on a partial system-wide basis where SAV occurs in abundance. Some questions remain about the applicability to certain segments of the UMR main channel, specifically the Open River where SAV is not as abundant. It has been the most useful in the segments and pools of the main channel where submersed vegetation is prominent. It was included in the Preliminary Bioassessment Thresholds Report (Miltner et al. 2011) along with fish and macroinvertebrates.

Algae
Periphyton was assessed in the main channel border as part of the EMAP-GRE program, and is also an EPA NRSA assemblage. At this point we have not found a publication that details an assessment tool, although algal IBI type indices are in use for other waterbody types. We expect it will be a potentially suitable indicator assemblage given its widespread usage in other lotic systems throughout the U.S. Phytoplankton is sampled by EMAP-GRE and LTRMP. Again, we have not found an assessment tool specific to this component of the algal assemblage, but they do exist for other waterbody types.

Zooplankton
This assemblage was included in the EMAP-GRE suite of indicators and it has been included as part of the LTRMP. However, we could find nothing about any analysis of this data that would lead to it being considered as a candidate for state usage.

Freshwater Mussels
A developmental project to include a mussel assemblage assessment within the LTRMP in UMR pools 5, 6, and 18 was recently published by Newton et al. (2011). This project utilized the LTRMP stratified random design. Incorporating mussels as a routine main channel biological assessment tool will require additional testing and application, but its potential as an additional assemblage has been enhanced by the recent work of Newton et al. (2011).

Summary of Available Assemblages
Based on our review of the extant bioassessment programs, fish and macroinvertebrates emerge as the two assemblages that are in a comparative state of "readiness" to support a near-term biological assessment for the entirety of the UMR main channel. Submersed aquatic vegetation (SAV) comprises a realistically applicable third assemblage and has only recently become available in terms of a comparative state of readiness. Algae and the remaining assemblage groups are each in various stages of development, testing, and refinement. Each will need to have a readily available assessment mechanism, i.e., a calibrated index that is relevant to the BCG, to serve the goals of this project.

Table 5 provides a summary of available methods and indices for UMR assemblages. Those in bold indicate the preferences of the workgroup at the project's second work session and those investigated to a greater extent in this project's preliminary bioassessment thresholds report (Miltner et al. 2011).

Table 5: Currently Available Methods and Indices for UMR Biological Assemblages.

Indicator Assemblage	Methods	Index
Fish	**LTRMP** **EMAP-GRE** EPA-NRSA WI DNR MN DNR MN PCA IL DNR	**GRFIn (EMAP-GRE)** IBI (WIDNR)
Macroinvertebrates	LTRMP (formerly) **EMAP-GRE** EPA-NRSA IL EPA MN PCA Metropolitan Council	**GRMIn (EMAP-GRE)**
Vegetation	**LTRMP** EMAP-GRE	**Submersed Macrophyte Index** (MNDNR)
Mussels	LTRMP	In Development (Newton & Ziegler 2011)
Algae (Periphyton/Phytoplankton)	EMAP-GRE EPA-NRSA LTRMP	None Known
Zooplankton	LTRMP	None Known

Selection of Preferred Approaches - Applicability of Protocols and Data Sets for a CWA Biological Assessment

The preceding discussion has clearly identified EMAP-GRE and LTRMP as the leading candidates for supporting a UMR CWA biological assessment, particularly in terms of program protocols. Additionally, EMAP-GRE indices (GRFIn and GRMIn) and an EMAP/LTRMP hybrid index (SMI) are the leading candidates for biological assessment tools. With these determinations in hand, the next steps are to consider which program protocols – EMAP-GRE or LTRMP – provide a better fit for a CWA assessment and how the indices perform in a CWA assessment setting. This question can be considered both in the near term, for an initial CWA biological assessment, and in the long term for sustaining an ongoing assessment.

Initial CWA Biological Assessment – Data Set and Examination of Indices
For the purposes of conducting an initial CWA biological assessment of the UMR, the EMAP-GRE data set is preferred over the LTRMP data set for the following reasons:

1. It represents the most spatially comprehensive database providing nearly continuous longitudinal coverage of the UMR main channel.
2. It is based on a standardized sampling of spatially contiguous sites that are the primary origin of the database. This allows upwards aggregation of the data from sites to reach and pool wide applications while retaining the individual sampling site as a key unit of assessment. This allows for the protocol to be used for local and reach scale assessments of specific places and stressors both now and into the future. It further allows for the detection, quantification, and characterization of the aggregate biological assessment responses to pollutional gradients along the continuum of the main channel via the delineation of proximate stressors.
3. It is of relatively recent origin (2004-6) thus using it to develop an initial condition assessment should still be relevant.
4. It includes two *spatially matched* assemblage indicators – fish and macroinvertebrates – therefore supporting a dual assemblage assessment approach.
5. The collective rigor of the methods and data analyses conforms to the highest standards for a bioassessment program (it is a Level 4 program after Yoder and Barbour 2009; see Appendix B).

Because of the preferred characteristics listed above, the EMAP-GRE data set (Angradi et al. 2009a, 2009b) was used in a preliminary analysis of biological condition thresholds for the UMR main channel, as discussed in more detail in Chapter 5 of this report.

The threshold analysis also provided an opportunity to test the ability of the indices – GRFIn, GRMIn, and SMI – to produce meaningful and CWA-relevant assessments of aquatic life condition. As will be discussed in Chapter 5, each index demonstrated both strengths and weaknesses in this regard.

Ongoing CWA Biological Assessment – Preferred Program Protocols
Of the readily available approaches, the EMAP-GRE program protocol not only best supports an initial CWA assessment but also currently provides the best "fit" for ongoing CWA biological assessment. This is primarily due to the reasons listed above – it is spatially comprehensive, has contiguous sample sites allowing for scalability and detection of pollution gradients, includes two assemblages, and is a level 4 bioassessment program. In addition, two of the leading index candidates – GRFIn and GRMIn – are compatible with the data produced using EMAP-GRE methods.

As evidenced by the critical technical elements review, the LTRMP protocol also incorporates many of the desirable attributes for a CWA bioassessment program. However, it is primarily limited in a CWA context by the spatial design of its fish sampling program (see discussion below) and the lack of any current macroinvertebrate assessment protocol. The preference of EMAP-GRE over LTRMP in a CWA context should not, however, be interpreted as a criticism of LTRMP, as its design reflects the specific purposes and goals for which it was created for the USACE EMP (which does not explicitly include a “CWA style” assessment).

The key component in selecting a preferred approach is the sampling site as the fundamental unit of assessment. As stated above, the advantages of the EMAP-GRE sample site approach coupled with the lack of a LTRMP macroinvertebrate assemblage tool, were influential in the identification of EMAP-GRE as a preferred protocol. There has been ongoing work to determine the comparability of the LTRMP fish sampling protocol and achieving “equivalency” with the 1.0 km GRE site protocol (Dukerschein et al., in progress). However, even with this in hand, the spatial make-up of the LTRMP design limits the application of these transformations to reach-wide assessments and does not allow for a site-based assessment. This limits the potential use of LTRMP data to making only broad condition assessments of reaches and pools, and confines the resulting analyses to general statements of condition within those spatial limitations. Also of note, the delineation of a sampling site is an important consideration should states wish to incorporate the SMI into their CWA assessments, as it currently does not rely on the EMAP-GRE delineation of a sample site.

The implications of the choice of the EMAP-GRE style approach for future assessments of the main channel are discussed further in Chapter 7 of this report.

Chapter 5: Determination of Biological Assessment Thresholds

A major emphasis of the project is to identify "potential impairment thresholds for the UMR main channel in determining the attainment of aquatic life uses," as identified in the project's work plan (MBI 2010). To that end, we conducted a preliminary assessment of the derivation of biological thresholds for review by the UMRBA Water Quality Task Force (WQTF) in January 2011. This was a previously unplanned, but crucial step in making progress in the development of the draft guidance, and also responded to the discussion at the project's second work session regarding an initial CWA biological assessment of the UMR. The empirical derivation of biological thresholds was accompanied by a preliminary biological condition gradient (BCG) analysis, which was conducted in order to provide a ecologically-based comparison point for the thresholds. The threshold and BCG analyses constitute perhaps the most technically detailed and unique portions of this project, as they explore in depth the mechanics of incorporating biological approaches into UMR CWA assessment.

Statistical Derivation of Impairment Thresholds

A stand alone report entitled *"Preliminary Analysis of Biological Assessment Thresholds for Determining Aquatic Life Use Attainment Status in the Upper Mississippi River Mainstem"* (Miltner et al. 2011) was produced that details the analyses that were conducted and preliminary results of bioassessment thresholds. As such this report contains options available to the WQTF for a initial CWA bioassessment of the UMR main channel. WQTF members reviewed an initial draft in advance of a January 2011 meeting and discussion. Feedback received from the WQTF in January 2011 indicated an interest in:

1) examining alternate approaches to establishing thresholds using EMAP-GRE indices (i.e., beyond the three condition stressor-based classes developed by GRE);
2) comparing GRE-based thresholds to those developed using alternate approaches and indices, including "reference" condition and "peer river" based approaches;
3) further exploring the sensitivity of the GRE suite of indices and other available indices to the suite of UMR stressor gradients; and,
4) evaluating the possibility of incorporating a submersed vegetation index into a UMR bioassessment.

Technical Approach

We focused primarily on EMAP-GRE fish and macroinvertebrate data because both were collected at similar spatial densities and over the same length of river (Yoder et al. 2010). We did, however, examine the potential utility of a recently developed submersed aquatic vegetation index (SMI), though our analyses were limited by the spatial density and longitudinal coverage of submersed vegetation sampling sites.

The EMAP-GRE indices and thresholds were compared to other indices and thresholds that were deemed potentially applicable to the UMR, examining each for their responsiveness to gradients of aquatic life stressors and their ability to accurately characterize the UMR. This included the Fish Assessment Community Index (FACI: Emery et al. 2007) and the Non-wadeable Macroinvertebrate Assemblage Condition Index (NMACI; Blocksom and Johnson 2009) that were developed for a REMAP project that included large rivers that are tributary to the Upper Mississippi and Ohio Rivers. We used multiple statistical methods to derive potential threshold values for the indices examined, and evaluated the appropriateness of those thresholds in light of aquatic life use attainment realities in the UMR main channel. The primary goals of the thresholds report were to:

1. Examine the sensitivity and suitability of various biological indices for assessing CWA aquatic life use attainment in the main channel of the UMR, including the identification of proximate stressors;
2. Aid the WQTF in visualizing the likely outcomes of different options including a mix of indices and thresholds for a biological-based assessment of aquatic life use attainment for the UMR;
3. Compare the EMAP-GRE developed indices and thresholds to other available approaches to aid the WQTF in its consideration of EMAP-GRE tools as the leading candidates for recommended bioassessment approaches on the UMR; and,
4. Assess the potential for integrating an additional assemblage (submersed vegetation) into a UMR biological assessment.

The rationale for a dual indicator approach is that it enhances the redundancy of the resulting bioassessment because individual assemblages have different temporal responses and sensitivities to various stressors existing along a disturbance gradient. Hence the accuracy of the overall bioassessment is improved with two assemblages. However, the dual assemblage approach is equally important for determining proximate causes of non-attainment. As such we evaluated the sensitivities of the two assemblage indicators, fish and macroinvertebrates, to environmental stressors in the UMR main channel. Optimally, indicators based on different assemblages will each track the overall stressor gradient in a generally similar manner, but will show different sensitivities to the magnitude and severity of the effects of individual stressors along the disturbance gradient.

Methods for Deriving Biological Thresholds

Using EMAP-GRE data, we conducted biological condition assessments of the UMR by using:

1. EMAP-GRE derived thresholds and indices (including Missouri River derived GRE indices);
2. Alternate thresholds for EMAP-GRE indices; and,
3. Regional EMAP (REMAP) indices and thresholds.

Table 6 provides a summary of the scenarios considered in examining biological thresholds. We also examined the effect of integrating a submersed aquatic vegetation index (SMI) into the condition assessment where that data was paired with the fish and macroinvertebrate results.

Table 6. Scenarios considered for determining biologically based aquatic life use thresholds.

Threshold	Indices Used	Rationale/Method	Description
Most Disturbed	GRE	Based on an empirical relationship, the boundary between most and intermediate disturbed should be achievable.	Biological index scores are plotted against a stressor gradient composed of land use, habitat, water quality and other environmental variables; the y-intercept of a quantile regression sets the upper bounds for index expectations, and the 5th percentile the floor. Trisection of the resulting range defines least, intermediate and most disturbed conditions.
Quadrisection	GRE	An *ad hoc* method that sets achievable expectations based on peers.	The 95th to 5th percentiles of index scores for a given river (in this case, scores from the upper impounded reach and the open river) from the ceiling and floor values, and the resulting range is quadrisected. The midpoint or first section boundary can set the threshold for attainment.
Quadrisection of Similar Rivers	REMAP	This method provides a partially independent method for determining a realistically achievable benchmark.	Index scores from similar rivers (size, drainage area, fauna, physical alteration) are quadrisected, and the midpoint sets the threshold.
Reference 25th Percentile	REMAP	Reference condition-based.	Reference sites were identified for the NMACI and FACI based on the 25th percentile of a stressor gradient (defined by environmental variables).
Change Point	REMAP & GRE	Least arbitrary of the methods for identifying an impairment threshold.	Index scores were plotted against the stressor index and a locally weighted regression line fit through the points. A deviance reduction method (classification and regression tree [CART]) was used to partition scores along the stressor index, and the point where the regression line crossed the partition was used as a threshold (see Figure 2).

Assessment Using EMAP-GRE Developed Indices (GRFIn and GRMIn) and Thresholds

Use of EMAP-Derived Thresholds: The EMAP-GRE program had derived biological condition thresholds for the GRFIn and GRMIn indices based on biological responses against an empirical stressor gradient constructed from land use, population density, habitat quality, and water chemistry indicators (Angradi et al. 2009a). Unlike a reference condition approach where percentiles of a least or minimally disturbed reference population form the basis for a set of condition thresholds, the empirical approach sets the baseline to the y-intercept in the relationship between the empirical stressor gradient and the respective biological assemblage endpoint (i.e., fish or macroinvertebrate assemblage index or attribute). Once the upper limit is defined (i.e., by the 95th percentile regression line y-intercept) the data range was trisected, thus forming three disturbance classes - least, intermediate and most disturbed (re: Figure 7 in Angradi et al. 2009a). An empirical approach was initially used for the main channel UMR because the extent of anthropogenic modifications precludes a direct reference condition approach - essentially the navigable mainstem from the Twin Cities to St. Louis is a series of regulated, modified navigational impoundments and a highly modified open river segment.

Initial Choice of "Most Disturbed" Threshold: In our analysis, we initially considered sites to be failing to meet the minimum CWA goal for beneficial aquatic life uses if either the fish or macroinvertebrate indicators were less than their respective EMAP-derived, most disturbed threshold. Sites with both assemblage indices classed at or above intermediately disturbed were considered in full attainment of the minimum CWA aquatic life use goal.

Site-Based Assessment: Assessments were made at the site level to specifically track longitudinal trends in status. Attainment status for the UMR from the confluence with the Ohio River upstream to St. Anthony Falls, MN was apportioned to miles of attainment/non-attainment by UMR assessment reach and by state using the weighted probability baseline of the original EMAP-GRE design and as an a linear interpolation of attainment status by sampling site.

Alternative Calculations: We also conducted alternative analyses using least disturbed as the CWA threshold, and a third comparing a distribution of index scores from the lower St. Croix River as a least impacted analog for the UMR main channel.

Potential Causes of Non-Attainment: Potential causes of non-attainment were identified by examining associations between biological index scores and ambient water chemistry, habitat, and stressor indicators (e.g., population density, upstream distance to wastewater treatment plants, etc. all based on data supplied by Dave Bolgrien, U.S. EPA). Statistical methods included rank-order correlations, classification and regression trees (CART), linear regression, and discriminant analyses. For the latter, condition classes formed the categorical variables, and only data from the impounded UMR were selected to allow comparisons between the two biological assemblages. Data from the unimpounded UMR (Open River) were not included because the GRMIn is not calibrated for that section of the UMR. Rank-order correlations revealed broad patterns in the data. Multiple regression goes a step further in assessing combinations of variables contributing explanatory power. CART models may reveal nonlinear and threshold responses undetected by linear regression, and discriminant analysis helps assess the results of the previously listed methods in terms of how well the stressors contribute to forming narrative classes.

Assessment Using Alternate Thresholds for GRE Indices

Quadrisection: Quadrisection was used as an alternate method to define an assessment threshold, wherein the 95th and 5th percentiles of GRFIn and GRMIn scores (GRE Indices) from the entire UMR (the upper impounded reach and the open river) set the ceiling and floor values, and the midpoint of the resulting range set as the impairment threshold. Data from the entire UMR were then included to expand the potential stressor gradient. The midpoint of the quadrisection was chosen as the impairment threshold because it sometimes functions in a

CWA context as a boundary between fair and good narrative quality (all internally-derived boundaries are necessarily arbitrary).

Change Point: Another method for defining an impairment threshold entailed arraying GRFIn and GRMIn scores along the stressor gradient (Angradi et al. 2009) and testing for a change point via the method of deviance reduction (i.e., the first split from a classification and regression tree [CART]). A locally weighted smoothing line (LOWESS, α=0.5) was fitted to the plot of GRFIn and GRMIn scores against the stressor index to find the respective index score corresponding to the change point . In both cases, the change point method resulted in a threshold similar to either the one given by quadrisection (GRFIn), or the existing disturbance threshold (GRMIn), and was therefore not used to tally impairment, as it was redundant.

Best UMR Reaches 2 and 3: An approach utilizing a "best of UMR" sites approach included basing trisections and quadrisections on GRFIn and GRMIn scores in reaches 2 and 3. These reaches consistently had the highest index scores and the rationale for this alternative would be that it demonstrates the potential for other UMR reaches with perhaps the exception of the open river reach. Figure 6 is an example of how the sections were developed with the suggested thresholds for impairment. Four options were developed using combinations of the GRFIn and GRMIn and GRFIn and SMI.

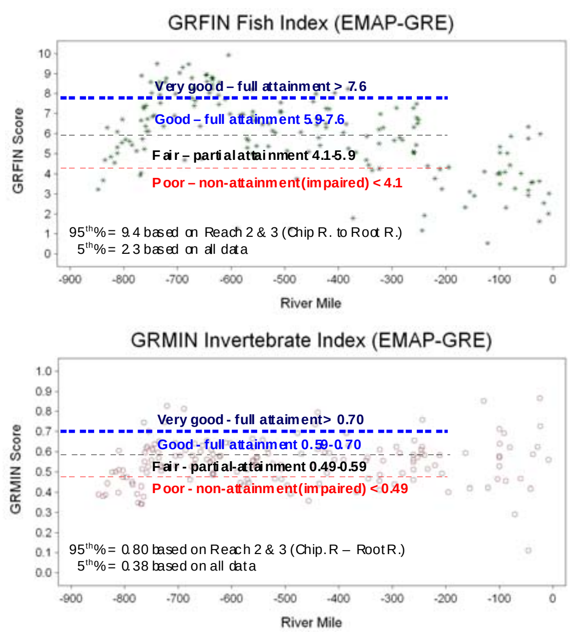


Figure 6. *Potential threshold conditions based on fish (GRFIN) and macroinvertebrate (GRMIN) index scores for EMAP-GRE data (2004-2006). Attainment classes are quartered values of the 95th percentile of Reach 2 & 3 minus the 5th percentile of all data.*

Assessment Using REMAP Indices (NMACI and FACI)
Indices Used: A separate set of biological indices developed for large Midwestern rivers via a Regional EMAP (REMAP) project (Emery et al. 2007) were applied to the UMR data to help identify a disturbance threshold, and to provide a second set of condition assessments for comparison to the condition estimates given by the GRFIn and GRMIn indices. The indices, the Non-wadeable Macroinvertebrate Assemblage Condition Index (NMACI; Blocksom and Johnson 2009) and the Fish Assemblage Community Index (FACI; Emery et al. 2007), were developed from sites sampled during the REMAP Large Rivers project.

Thresholds Used: Potential thresholds for each index were created by in the following ways: 1) quadrisecting scores calculated from the UMR data, 2) the 25th percentile from reference sites, (the 25th percentile of reference sites for the FACI was estimated from Figure 12 in Emery et al. [2007], and for the NMACI, from Figure 4 in Blocksom and Johnson [2009]), 3) the change point against the stressor index, and 4) quadrisecting FACI and NMACI scores from similar large Midwestern Rivers[4].

Assessment of Submersed Macrophyte Index (SMI)
Submersed aquatic vegetation is a third assemblage that recently became available for assessment reaches 0 through 5 of the upper impounded UMR. The Submersed Macrophyte Index (SMI[5]) was developed based on sampling main and side channels (Moore et al. 2011). A threshold for the SMI was established by quadrisection, using the midpoint of the range between the 95th and 5th percentiles. Because the macroinvertebrate indicators tended to have a limited range of response, the SMI was experimentally used within the threshold analyses to over-ride condition assessments that were rated as not attaining based on GRMIn or NMACI scores for sites where the SMI was available. The over-ride approach was used because the spatial coverage of the SMI was limited compared to that for the fish and macroinvertebrate assemblage, and was therefore assumed to be a supplemental indicator for the purposes of the analyses, rather than a substitute or third assemblage.

Preliminary Thresholds Report Conclusions
The thresholds report examined the potential applicability of existing biological indices in making UMR CWA assessments. As described in the preceding section, the EMAP-GRE indices developed specifically for the UMR (GRFIn and GRMIn) and the lower Missouri River were examined in detail and resultant thresholds compared to those from a REMAP-developed suite of large river indices (FACI and NMACI). The potential utility of a vegetation index (SMI) was also examined. The following general conclusions can be drawn from the examination of these indices:

[4] The Saint Croix River below Taylor Falls, the Wisconsin River below Lake Wisconsin, the Minnesota River downstream from New Ulm, the Wabash River downstream from the confluence with the Vermillion River, the Illinois River, and one impounded site on the Muskingum River (FACI only).
[5] SMI data were supplied by Heidi Langrehr, Wisconsin DNR.

- Both GRFIn and the GRMIn hold promise for CWA aquatic life use assessment on the UMR as both are calibrated specifically for the UMR. This holds especially true for the Impounded Reaches for both indicators, and for the GRFIn only in the Open River reach. The development of an Ad Hoc macroinvertebrate index holds promise for improving the GRMIn (see below) throughout the UMR.
- GRFIn is responsive to the different types and gradients of stressors in the UMR main channel, and appears to track a wide range of condition. The Ad Hoc Index improved the range of response.
- The GRMIn suggests a more narrow range of condition in the UMR compared to the GRFIn, and tended to be less responsive to measured environmental variables compared to the GRFIn. However, the GRMIn compliments the GRFIn by responding to different stressors (i.e., GRFIn was more responsive to habitat stressors, GRMIn to selected water quality stressors). The GRMIn needs to be tested via a longitudinal survey against a local, known stressor (e.g., a reach subject to combined sewer overflows or other such severe point source impacts) to evaluate whether the apparent narrow response is a function of the overall condition of the UMR (as read by the GRMIn) or a true limitation of the GRMIn. Individual components of the macroinvertebrate community were more responsive to environmental stressors than the composite GRMIn index, suggesting that structural improvements to GRMIn may improve its utility for UMR bioassessment. The Ad Hoc index developed as part of the thresholds report is a good first step in fulfilling that need.
- The REMAP derived FACI was responsive to the range of stressors in the UMR much the same as GRFIn, and it correlated well with the GRFIn in the UMR. However, since it was derived from smaller, albeit large rivers, its application in the UMR needs to take that difference into account. This is particularly true for the OR reaches that may be outside of its derivation and calibration domain.
- The REMAP NMACI was the least responsive of all the indices examined against the stressor gradient for the UMR main channel, and is therefore not recommended for use at this time.
- SMI was sensitive to a number of different stressors and therefore represents another candidate assessment tool that can be used either with or in lieu of the GRMIn. However, because a direct evaluation of the SMI in a dual indicator approach was limited by spatial differences between where vegetation and fish or macroinvertebrates were sampled, further study is encouraged to demonstrate the utility of the SMI to the remainder of the impounded UMR.
- The Missouri River scoring was applied to the Open River and resulted in less variation in GRFIn and GRMIn scores, suggesting that component metrics of the Missouri River indices could supplement or replace existing OR Mississippi River GRFIn or GRMIn metrics. The metrics for the Missouri River GRMIn have several in common with the Ad Hoc index. These options are further described in Table 8b.

Evaluating Bioassessment Threshold Options

As described above, various statistical threshold determination approaches (empirical/ trisection, quadrisection, 25% of REMAP reference, best available "peer" examples, best of UMR, change point) were applied using the GRFIn, GRMIn, FACI, and NMACI indices (Miltner et al. 2011). The following conclusions can be drawn from the examination of these threshold determinations:

- All internally-derived threshold analyses produced >50% non-attainment for the UMR main channel as a whole. This result is not surprising given that threshold scores for all indices as a percentile of their range in the UMR were near the median or higher, except for the most-intermediate disturbance thresholds for the GRMIn and GRFIn. This is an obvious change in the characterization of the UMR given that only 4 of 13 interstate assessment reaches are currently reported to have aquatic life use impairment using the current non-biological approach.
- Threshold calculations based on best available conditions from other rivers (i.e., NMACI & FACI 25th % of "reference") result in greater proportions of the UMR in non-attainment (>80%) and do not show good agreement on condition status between fish and macroinvertebrates. As such, these thresholds are not reasonably attainable and therefore they are not effective approaches for a UMR CWA assessment.
- The percentages of non-attainment from all other threshold approaches ranged from 24.3% to 72.9% non-attainment. These approaches may offer more plausible options for threshold determination.
- Among the more plausible options, the EMAP-GRE most disturbed/intermediate disturbed threshold presents a viable option for an "upper tier" threshold, as it produces an attainable upper tier goal in the context of comparing the UMR to its nearby "peer" rivers. Similarly, another suitably protective "upper tier" threshold (e.g., the 25th or 50th percentile) can be identified from the Lower St. Croix. These thresholds may be appropriate for establishing tiered biocriteria as a forerunner of tiered aquatic life uses (TALU) for the UMR.
- Given the uncertainties with the current GRMIn index, a workable minimum threshold is suggested by comparing the UMR to the wider collection of its regional peers. Application of this threshold (at the 16th percentile) to the GRMIn is recommended as a default until the uncertainties with the GRMIn can be sorted out and would still result in a condition assessment for the UMR comparable to (i.e., the lower St. Croix) or better than that of the other regional rivers.
- The analyses of threshold choice in the thresholds report were largely been driven by the comparison of outcomes (i.e., non-attainment percentages) across threshold options to determine attainability and comparability to other systems. As such, it seems useful to integrate these outcomes with the BCG analysis (Rankin and Yoder 2011) for the selection of CWA threshold for the UMR, as this provides an ecologically-based "check point" on the empirically-derived threshold choices.

In summary, among the more plausible options, the EMAP-GRE most disturbed/intermediate disturbed threshold and a percentile of the lower St. Croix River offer two viable options for "higher-end" thresholds, as each produces an attainable goal in the context of comparing the UMR to its nearby "peer" rivers. However, given the uncertainty in the current GRMIn index, a workable lower-end threshold is suggested by comparing the UMR to the wider collection of its peers and this scenario was therefore added for consideration (as reflected in Table 8a). Application of these two thresholds – a "higher end" and a "lower end" - may be appropriate for establishing tiered bioassessment thresholds as a forerunner to tiered aquatic life uses (TALU).

The UMR Open River (assessment reaches 12 and 13) represented a challenge in that it is substantially different from the impounded UMR (assessment reaches 0-11) and is a highly modified river channel in terms of physical and flow characteristics. As such the GRE developed different stressor gradients and fish and macroinvertebrate IBIs for the OR reaches as a result. In addition to using the GRE derived indices for the OR, we also applied the lower Missouri River GRFIn and GRMIn to this part of the UMR based on suggestions received at the June 2011 working session. The concept herein is that the Missouri River so substantially influences the UMR that it retains similar physical and flow characteristics, thus applying the GRE indices developed for the lower Missouri may well have better applicability than the UMR OR derived indices. In addition, the calibration basis of the Missouri river indices is broader when compared to the OR. This addresses concerns that the OR derived indices are too narrow to function properly. This notion is supported by the finding of the BCG analysis for the OR reach, as described in the following section. Thus several of the options included in the thresholds report included different iterations using the lower Missouri River GRFIn and GRMIn (see Table 8b).

Addressing Uncertainties Regarding Threshold Options

Any attempt to develop a threshold for biological measures in response to the intent of CWA Section 101[a][2] for the ". . . protection and propagation of fish, shellfish, and wildlife" is necessarily seen as having an inherent level of subjectivity and judgment. However, we would assert here that a well organized and developed empirical process can aid in setting such thresholds.

We took two different approaches to reach this point in the project regarding the empirical approach. The statistical derivation of numeric thresholds or biocriteria as just described was attempted using a number of different combinations of biological indices that included both "externally" and "internally" driven methods. In addition, to augment the empirical approach, we undertook an initial biological condition gradient (BCG) analysis for the UMR.

Biological Condition Gradient for the UMR

We developed a Biological Condition Gradient (BCG; Davies and Jackson 2006) approach as a fundamentally different way of determining the current status of the UMR fish assemblage compared to historical conditions (Rankin and Yoder 2011). The BCG is a conceptual, but scientifically based model of biological response to the increased effect of stressors and is specific to the aquatic resource type of interest (e.g., wadeable streams, large rivers, great rivers, wetlands, etc.). It was developed in response to prior difficulties in communicating effectively about the ecological meaning and management relevance of different quantitative measures of ecological condition. The BCG encompasses the complete range, or gradient, of aquatic resource condition from "as naturally occurs", e.g., undisturbed or minimally disturbed conditions, through increasing levels of alteration to severely degraded conditions. It describes changes in 10 ecological attributes along the BCG that respond to increasing levels of stressors. The BCG is divided into six condition levels, with level 1 representing natural, or undisturbed conditions, level 6 representing severely degraded conditions, and the other levels representing conditions in between (Figure 7). The ecological attributes for each BCG level are characterized by how each is expected to change as biological condition transitions to successively stressed levels. The intent was to tailor the make-up and response of the ecological attributes to the system in question. The UMR BCG analysis was restricted to the fish assemblage out of necessity and is an example of what might also be accomplished for the other preferred assemblages.

Using the BCG to Evaluate Threshold Options

Setting reasonably protective and attainable CWA attainment thresholds depends on being able

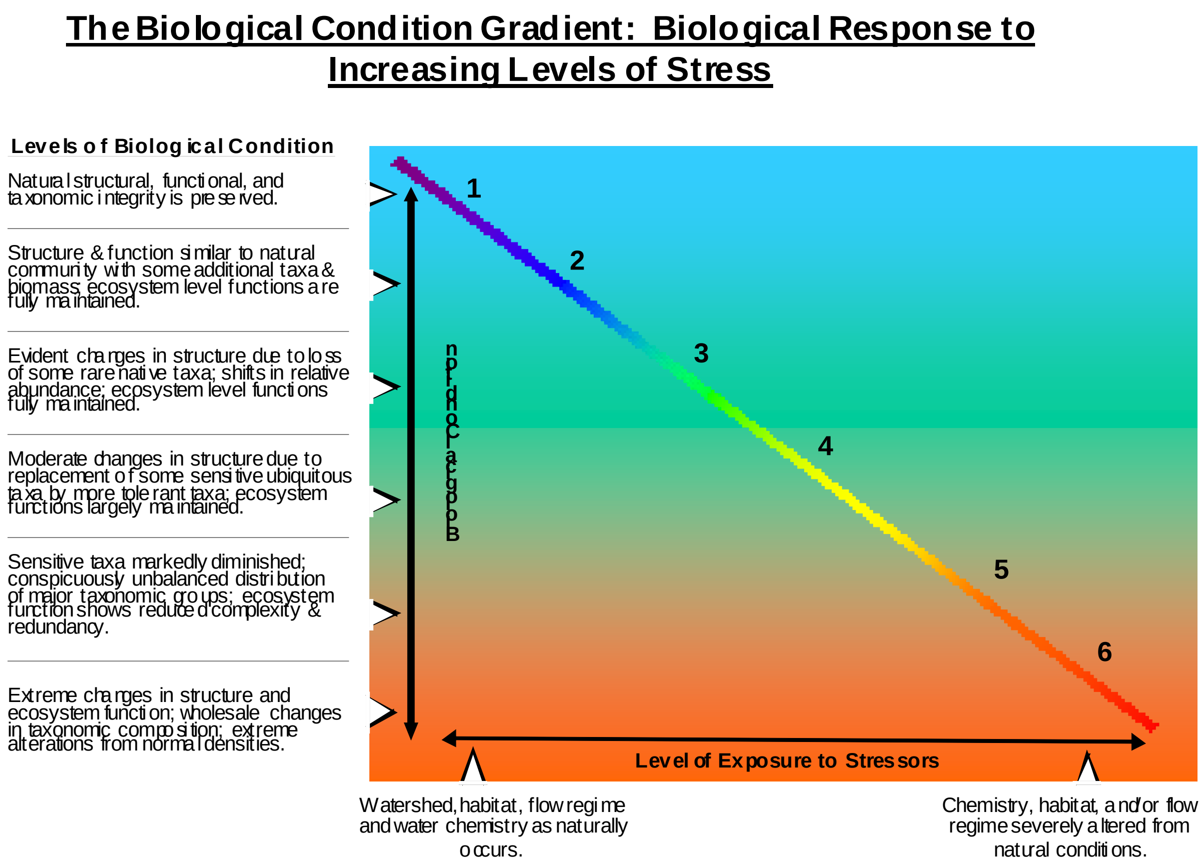


Figure 7. *The Biological Condition Gradient (BCG) conceptual model that depicts six levels of change in key biological attributes in response to the increasing effect of stressors (modified from Davies and Jackson (2006).*

to quantify both the biological and stressor gradients that exist across a region. Herein we suggest that understanding the historical "anchor" or "as naturally occurs" condition is a fundamental component of this primary task. Natural conditions are the conceptual upper end "anchor" of any BCG process even where such conditions no longer exist due to human caused legacy alterations of the landscape and river system. The ecological condition to support an aquatic life use for a waterbody can be described in terms of the BCG levels. For example, the ecological condition needed to support a high quality virtually "pristine" waterbody will be at level 1 or 2. Whereas the support of a sustainable assemblage in a historically altered waterbody may span levels 2-4. The ecological attributes that correspond to the BCG levels are measurable with common biological assessment methodologies and the resultant expression of quality via indices or other tools can be directly linked to an aquatic life use. As such the BCG is used here as an independent method for evaluating the ecological meaning of quantitative thresholds derived by empirical means (Miltner et al. 2011). The BCG therefore provides a rational and consistent means for helping determine appropriate aquatic life uses for the purpose of setting biological impairment thresholds.

Oftentimes a sufficiently broad disturbance gradient exists that helps to define and visualize the biological responses observed along one or more stressor gradients. Reference sites are typically used to empirically derive attainable goals for smaller streams and rivers (i.e., the regional reference condition approach). However, for large and great rivers, the widespread alterations of these waters makes the regional reference condition approach difficult at best due to a dearth of actual reference analogs (Angradi et al. 2009a). Large and great rivers are frequently and directly modified by dams, other hydrological modifications, and chemical impacts (e.g., effluents, runoff). A good description of the historical changes that have occurred in the UMR is available in Pitlo and Rasmussen (2004). Both the impounded and unimpounded reaches represent highly modified conditions with the unimpounded reaches largely cut off from their historical floodplains (Barko et al. 2003, 2004). Minimally disturbed reference sites typically do not exist for large and great rivers in which the integration of the effects of human disturbance in upstream watersheds and identification of attainable goals has relied on the use of "internal" reference sites (Emery et al. 2003; Angradi et al. 2009a). As a supplementary approach, the development of a BCG as a historical anchor can broaden the environmental gradient along which to measure biological performance and better understand the influence of stressor gradients.

The Upper Mississippi River is a historically diverse system with 163 fish species recorded based on a report by Steuck et al. (2010). The list was compiled from a variety of current and historical sources and was used herein as our primary source for the historical occurrence of fish species in the UMR. Expected fish distributions and abundances change with stream size and location so we are following the demarcation used by UMR restoration programs and described in UMRBA 2011 in our analyses, dividing the UMR into the following three reaches:

Upper Impounded Reach: This reach starts upstream on the UMR at St. Croix River and goes downstream to Lock and Dam 13. This includes CWA assessment reaches 1-6 and encompasses river miles 812-523.

Lower Impounded Reach: This reach starts upstream on the UMR at Pool 14 and goes downstream to the Missouri River. This includes CWA assessment reaches 7-11 and encompasses river miles 523-196.

Un-impounded Reach ("Open River"): This reach starts upstream on the UMR at the confluence with the Missouri River and goes downstream to the confluence with the Ohio River. This includes CWA assessment reaches 12-13 and encompasses river miles 196-0.

BCG Methodology

Our analysis, Rankin and Yoder (2011), used distributional information in Steuck et al. (2010) as the "universe" of historical species distributions in the UMR. Steuck et al. (2010) reported, by UMR reach, the presence and relative abundance of each fish species recently or historically collected in the UMR. Presence and abundance codes include:

Table 7. Key to species status codes for the UMR reported by Steuck et al. (2010).

O	Occasionally collected, not generally distributed, but local concentrations may occur.
C	Commonly taken in most sample collections; can make up a large portion of some samples.
A	Abundantly taken in all river surveys.
X	Probably occurs only as a stray from a tributary or inland stocking.
H	Records of occurrence are available, but no collections have been documented in the last ten years.
R	Considered to be rare. Some species in this category may be on the verge of extirpation.
U	Uncommon, does not usually appear in sample collections, populations are small, but the species in this category do not appear to be on the verge of extirpation.

Creating a "Synthetic" Data Set

To extrapolate fish species and abundances during a historical pre-disturbance time period Rankin and Yoder (2011) used existing large river data from throughout the Midwest (most heavily weighted by Ohio and Indiana databases) to estimate; 1) the frequency of occurrence of a species by biological condition range based on the Ohio IBI; and, 2) the relative catch rates using boat electrofishing methods of each species by biological condition range (i.e., numbers per km). This information was combined with the historical fish distribution information (e.g., historical and present occurrence of species in the upper impounded reach, lower impounded reach and unimpounded reach) of Steuck et al. (2010), knowledge of life history information, and descriptions of fish populations from historical snippets (e.g., Carlander 1954) to derive extrapolated potential catch frequencies and abundances that likely occurred during historical periods (e.g., pre-impoundment). These frequencies and average estimates of abundances were then used to create a "pool" of fish to "sample" using a random selection process without replacement. This was done for 10 iterations for each of the three UMR reaches and then the GRFIn, FACI, and the new Ohio Continuous IBI for boatable rivers were calculated for this data.

Inferring Stressor Levels from Species Assemblages

For the UMR, multivariate and correlative measures using GRFIn and GRMIn, the FACI and other measures to identify limiting stressors is one way to understand stressor impacts - this is a "top-down" approach (Miltner et al. 2011). An alternative, but complementary, approach is to use information about individual species responses to stressors gained from broad-scale studies of species sensitivities to infer: 1) which stressors are most limiting; 2) understanding the limiting nature of stressors; and, 3) predict species occurrence and distribution under various stressor reduction scenarios – this is a "bottom-up" approach.

Assumptions

It was assumed that if habitat conditions were close to "as naturally occurs" that sampling would occur along the main channel border and would reflect the availability of historical species from backwater and side channel habitats, many of which [species] are now rare or extirpated. This same assumption has been made by others sampling the main channel and they concluded that such sampling was typically representative of the conditions in the backwaters and secondary channels (Angradi 2006; Thorp 1992). Most of the connections to the floodplain and backwaters have been substantially altered (Weigel et al. 2006).

It may well be that the conditions that occurred during the historical period are not presently attainable because of habitat limitations related to the impounded and disconnected nature of both the impounded and unimpounded UMR. What these extrapolated indices do is provide an anchor point along a continuum of change that "reaches down" to current conditions. It allows the consideration or development of hypotheses about the degree of restoration that might be feasible. This could vary within an impounded pool where the upper portions may be more amenable to restoration compared to lower reaches that are more permanently inundated. In any case the extrapolated data can be used to develop a continuum along one or more disturbance axes.

UMR BCG Development

A complete BCG process depends on input from a panel of regional or system-wide experts on biological assemblages. What is used herein is a modification of an initial BCG exercise conducted on the Wabash River in Indiana (Armitage et al. 2009) modified to fit conditions in the Upper Mississippi River. It is presented as a starting point, not a final BCG product. As with the synthetic community exercise it is designed to help anchor the biological assemblages of the UMR in a historical natural condition and to understand how these assemblages have changed and might change in response to changes along various stressor gradients in the system, thus it is complementary with the stressor analyses done in the other parts of this study.

Determining a Reference Condition for the UMR

Stoddard et al. (2006) summarized the stages of reference conditions that can be used in the management of flowing waters. For all of the major Midwestern rivers it is unlikely that any, or even reaches of these rivers, could be classified as having Minimally Disturbed Conditions (MDC). Existing conditions, depending on the river and setting, would likely be described as in a Least Disturbed Condition (LDC) at best and more typically as Best Available Condition (BAC). For the UMR this has been discussed in the efforts to derive the multimetric fish and invertebrate indices (GRFIn and GRMIn; Angradi et al. 2009b). Describing the Historical Condition (HC) and extrapolating from historical descriptions to an approximate MDC can be used to determine the potential to shift river fish assemblages towards these conditions. We used the BCG exercise, based on historical data from the UMR (Steuck et al. 2010), to guide in establishing and quantifying what historical conditions may have approximated in terms of the fish assemblages that existed. The goal of this exercise is not to set a pristine or natural goal for aquatic life impairment in the UMR, but rather to create a dataset to derive a *trajectory* between existing and historical conditions. By projecting what would be feasible in terms of stressor reduction we can be at least partially predictive in terms of what biological goals are attainable for the UMR.

Extrapolation of Fish Assemblages to Pristine and Pre-Settlement Historical Conditions in the UMR System

One of our goals was to be able to understand the historical fish assemblage condition and biodiversity in the UMR to provide an endpoint or anchor point for extrapolating between existing conditions. This concept is illustrated in Figure 8 - the dark blue points represent the existing conditions in the UMR along a generalized "stressor gradient" along the x-axis (e.g., the stressor gradient of Angradi et al 2009a). This stressor gradient represents the cumulative stressor "load" that influences the biota of the UMR. The green and grey points reflect hypothetical pre-settlement and immediate post-settlement conditions in the UMR. Because of the magnitude of landscape changes and impoundment of much of the UMR that has occurred, these conditions may not be realistic or even desirable societal goals. Hence the expectations for

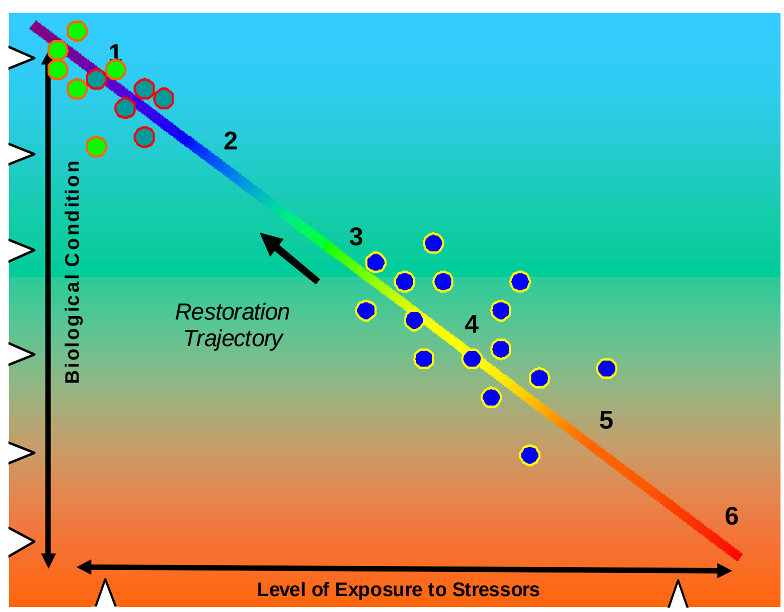


Figure 8. *Hypothetical plot of biological condition (y-axis) vs. a stressor gradient (x-axis; modified from U.S. EPA 2005). On this graph we have superimposed points presenting existing conditions in the UMR mainstem (blue points) and two groups of points representing pre-settlement (green points) and post-settlement conditions (grey points).*

biological assemblages will change in accordance with either. A principal goal of a BCG is to establish 6-7 categories or levels that represent a biological gradient from pristine or MDC along one or more stressor gradients either specific (e.g., chemical, habitat) or along a more "generalized" stressor gradient that combines various measures of anthropogenic stress that typically include landscape measures, population measures (e.g., actual population, housing density, impervious surface, road density, etc.).

Synthetic Assemblage Results

The basis for deriving a "synthetic" historical fish assemblage is the observation that the probability of capture and average abundance of a species is related to the array of stressors present in a reach and is reflected in the biological indices used as response indicators (e.g., GRFIn, IBI, FACI). We used this information to derive probabilities of capture and extrapolated abundances during historical periods prior to much anthropogenic disturbance as well as a time period (e.g., 1960s) with poor-very poor conditions related to inadequate effluent controls. Rankin and Yoder (2011) illustrated changes in the probability of capture, relative abundance, and abundance by capture rate with Ohio IBI for three riverine species in the UMR, the blue sucker, river darter, and black buffalo (see their Figure 4). The extrapolation to historical data used to derive the pool of potential fish for our historical IBI was developed using the trend of actual data, the historical reports of occurrence and distribution from Steuck et al. (2010), and life history information and other historical

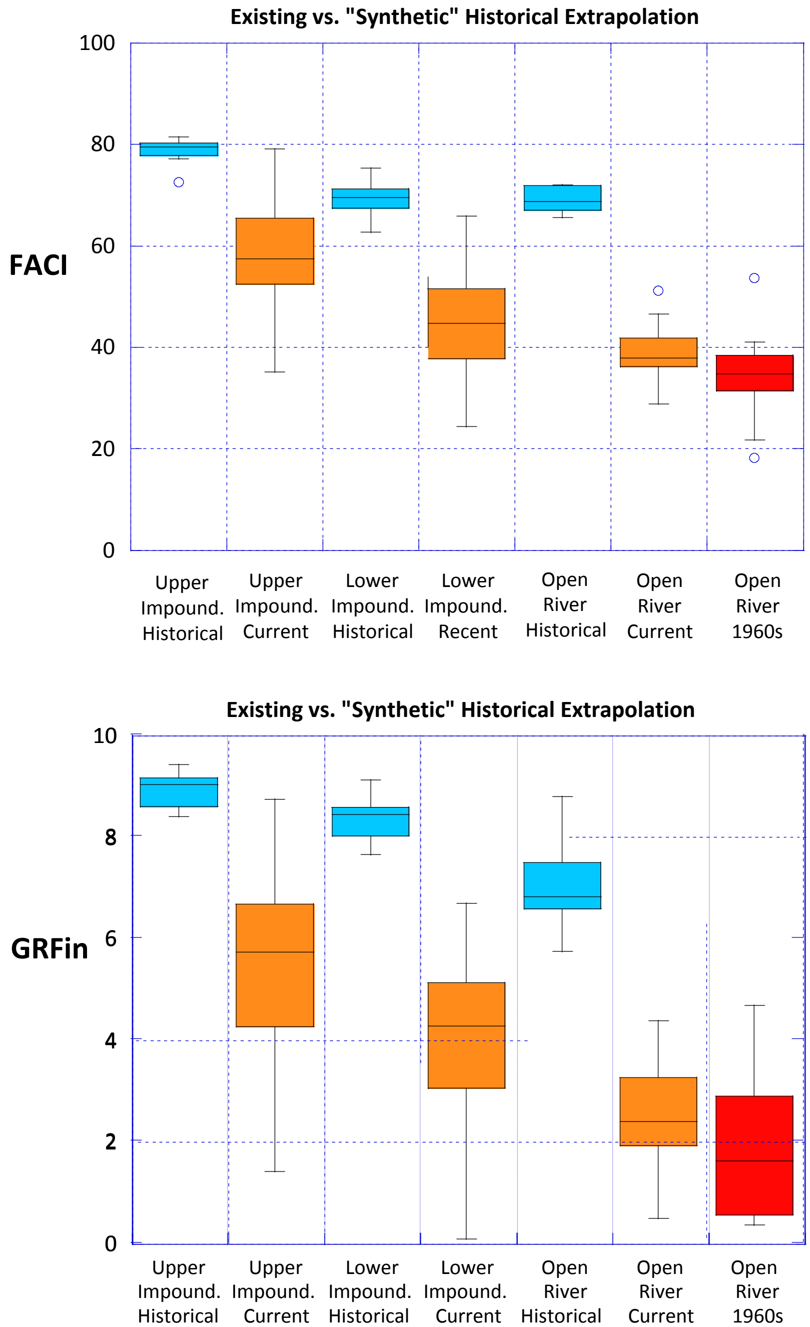
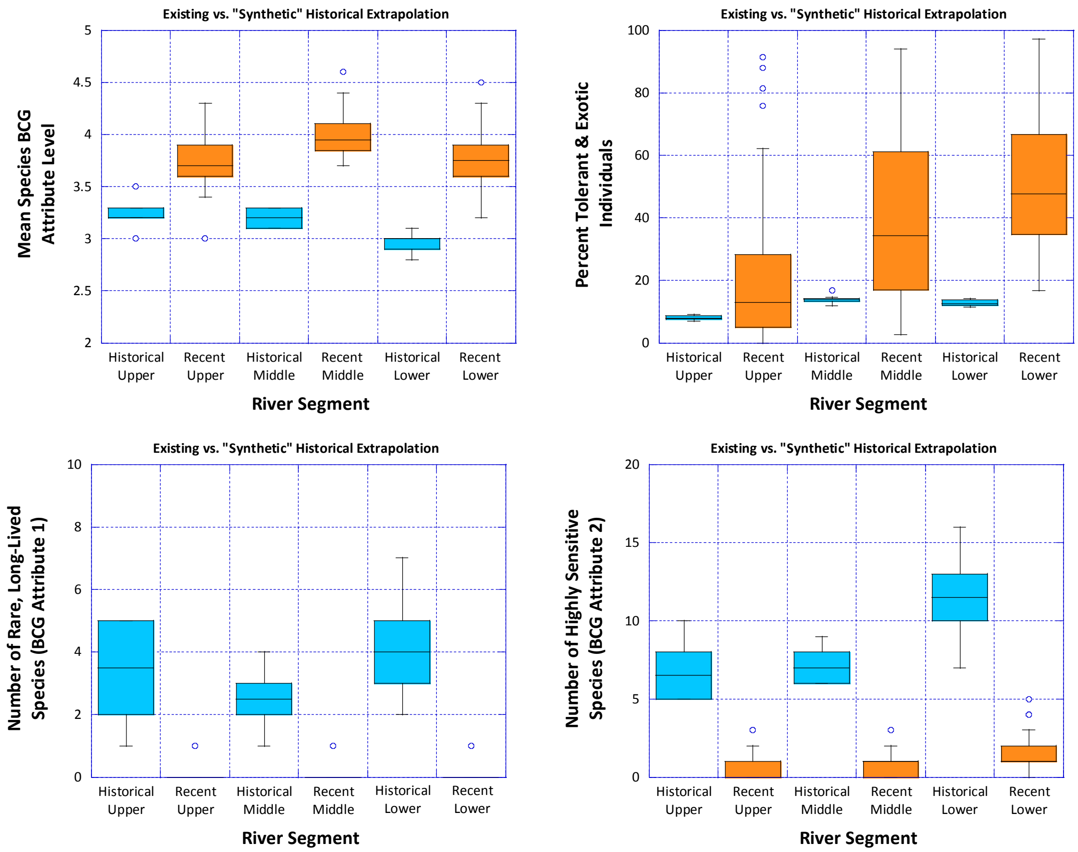


Figure 9. *. Box and whisker plot of FACI scores (top) and GRFin scores (bottom) for historical "synthetically" derived fish assemblages (blue) and recent data (orange) for the upper impounded reaches, lower impounded reaches, and the open river reaches of the UMR. Red box is synthetic data estimating score in the open river during the 1960s prior to CWA point source controls.*

sources that discussed the occurrence of fish species in large Midwest rivers prior to the anthropogenic impacts of the past two centuries. The extrapolation was done separately for each of the three reaches of the UMR (upper impounded, lower impounded, and unimpounded Open River).

For this analysis Rankin and Yoder (2011) calculated the GRFIn, Regional FACI score, and the new "continuous" Ohio boatable IBI (CIBI). The Ohio CIBI was used because it was calibrated to allow scoring above existing LDC reference although it was originally developed for rivers smaller than the UMR. Furthermore, this represents a 30 year database that contains multiple assessments of rivers where changes in quality have been documented in response to CWA abatement measures (Yoder et al. 2005). As such it provides the needed trajectory of change that can be expected for rivers in general. We modified the GRFIn indices by dropping the biomass-based metrics because we did not extrapolate biomass for this exercise. We also recalibrated the scoring expectations by including the historical data (both pre-settlement and 1960s). As expected the synthetic data resulted in higher GRFIn and FACI scores than the current data (Figure 9). Two metrics of the FACI that "under-performed" compared to what might have been expected in a "historical" assemblage included the proportion of deep-bodied suckers and round-bodied suckers. The estimates of abundance were based on abundances of each species at the best existing sites and for rare species extrapolations based on species life histories and anecdotal descriptions of abundances where available. The FACI was calibrated based on existing data in the Midwest and generally on larger rivers. While deep-bodied suckers and round-bodied suckers were abundant in the synthetic results, other species, particularly rare species, were likely higher in abundance in the synthetic assemblage and round-bodied and deep-bodied suckers thus had lower proportional abundances which depressed these metrics. Whether this is a reasonable assumption is not completely certain, but calibration of the FACI from a historical perspective would have captured this difference and these metrics would have scored higher (FACIs in high 80s and 90s). The CIBI has a similar issue with the proportion of round-bodied sucker and simple lithophilic spawners metrics "underperforming" compared to the species composition extrapolated from the model. The scoring of these metrics would have been adjusted or calibrated differently if this data was used in the original calibration of FACI or the CIBI. In any case, the synthetic assemblages scored substantially higher than the present-day data, as might be expected. An understanding of which stressors are limiting to species common in the historical data, but rare or uncommon the recent data would be the basis for understanding whether and where restoration may be feasible.

One thing is readily apparent, that all three reaches have about the same levels of historical BCG attributes. The unimpounded or open river is especially enlightening because it is presently the most highly modified of all three reaches examined here and that is especially reflected by a higher proportion of tolerant and exotic individuals (Figure 10). It also raises questions about whether or not to extend the impounded indices and stressor gradients to the open river. Presently, the open river was treated separately from the impounded UMR in the derivation and calibration of the GRE fish index (GRFIn) and the distinct stressor gradient will result in different thresholds. The BCG analysis suggests that the attribute I-III and attribute VI



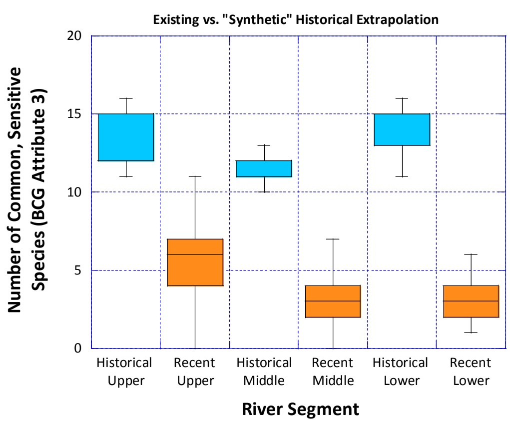


Figure 10. *Box and whisker plots of BCG attribute data for recent data (orange) and historical "synthetic" data for the "upper" impounded (RM 523-812), "middle" impounded (RM 523-196) and "lower" unimpounded (RM 196-0) UMR for BCG attribute measures: mean species BCG attribute (upper left), mean tolerant and exotic species (upper right), number of rare, long-live species (middle, left), highly sensitive species (middle, right) and sensitive common species (bottom, right).*

characteristics of the historical fish assemblage were similar between the impounded and open river, thus raising questions about having two fish indices and two stressor gradients. The OR GRFIn has eight metrics compared to the impounded GRFIn having 10 metrics presumably due the inherently different character of the OR fish assemblage. While each GRFIn lacks metrics that include attribute I-III species (except indirectly via other metrics), the historical presence of these “higher value” BCG attributes suggests that the contemporary OR GRFIn may be lacking in its “coverage” of these attributes. At a minimum this argues for extending the impounded GRFIn into the OR and also for placing value on the FACI which contains some of these connections.

To relate these attributes to the currently used assessment tools the number of sensitive, common species of BCG attribute 3 was correlated to the FACI index and the result was a strong correlation. There was, however, was no apparent correlation between BCG attributes 1 and 2 and the FACI. The strong correlation with the sensitive common species suggests that this attribute could be useful in understanding the consequences of choosing various statistically derived thresholds. The lack of correlation with attributes 1 and 2 may well be explained to the large difference between historical conditions and existing conditions in the UMR. It also could be that populations of these species in the main channel samples, many of which are off main channel habitat dependent, may be more related to the losses of connectivity with the off main channel habitats than with the conditions in the main channel itself. If this is indeed true then it emphasizes the importance in understanding the connections with historical conditions and building this into the assessment process on larger rivers.

Using the BCG to Underpin Selection of Tiered CWA Goals for the UMR

The practical utility of these analyses to the present task of setting appropriate and attainable thresholds for UMR CWA bioassessment is greatly enhanced if threshold options are linked to the BCG, if even only indirectly. Herein we examine using our BCG to underpin the selection of tiered CWA thresholds derived using multimetric indices such as the GRFIn and FACI in the UMR. BCG levels 1 and 2 essentially reflect natural or near-natural conditions. Levels 3 and 4 represent conditions that reflect biological assemblages that have been subjected to increasing levels of stress, but still retain sufficient biological attributes that are consistent with the interim goal of the CWA (e.g., protection and propagation of fish). Level 5 represents a level of stress that is increasingly inconsistent with this goal and is therefore unacceptable and considered as impaired. This suggests the need for remediation such that the overall biological condition improves the assemblage condition to level 4 at least. If such improvement is precluded by activities that cannot be remediated and which “qualify” under the 40 CFR Part 131 existing use and use attainability analysis provisions, a Use Attainability Analysis must be conducted to demonstrate that the assemblage condition cannot be feasibly restored and to document the factors that are most limiting.

Using the initial UMR BCG constructed for this project, the majority of UMR sites currently fall between BCG levels 3 and 5 with few in levels 1, 2 or 6. Without a site specific identification of BCG level for each site we can substitute a weighted mean BCG level for the species found at each site as well as examining the relationship with BCG attributes such as the number of

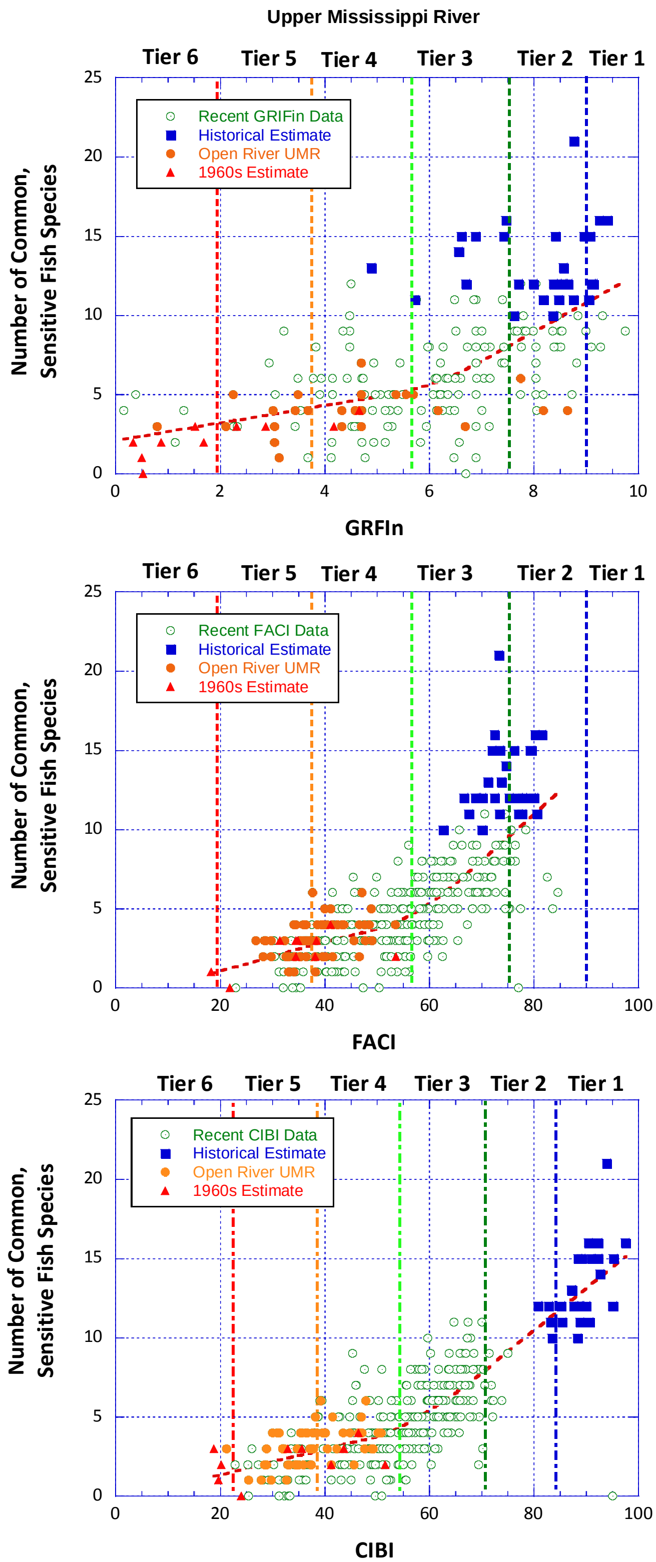


***Figure 11**. Plot of the GRFIn (top), FACI (middle), and CIBI (bottom) vs. the number of BCG attribute 3 species (common, sensitive) for sites in the UMR. Historical calculations were not available for the GRFI, but were extrapolated from correlations with FACI.*

sensitive, common fish species. A lower weighted mean would indicate an assemblage is skewed towards sensitive or intolerant individuals and a higher weighted mean would indicate more tolerant or alien species at a site. Based on this measure and BCG attribute 3, we estimate BCG tier cutoffs for the GRFIn, FACI, and CIBI. Figure 11 reflects plots of the GRFIn, FACI and CIBI vs. the number of BCG attribute 3 species for each UMR site. The "synthetic" data is coded with blue squares to distinguish it from the current sampling data (green circles), we coded the lower, un-impounded river with solid orange points, and the 1960s synthetic data with solid red triangles. The GRFIn, FACI, and CIBI reflect a positive relationship with the attribute 3 BCG measure (Figure 11); the curve is a locally weighted regression that minimizes the effect of outliers. The breaks in these curves illustrate some patterns in these relationships that can be used to support various options for selecting appropriate thresholds. The break in the curves in these relationships with the BCG attribute 3 is aided by the availability of the synthetic data to "complete" the curve (Figure 11, middle).

The CIBI was formulated to better separate high and low performing sites beyond the current range of that index based on currently available data. The red triangles reflect data on the UMR from the time period when point source water quality stressors were severe (1960s) and help to "anchor" the lower part of the distribution. The availability of such data, which we modeled as we did the "pristine" historical conditions, can be used to improve change point analyses and to calibrate multimetric indices to make them more sensitive to extremes of the disturbance gradient (e.g., IBI). Although the CIBI was originally calibrated for smaller rivers the application of the method can illuminate change points since the other indices were calibrated to accommodate estimates of historical assemblage condition.

Integrating Statistical and BCG Approaches

Our initial BCG for the UMR provides a theoretical "endpoint" for consideration of condition thresholds and it therefore can form a basis for better informed discussions of attainability. As such, the BCG can be used to provide a narrative "reality check" to the statistical impairment thresholds derived by Miltner et al. (2011). The "distance" between the present environmental conditions and the "natural" environmental conditions that once existed in the UMR leaves much room for restoration, but this varies by reach and even location within the UMR. We also need to be aware of where the current UMR fish assemblage "is at" with respect to the BCG and as measured by currently used indices such as GRFIn and FACI. The fact that many of the historically common fish species are still present indicates that habitats still exist to support at least relict populations of the BCG attribute 1-3 species.

Evaluating Threshold Options

The 13 most plausible options for the impounded UMR and 8 options for the open river UMR that resulted from of the statistical derivations by Miltner et al. (2011) were ranked from "least stringent" to "most stringent" based on the resulting GRFIn threshold for each. The percentile that the corresponding threshold for GRFIn, GRMIn, and the SMI (where it was applicable) each represents in terms of the UMR database is provided as well. These options are summarized in Tables 8a (impounded UMR) and 8b (unimpounded open river UMR). We added the BCG level

Table 8a. Options for numeric thresholds delineating condition boundaries for the impounded UMR (CWA assessment reaches 0-11). Options are ranked in general order from least to most stringent as CWA threshold with the corresponding BCG level that the biocriterion represents (fish only) based on analyses by Rankin and Yoder (2011), as well as pro and con statements for each. Additional options that were deemed infeasible in Miltner et al. (2011) were not included here.

Option[1]	Threshold Basis[2]	Indices	Biocriteria Score (% rank in UMR)[3]	BCG Level	Rationale (Pro vs. Con)
Impounded UMR (Reaches 0-11)					
1	"Peer Rivers" GRFIn & Ad Hoc at 16th % of UMR Range	GRFIn	38 (16)	4.0	**Pro**: Only option that is "external" to the UMR. Most defensible threshold given the absence of contemporary reference conditions; derived based on performance relative to regional peers; Ad Hoc index in place of GRMIn. **Con**: Assumes condition of REMAP peer rivers as measured by FACI is directly transferable to UMR.
		Ad Hoc	48 (16)		
2	"Peer Rivers" GRFIn & GRMIn at 16th% of UMR Range	GRFIn	38 (16%ile)	4.0	**Pro**: Only option that is "external" to the UMR. Most defensible threshold given the absence of contemporary reference conditions; derived based on performance relative to regional peers; down-weights chance of Type I errors to which the narrow response range of GRMIn is susceptible. **Con**: Assumes condition of REMAP peer rivers as measured by FACI is directly transferable to UMR.
		GRMIn	44 (16%ile)	NA	
3a	GRFIn & GRMIn Lower Bounds of UMR Quadrisection	GRFIn	39 (17%ile)	4.0	**Pro**: Defensible in that high percentages of sites in the UMR surpass these internally-derived thresholds; 1st section boundary of quadrisection reduces propagation of type I errors. **Con**: Susceptible to similar issues with other internally derived thresholds (e.g., no perspective as compared to outside peers).
		GRMIn	47 (26%ile)	NA	

[1] These are listed in a general order of defensibility from most to least.
[2] Combinations of thresholds are described in Table 6.
[3] For the given index score the percentile rank from all scores in the UMR is listed.

Table 8a. (continued)

Option	Threshold Basis	Indices	Biocriteria Score (% rank in UMR)	BCG Level	Rationale (Pro vs. Con)
3b	GRFIn & GRMIn 1st Section Boundary of UMR Quadrisection (empirical y-intercept quadrisection)	GRFIn	46 (27%ile)	4.0	**Pro**: Defensible in that high percentages of sites in the UMR surpass these internally-derived thresholds; upper boundary of quadrisection empirically derived by stressor index; 1st section boundary of quadrisection reduces tendency toward Type I errors. **Con**: Susceptible to similar issues with other internally derived thresholds (e.g., no perspective as compared to outside peers).
		GRMIn	44 (17%ile)	NA	
4	GRFIn & Ad Hoc Quadrisection (1st section boundary)	GRFIn	39 (17)	4.0	**Pro**: Defensible in that high percentages of sites in the UMR surpass these internally-derived thresholds; 1st section boundary of quadrisection reduces tendency toward Type I errors. **Con**: Susceptible to similar issues with other internally derived thresholds (e.g., no perspective as compared to outside peers).
		Ad Hoc	50 (27)	NA	
5	GRFIn & GRMIn Lower Bounds of UMR Quadrisection (empirical y-intercept quadrisection)	GRFIn	46 (27)	3.5	**Pro**: Defensible in that high percentages of sites in the UMR surpass these internally-derived thresholds; lower bound of quadrisection reduces propagation of type I errors. **Con**: Susceptible to similar issues with other internally derived thresholds (e.g., no perspective as compared to outside peers).
		GRMIn	44 (17)	NA	
6	GRFIn & GRMIn Intermediate-Most Disturbed Threshold on GRE Stressor Gradient	GRFIn	53 (38)	3.5	**Pro**: Defensible based on relationship with stressor index; GRMIn threshold relatively high as percent of range in UMR. **Con**: May propagate unacceptable rate type I error with respect to reality of this threshold as a minimum CWA goal; internally based, no perspective to attainment compared to outside peers.
		GRMIn	52 (41)	NA	

Table 8a. (continued)

Option	Threshold Basis	Indices	Biocriteria Score (% rank in UMR)	BCG Level	Rationale (Pro vs. Con)
7	GRFIn Most Disturbed & Ad Hoc Trisection	GRFIn	53 (38)	3.5	**Pro**: Defensible based on relationship with stressor index; GRMIn threshold relatively high as percent of range in UMR. **Con**: May propagate unacceptable rate type I error with respect to reality of this threshold as a minimum CWA goal; internally based, no perspective to attainment compared to outside peers.
		Ad Hoc	53 (37)	NA	
8	GRFIn & Ad Hoc Quadrisection (2nd section boundary)	GRFIn	55 (45)	3.5	**Pro**: Defensible in that high percentages of sites in the UMR surpass these internally-derived thresholds. **Con**: May propagate unacceptable rate type I error with respect to reality of this threshold as a minimum CWA goal; internally based, no perspective to attainment compared to outside peers.
		Ad Hoc	60 (48)	NA	
9	GRFIn & GRMIn Mid-point of quadrisection UMR	GRFIn	56 (53)	3.5	**Pro**: Defensible in that high percentages of sites in the UMR surpass these internally-derived thresholds . **Con**: May propagate unacceptable rate type I error with respect to reality of this threshold as a minimum CWA goal; internally based, no perspective to attainment compared to outside peers.
		GRMIn	55 (45)	NA	
10	GRFIn + SMI Reaches 0-5 Mid-point of quadrisection UMR; GRFIn & GRMIn Reaches 6 - 13	GRFIn	56 (53)	3.5	**Pro**: SMI is sensitive to multiple stressors in the UMR; higher thresholds suggested to coincide with best performing reaches. **Con**: SMI is presently applicable only to Reaches 0 – 5; may propagate unacceptable rate of type I error with respect to reality of this threshold as a minimum CWA goal; internally based, no perspective to attainment compared to outside peers.
		SMI	44 (41)	NA	

Table 8a. (continued)

Option	Threshold Basis	Indices	Biocriteria (% rank in UMR)	BCG Level	Rationale (Pro vs. Con)
11	GRFIn @ median of Lower St. Croix; GRMIn at corresponding UMR percentile	GRFIn	62 (56)	3.0	**Pro**: Externally derived as compared to adjacent river that is perceived as being less impacted than most of UMR; may function best as an upper tier threshold as opposed to minimum CWA goal. **Con**: May propagate unacceptable rate type I error with respect to reality of this threshold as a minimum CWA goal for UMR.
		GRMIn	57 (56)	NA	
12	GRE Indices trisection of 95-5%ile of "Best UMR" Reaches 2&3	GRFIn	70 (69)	3.0	**Pro**: Sets thresholds at good-fair-poor; provides for upper tier threshold. **Con**: May incur type I errors, especially for GRMIn, as a baseline CWA threshold.
		GRMIn	59 (70)	NA	
13	GRE Indices quadrisection of 95-5%ile of "Best UMR" Reaches 2&3	GRFIn	67 (65)	3.0	**Pro**: Sets thresholds at excellent-good-fair-poor; provides for upper tier threshold. **Con**: May incur type I errors, especially for GRMIn, as a baseline CWA threshold.
		GRMIn	58 (59)	NA	

Table 8b. Options for numeric thresholds delineating condition boundaries for the unimpounded Open River reaches of the UMR generally ranked from least to most stringent in terms of CWA threshold with enhanced pro and con statements for each and the corresponding BCG level that the biocriterion represents (fish only) based on analyses by Rankin and Yoder (2011). Options are ranked in order from least to most stringent as CWA threshold with exception of option 8.

Option[4]	**Threshold Basis**[5]	**Indices**	**Biocriteria Score (% rank in UMR)**[6]	**BCG Level**	**Rationale (Pro vs. Con)**
1	"Peer Rivers" Missouri River GRFIn and GRMIn 16th percentile of UMR	MO River GRFIn	38 (16)	4.0	**Pro**: Only option that is "external" to the UMR. Assumes conditions in lower Missouri River are analogous to that in OR reach of UMR. Defensible given the absence of contemporary reference conditions, derived based on performance relative to peers; down-weights chance of Type I errors. **Con**: Uncertainties are associated with degree that OR is similar to lower Missouri River.
		MO River GRMIn	39 (16)	NA	
2	"Peer Rivers" GRFIn at 16th% of UMR Range	GRFIn (Imp. UMR)	38 (16)	4.0	**Pro**: Only option that is "external" to the UMR. Most defensible threshold given the absence of contemporary reference conditions in OR and highly degraded state based on BCG; derived based on performance relative to regional peers; opens needed consideration of UAA issues. **Con**: Assumes REMAP reference condition is directly transferable to UMR OR reaches; extends Impounded GRFIn to OR; not locally derived and applicability of REMAP rivers to OR is a potential issue; may necessitate UAA considerations; single assemblage application.

[4] These are listed in a general order of defensibility from most to least.
[5] Combinations of thresholds are described in Table 5.
[6] For the given index score the percentile rank from all scores in the UMR is listed.

Table 8b. (continued)

Option	Threshold Basis	Indices	Biocriteria Score (% rank in UMR)	BCG Level	Rationale (Pro vs. Con)
3	Missouri River GRFIn and GRMIn Reaches 12 – 13 Lower Bounds of UMR Quadrisection	MO River GRFIn	39 (17)	4.0	**Pro**: Assumes conditions in lower Missouri River are analogous to that in OR reach of UMR. Lower boundary of quadrisection down-weights chance of Type I errors. **Con**: Uncertainties are associated with degree that OR is similar to lower Missouri River.
		MO River GRMIn	47 (34)	NA	
4	Missouri River GRFIn and GRMIn Lower Bounds of Trisection UMR	MO River GRFIn	44 (26)	4.0+	**Pro**: Assumes conditions in lower Missouri River are analogous to that in OR reach of UMR. Lower boundary of trisection down-weights chance of Type I errors. **Con**: Uncertainties are associated with degree that OR is similar to lower Missouri River.
		MO River GRMIn	50 (40)	NA	
5	GRFIn Intermediate-Most Disturbed Threshold on GRE Stressor Gradient	GRFIn (Imp. UMR)	53 (38%ile)	3.5	**Pro**: Defensible based on relationship with impounded UMR stressor index; GRFIn threshold relatively high as percent of range in UMR. **Con**: Applies impounded GRFIn outside of its calibration domain; result is propagation of type I error; single assemblage application.

Table 8b. (continued)

Option	Threshold Basis	Indices	Biocriteria (% rank in UMR)	BCG Level	Rationale (Pro vs. Con)
6	Missouri River GRFIn and GRMIn Impounded UMR Most Disturbed	GRFIn	52 (38)	3.5	**Pro**: Assumes conditions in lower Missouri River are analogous to that in OR reach of UMR. Defensible based on relationship with impounded UMR stressor index; GRFIn and GRMIn thresholds relatively high as percent of range in UMR. **Con**: Applies impounded stressor gradient outside of its calibration domain; result is propagation of type I error; uncertainties are associated with degree that OR is similar to lower Missouri River.
		GRMIn	53 (49)	NA	
7	Missouri River GRFIn and GRMIn Middle Bounds of UMR Quadrisection	GRFIn	55 (48)	3.5	**Pro**: Assumes conditions in lower Missouri River are analogous to that in OR reach of UMR. Defensible based on relationship with impounded UMR stressor index; GRFIn & GRMIn thresholds relatively high as percent of range in UMR. **Con**: Applies impounded stressor gradient outside of its calibration domain; result is propagation of type I error; uncertainties are associated with degree that OR is similar to lower Missouri River.
		GRMIn	56 (59)	NA	
8	OR GRFIn Intermediate-Most Disturbed Threshold on GRE Stressor Gradient	GRFIn (OR UMR)	36 (14%ile)	4.5	**Pro**: Defensible based on relationship with OR UMR stressor index. **Con**: Issues identified by BCG analysis infer that using OR GRFIn and associated stressor index sets threshold too low for CWA interim goal; could obscure attainability uncertainties.

from the analyses in Rankin and Yoder (2011) that corresponds to the numeric GRFIn threshold for each statistically derived threshold option. The pros and cons of each option are also summarized hence the table forms a basis for making informed decisions about which threshold options are the most plausible as CWA goal thresholds for the UMR main channel. Besides considering the basis for each, the statistical derivations alone seem insufficient for making choices about which sets of thresholds should be considered as the minimum CWA threshold for the UMR. At a minimum, even the best informed choice would be somewhat subjective especially considering that there is no minimally disturbed or least disturbed population of reference sites upon which to extract a biocriterion from the UMR. REMAP offers a somewhat analogous alternate, but that too has associated uncertainties. As such any choice on a purely statistical basis is inherently subjective.

States that have established biocriteria thresholds as numeric WQS or as translators of a narrative biocriterion have used percentiles of a population of reference sites. This has been done mostly for wadeable streams and smaller rivers although a few have developed large river biocriteria (e.g., Ohio, Maine, Wisconsin, ORSANCO) following an approach that attempted to emulate the Stoddard et al. (2006) descriptions. However, a few states have also developed biological assessment thresholds using a BCG approach and usually in concert with the more traditional multimetric indices. EPA has recently expressed concern about where states are setting impairment thresholds for the minimum CWA goal use fearing that they are being set too low due to a lack of consideration about their reference population or opting to err on the side of making a type I error given the implication for TMDLs in impaired segments. While no guidance or policies have been forthcoming from EPA, there seems to be a consensus among the states that BCG level 4 is consistent with the minimum acceptable CWA goals for aquatic life. As such, the states that have developed BCG models have an alternate method to ground truth their more statistically driven derivation of index based thresholds.

UMR Threshold Option BCG Levels

With the development of a preliminary BCG for the UMR by Yoder and Rankin (2011) and the synthetic fish model that it accompanies, an alternate and independent means exists for evaluating the statistically derived options of Miltner et al. (2011). In terms of the BCG the options range from level 3.0 to 4.0 in the impounded UMR (Table 8a) and 3.5 to 4.5 for the unimpounded open river UMR (Table 8b). For the impounded UMR options 1-4 are consistent with BCG level 4 for the fish assemblage and the associated biocriterion for GRFIn ranges from the 16^{th} to 19^{th}%ile of the UMR. The range of GRFIn scores is 38 for the only externally derived option to 41 for two of the “best of UMR” reach 2 and 3 options. GRMIn scores were somewhat more variable ranging from 41 (19^{th} %ile) for options 3 and 4 to 47 (26^{th} %ile) for option 2. With the exception of the slightly higher GRMIn percentile for option 2, these are similar in their representation as is the GRFIn. The remaining options coincide with BCG levels 3.5 and 3.0 and as such have higher GRFIn and GRMIn criteria and higher percentiles. Setting the minimum CWA threshold at these levels will result in a higher proportion of UMR sites as impaired and if the BCG is taken into consideration it may incur a greater type I error. However, there are a significant number of sites in the UMR that already meet these higher thresholds. Hence this seems like a good opportunity to develop “upper tier” thresholds so that those sites are protected to their current condition. It would also allow sites and segments that respond to restoration in a better than expected manner could gain the protection of a higher tier threshold. This would provide more utility and meaning for the UMR bioassessment and would hopefully lead to the development of TALUs. Option 10 or a variation thereof seems like a good candidate for an upper tier threshold.

The unimpounded open river UMR presents an added challenge in threshold setting. While the open river GRFIn was specifically derived and calibrated for the OR reaches, there has been some concern expressed about this methodology. The concern is that the OR is highly modified for navigation and flood control hence it is already impacted and potentially unsuitable for developing biological expectations consistent with the CWA minimum goal for aquatic life. The BCG analysis indicates that the historical OR had the same species richness and attributes I and II potential as the impounded UMR. Figure 11 from the BCG analysis also suggests that the OR GRFIn does not match up very well with the BCG with scores ranging through all six levels. Contrast this to the more widely calibrated FACI and Ohio CIBI that cluster fairly tightly in BCG levels 4 and 5 with a few sites in level 6. This strongly suggests that the OR GRFIn is not reliable to recognize the different BCG levels due to its comparatively narrow calibration domain. The OR GRFIn biocriteria that resulted from a trisection of the OR stressor gradient coincided with BCG level 4.5 which suggested that it would not be sufficiently protective.

In responding to the issues identified with the open river GRFIn and GRMIn, the Missouri River indices developed by EMAP-GRE (also called GRFIn and GRMIn) were incorporated into the analyses of threshold options for the open river UMR. This approach provided some promising options, as Option 2 in Table 8b yielded the same results as option 1 in Table 8a and seems to be a reasonable threshold given what we know about this reach from the BCG analysis. The only remaining option sets the threshold at the 38^{th} %ile which seems too high for a minimum CWA biocriterion and it coincides with BCG level 3.5. One issue that may arise in the future is with impaired sites and reaches not being addressable due to the legacy modifications of habitat for navigation and flood control. If this indeed present a restoration conflict then it suggests the need for a use attainability analysis (UAA; 40CFR Part 131). However, each instance will need to be evaluated for the feasibility of restoration and in consideration of the most limiting factors. This could also apply to portions of the impounded UMR where legacy or socioeconomic dependent modifications have been made.

In summary we recommend the following for detailed consideration as UMR thresholds:

1. **We recommend a unified threshold option (i.e., same BCG level expectation) for both the impounded and unimpounded UMR based primarily on the results of the BCG analysis.** Historically speaking the impounded and unimpounded open river reaches had remarkable similarity in BCG attributes and characteristics.

2. **For the Impounded UMR any of the option 1-4 thresholds would be feasible as a minimum CWA biocriteria based on two assemblages.** At this time we are recommending GRFIn and the Ad Hoc macroinvertebrate index with the caveat that the

SMI could be used as a third assemblage. This also assumes that the Ad Hoc index will be more fully developed in the near future.

3. **Of these, options 1 or 2 (“Peer Rivers” [GRE indices at 16%ile]) have some compelling reasons for selection, one of which is its external basis for derivation.** This is the closest option that is available that emulates a reference condition approach.

4. **Option 3 (GRFIn and GRMIn lower bounds of UMR quadrisection) could also serve as a basis for the minimum CWA goal and it is derived from the GRE stressor gradient from which the GRFIn and GRMIn were derived.** Its weakness is being internally derived and the issues that raises.

5. **Options 5-10 each result in higher minimum CWA thresholds and thus risk the propagation of type I errors which could result in erroneous impairment listings. These options correspond to BCG level 3.5.**

6. **Options 11-13 correspond to BCG level 3.0 and provide an opportunity to establish an upper tier threshold that would be used in addition to the minimum threshold. Option 11 or a variant thereof seems to offer the best rationale for such a threshold.** The result will be the recognition and protection of sites, segments, and reaches that are performing better than the CWA minimum. As such this would constitute a “protection use” in addition to the CWA minimum functioning as a “restoration use”. A methodology for determining how and where to apply this upper tier threshold will need to be developed.

7. **For the Open River options 1-3 all offer about the same result in terms of an impairment threshold and BCG level (4.0). Option 1 is more compelling if the notion that the OR is more like the lower Missouri River.**
   While the results for the fish assemblage among options 1-3 achieve about the same result, options 1 and 3 offers a dual assemblage approach. Option 2 is based on extending the impounded UMR GRFIn into the OR based on the BCG conclusions, but it offers a single assemblage approach.

Importantly, the leading threshold options for both the unimpounded (Table 8a, options 1-4) and the open river (Table 8b, options 1-3) produce comparatively similar results, in terms of biocriteria thresholds. As such, structurally different methods for deriving thresholds appeared to indicate close agreement regarding an appropriate minimum CWA threshold for the UMR. While further refinement and revision to the supporting indices, the BCG, and threshold determination may be needed, what emerges from our work herein is a promising conclusion that threshold identification is feasible, robust, and meaningful in a CWA context for the UMR.

Chapter 6: Implications of Adopting Biological Assessment for the UMR

Background and Scoping Report Prediction of Implications
Since the current CWA aquatic life use status assessment relies on chemical/physical surrogates, and based on what we know about how biological approaches compare as described in Chapter 3 (Karr and Yoder 2004), some implications of applying a standardized and rigorous biological assessment approach to the UMR main channel as predicted in this project's Background and Scoping Report include, but are not limited to, the following:

1. Biological endpoints will serve as the arbiter of status, which means that prior determinations based on chemical/physical surrogates will almost certainly change – the expected result is that the quantity of impairment (i.e., lineal miles of mainstem, severity of impairments) will increase based on what we have learned elsewhere;
2. The more comprehensive assessment provided by a biologically based approach will add new sites, segments, and reaches to state 303[d] listings, however, we cannot rule out the delisting of currently listed segments;
3. The type of integrated biological assessment that is envisioned by this guidance will lead to a sharper definition of causes and sources of impairment and the capacity to "proportionalize" each based on the severity and extent of measured biological impairments;
4. The more refined spatial definition of impairments, as specified by the adequate monitoring and assessment concepts and the increments of biological condition communicated by multimetric indices, will likely lead to opportunities to refine longitudinal reporting reaches for the main channel, and this may include lateral distinctions in the future;
5. The potential to apply the concepts embedded in a "tiered aquatic life use (TALU) approach" to monitoring and assessment and water quality standards (WQS) will lead to further refinements in the determination of status and also in addressing legacy impacts that will be dealt with by the "use attainability" aspects of this approach.

The Background and Scoping Report (Yoder et al. 2010) suggested that, once a preferred biological assessment methodology is identified, a pilot should be conducted to determine how 305[b] assessments and 303[d] listings will change from their current delineations. This led to the development of the thresholds report (Miltner et al. 2011), as well as the biological condition gradient report (Rankin and Yoder 2011). These reports, coupled with the discussion and conclusions in chapter 5, can serve as the basis for an initial biologically based CWA assessment of the UMR main channel.

Thresholds Report Findings
Relative to the predictions of the Background and Scoping Report, the results of the thresholds report indicate the following about each of the five above predictions.

1. *Biological endpoints will serve as the arbiter of status, which means that prior determinations based on chemical/physical surrogates (see Appendix A) will very likely*

change – the expected result is that the quantity of impairment (i.e., lineal miles of mainstem, severity of impairments) and associated causes will increase based on what we have already learned elsewhere.

All of the UMR-derived threshold analyses produced significantly greater non-attainment for the UMR main channel as a whole (less in some reaches, more in others) than the current non-biological 303[d] list. Even the most forgiving externally-derived approaches, which indicated a percentage of impairment, river-wide, similar to current listings, showed at least some level of impairment in the majority of assessment reaches. As such, the application of a biological assessment to the UMR main channel will comprise a significant change with regard to aquatic life use support, as only 4 of 13 interstate assessment reaches are currently reported to have an aquatic life use impairment under the current non-biological approach.

2. *The more comprehensive assessment provided by a biologically based approach will add new segments to state 303[d] listings, however, we cannot rule out the delisting of currently listed segments.*

   The findings of the thresholds report indicate that, under various threshold calculation scenarios, most or all of the UMR assessment reaches have varying degrees of biological impairment, thus new segments would certainly be added under any one of the available biological assessment options in chapter 5. All of the currently listed reaches show at least some impairment under all of the threshold calculation options. As such, delisting of currently listed segments seems unlikely given the extent of the biological impairments under the various threshold options.

3. *The type of integrated biological assessment that is envisioned by this guidance will lead to a sharper definition of causes and sources of impairment and the capacity to "proportionalize" each based on the severity and extent of measured biological impairments.*

   The analyses utilized in the thresholds report also led to the determination of proximate stressors. Furthermore, the dual assemblage approach broadened the capacity of the biological assessment approach to capture a wide array of stressors. The stressors that corresponded to biological impairment include both "pollution" (e.g., habitat and flow related stressors) and "pollutants" (e.g., ammonia), with the added advantage of being able to understand where different stressors were potentially interacting to affect the biota. Such knowledge is advantageous when considering abatement and restoration practices to address biological impairments and can extend the value of the biological assessment beyond traditional CWA programs.

4. *The more refined spatial definition of impairments, as specified by the adequate monitoring and assessment concepts and the increments of biological condition communicated by multimetric indices, will likely lead to opportunities to refine*

longitudinal reporting reaches for the main channel, and this may include lateral distinctions in the future.

While the thresholds report was necessarily focused at the UMR assessment reach level, the design of the GRE spatial design allows for more detailed examination of impairments and their extent and severity along a longitudinal continuum. While this would require additional effort and analysis, it would enhance the WQTF states capacity to apply the results beyond their initial 305[b] and 303[d] obligations. This type of application is needed to better understand the resource and to more appropriately consider the application of incremental assessment as an alternative to the more tradition “pass/fail” framework.

The analyses conducted for this project were primarily focused on the identification of a minimum CWA goal threshold. In addition, the opportunity exists to begin the development and refinement of a TALU framework which would include upper tier thresholds to protect already “better performing” sites, segments, and reaches seem feasible. In turn, sites, segments, and reaches that perform below the minimum CWA goal threshold can now be evaluated for attainability given the widespread and legacy modifications of the main channel for navigation and flood control. This seems especially applicable to the unimpounded open river reaches, but could also be an issue locally in the impounded reaches.

This project was also focused on the UMR main channel. As discussed in the WQTF’s draft report on aquatic life designated uses (UMRBA 2011), a CWA monitoring and assessment approach for the off channel habitats will also be needed. The concepts and analyses contained in this guidance – though not necessarily the specific sampling protocols and indices – are generally applicable to the off channel habitats as well. As such, a similar process could be undertaken in the future to explore approaches for the UMR’s off channel habitats.

5. *The potential to apply the concepts embedded in the “TALU approach” to monitoring and assessment and WQS will lead to further refinements in the determination of status and also in addressing legacy impacts that will be dealt with by the “use attainability” aspects of this approach.*

While this is also linked to the framework discussed in number 4 above, the integration of the thresholds analyses with the BCG analyses represents the critical first steps of a TALU based approach which consists of . . . “tiered aquatic life uses based on numeric biological criteria and implementation via an adequate monitoring and assessment program that includes biological, chemical, and physical measures, parameters, indicators and a process for stressor identification”. Thus far the candidate biological criteria are represented by the GRE biological indices, implementation by adequate monitoring and assessment is largely exemplified by the GRE sampling design, and the stressor identification process employed in the thresholds report is exemplified by the proximate stressor analysis. While all three are essential ingredients of a TALU based approach, the key missing piece is the tiered aquatic life use framework. Threshold options 8-10 for the impounded river (see Table 8a) present viable scenarios for an upper tier impairment threshold, as each produces an attainable goal in the context of comparing the UMR to its nearby “peer” rivers. Application of two thresholds, one for the minimum CWA goal and the other for better performing sites, segment, and reaches seems appropriate and could be the forerunner for establishing tiered aquatic life uses (TALU). While this project’s scope has been primarily focused on the minimum CWA threshold goal, future work could continue to address the establishment of UMR aquatic life use tiers.

Overall, the anticipated implications of adopting CWA biological assessment for the UMR were borne out and made more specific by the work done in the development of this guidance document. This work has shown that an initial UMR CWA biological assessment utilizing a minimum CWA threshold is feasible with existing tools, that the results of biological assessment will significantly affect the CWA impairment characterization of the river, that stressors can be identified and assessments conducted at multiple spatial scales, and that a TALU framework can be developed by building on this project. As such, the implications of biological assessment become a more concrete and contemporary consideration for the UMR states.

Chapter 7: Implementation Considerations and Next Steps

Principles for Future UMR CWA Biological Assessment

We recommend that any future CWA based monitoring and assessment program adhere to the following principles:

1. Follow the adequate monitoring and assessment approach outlined by Yoder (1998). Essentially this outlines an approach for identifying the parameters and indicators that are the essential ingredients of a TALU based approach[1];
2. Develop a system-wide strategy that fosters the coordination and standardization of methods and implementation should multiple entities become the primary data collectors for sustained UMR CWA biological assessment;
3. Utilize a sampling design that provides spatially sufficient and robust coverage so as to detect the myriad of pollutional, stressor, and habitat gradients that occur along the UMR – this means that an intensive, longitudinal "pollution survey" type of design is preferred;
4. Sampling and analysis of the resulting data is executed by qualified professionals and within a programmatic framework that exhibits the characteristics of a level 4 program as defined by the critical technical elements process;
5. Data is transformed an analyzed not only to produce reach scale status assessments, but which can also support investigations and analyses at the site, segment, and reach scales. This allows for the entire effort to fulfill the goal of providing "day-to-day" support to both CWA and non-CWA programs.

Developing a sustained program based on these principles brings the focus on providing a measurement framework that can assess current conditions, but also detects changes in increments of condition and serves as a feedback to the various management programs that are working to restore and maintain the biological quality of the UMR.

While the development of thresholds is a critical component of this framework, it is a result of the quality and characteristics of the overall monitoring and assessment program that will eventually be applied to the UMR. Developing a comprehensive strategy that actually leads to the execution of this type of monitoring and assessment program is an essential next step.

EMAP-GRE as the Preferred Existing Model

This project's primary task was to evaluate existing tools for their suitability in supporting a biological assessment of the UMR under the Clean Water Act. The options for an initial CWA biological assessment of the UMR main channel, as examined in this report, depended on the monitoring and assessment conducted by the U.S. EPA EMAP-GRE program, as it provided the only longitudinally complete data set for the UMR. Additionally, the EMAP-GRE developed

[1] The TALU based approach includes tiered aquatic life uses (TALU) based on numeric biological criteria and implementation via an adequate monitoring and assessment program that includes biological, chemical, and physical measures, parameters, indicators and a process for stressor identification.

indices, GRMIn and GRFIn, along with the SMI, are the most promising assessment tools currently available for the UMR. Moreover, the EMAP-GRE protocols and assessment tools, taken in sum, closely reflect the bioassessment principles outlined above.

Therefore, while other bioassessment programs exist on the UMR, the EMAP-GRE approach (with appropriate modifications) is preferred for conducting a CWA assessment of aquatic life use support not only in near term, but in the long term, as is discussed in Chapters 4 and 5. As such, it fulfills the desire of the WQTF for a biologically based assessment for CWA purposes using existing tools.

However, as it stands, EMAP-GRE is a one-time effort with no plans for a repeat survey in the future. This means that to sustain anything beyond the current one-time assessment a strategy for conducting future assessments of the UMR is needed.

Possible Options for a Sustained Bioassessment of the UMR

Presently none of the programs that currently collect biological data on any portion of the UMR provide a *seamless substitute* for the GRE design, as described below. The states, therefore, will need to consider if and how to work from existing programs in implementing UMR biological assessment if a new "GRE like" program is not forthcoming.

Existing Programs

<u>USACE EMP-Long Term Resource Monitoring Program:</u> The USACE EMP-LTRMP presently collects data on selected pools and under a design that is fundamentally different from the preferred CWA bioassessment approach. As such it is not a seamless substitute. Differences in addition to a fundamentally different spatial design also include incomplete indicator assemblages (LTRMP does not include macroinvertebrates). While the technical and professional capacity of LTRMP is entirely sufficient (see Appendix C), continuing the data collection needed for the preferred CWA bioassessment design would require supplemental sampling and data analysis efforts. The LTRMP field stations have already demonstrated the capacity to execute a spatially sufficient coverage of the UMR (the field stations collected most of the GRE data in 2004-6). However, sampling under the GRE design and by the protocols of GRE is not currently within the LTRMP mission. In considering the potential for LTRMP to support UMR CWA assessment, it is important to recall that the fundamental assessment unit of the preferred sampling design is the 1.0 km site and that this forms the basis of the resulting data from which all else is derived. The Wisconsin DNR LTRMP field station has already anticipated this issue by conducting detailed comparability studies for the various fish sampling programs on the Wisconsin portion of the UMR (T. Dukerschein, personal communication). Part of these developmental efforts includes investigating if and how extant LTRMP fish sampling data could be aggregated to emulate the GRE protocol. However, based upon our detailed examination of the various programs in the Background and Scoping Report, a partial pool or reach-wide assessment is the only option for utilizing the current LTRMP sampling and resulting data and it would be limited to fish and submerged vegetation data. Thus, any approach relying on the current LTMRP design and assemblages, even if extended beyond the

current LTRMP study reaches, would have significant limitations in meeting the desired characteristics of a CWA assessment.

State Bioassessment Programs: These presently range from virtually non-existent to in-development to being practiced on a routine basis. The Wisconsin DNR Large Rivers Program includes the UMR portion in Wisconsin in addition to all of their inland rivers. However, the methodology and its execution are sufficiently different that any use of this data would not include a direct translation to the GRE suite of indices, although an approximate "pass/fail" bioassessment might be had with this program. Other states such as Minnesota have active and growing CWA bioassessment programs, but these are a likely a few years away from being viable for sustained CWA bioassessment support. Any state program will inherently be limited to its portion of the UMR and we have seen in the analyses that an interstate strategy is essential to extract the full benefits of an integrated biological assessment. Therefore, an approach utilizing state bioassessment programs would require both significant developments within these programs while ensuring standardization between programs. While it may be preferential that a single entity manage and conduct the sampling, reality may dictate the need to have multiple entities collaborating on a unified strategy, such as would need to be done if multiple state programs have a role in UMR bioassessment. In this event, standardization of sampling and its execution will be critical and will require some level of river-wide agreement and oversight.

A New Program

Ideally it would be preferential for a single entity to execute and manages all aspects of a future UMR bioassessment program. A single entity approach would also be accompanied by a (yet-to-be-developed) comprehensive UMR CWA monitoring strategy, though such a strategy is also needed to better inform the LTRMP and state-based options described above. The rationale for this approach includes addressing and minimizing concerns that naturally accompany a multiple entity approach. Not only does this include the obvious data collection concerns, but also extends to study design, data management, and data analysis concerns. While there are at present no real prospects for this type of approach, we recommend that this be considered and detailed as an option as part of any future monitoring strategy development. As mentioned above, a monitoring strategy blueprint for every aspect of a unified and standardized bioassessment of the UMR main channel would also be necessary in the event multiple entities carry out the work.

Assemblage/Index Recommendations

Regardless of which entity or entities execute the sampling consistency in the application of biological indicator assemblages and indices is critical to a valid assessment of the UMR main channel. The following are our recommendations for the preferred assemblages and indices, among existing tools, for biological assessment of the UMR. We have chosen to divide the recommendations into the UMR floodplain reaches in keeping with the draft UMRBA WQTF report regarding aquatic life uses and the methods of this project's preliminary BCG analysis; Upper Impounded, Lower Impounded, and Open River reaches.

- ✓ **Upper Impounded Reach**: Impounded GRFIn, Ad Hoc macroinvertebrate index, SMI
- ✓ **Lower Impounded Reach**: Impounded GRFIn, Ad Hoc macroinvertebrate index
- ✓ **Open River**: Missouri River GRFIn and GRMIn

These recommendations take their cue from the performance of each index determined from the thresholds report, but the essential finding is that these are the aquatic assemblages that need to be part of a long term monitoring strategy. Improvements in the various indices are inevitable and will be powered by the aggregation of long term UMR datasets and developments in bioassessment science. Therefore, we recommend that the states move forward with these indices now, as they provide adequate tools to conduct a meaningful UMR CWA assessment, recognizing that further adjustments and improvements in the future are to be expected.

Assessment Thresholds

Our analyses have determined that the currently available databases and indices allow for the selection of meaningful, attainable biological thresholds for the UMR main channel. Fundamentally different methods of derivation yielded similar answers, indicating that the process we employed was robust. Adding a preliminary BCG aided in better visualizing where along the BCG these thresholds occurred, lending more confidence to the selection of the minimum thresholds that portray the CWA 101[a][2] interim goal which is a major objective of this project. Clearly some combinations of thresholds and indices were unrealistic and these should not be considered any further. However, the thresholds that corresponded to BCG level 3 should be considered as an "upper tier" threshold as the forerunner of applying TALU concepts to the designation of aquatic life uses in the UMR main channel. A more formal BCG analysis could help refine these even further and is recommended as a future project.

What this project did not address is some of the important bioassessment implementation issues such as when does a departure from a biocriterion signify an impairment. Biological data like any other environmental parameter has an inherent degree of variation that lends to sampling methodology and inherent other sources of variation. While some of this is addressed via the critical technical elements process, this only signifies if the existing protocols are carried out in a professional and consistent manner. The states will need to make additional choices in an assessment methodology to calculate when excursions from thresholds constitute an impairment that triggers a 303[d] listing.

Sampling Design and Implementation Issues

While the conclusions herein are that a GRE style of monitoring design is desired, other important details will need to be considered in UMR monitoring strategy development. A partial list of these issues follows:

1. The GRE design is principally a stratified-random design in keeping with the fundamentals of EMAP. However, to serve the needs of multiple programs we are recommending a pollution survey design that is consistently more spatially intensive than the GRE design. Under the current GRE sampling design, some UMR assessment

reaches had as few as 3 sampling sites which is insufficient for local scale and reach scale assessment.

2. The integration of more compatible chemical/physical monitoring will need to be added to the monitoring strategy tasks. While the thresholds report did identify some chemical/physical parameters as proximate stressors, the chemical database in particular was "weak" from a frequency perspective. GRE collected one chemical sample per site in keeping with the EMAP mantra of visiting a site only one time. However, the dynamics of water chemistry dictate the need for multiple samples at a site within a seasonal index period to account for the inherent within season variability of most parameters. We are not advocating that chemical/physical data be used as a surrogate indicator for aquatic life attainment status, but rather in the role of a stress and exposure indicator to help explain the biological results.
3. A habitat assessment protocol is needed that generates an assessment of the comparative quality of UMR main channel habitat at the site level. While GRE collected various attributes of habitat, it did not assemble these variables into an index that conveys and overall quality of habitat. Such a tool will be needed as part of a longer term strategy.
4. The UMR monitoring strategy should recommend a "rotating" approach to systematic monitoring of the UMR main channel. It is highly doubtful that the entire UMR mainstem can be sampled in one year; it required three years for GRE to sample the entire mainstem. While this is resource dependent, it is reasonable to assume that multiple years will be required to pass through the UMR mainstem one time especially if modification of the GRE design adds sites. We would recommend that this "rotation" be no more than 5 years.

Data Management

The management, retrieval, and analysis of data is fundamental to any monitoring program and its execution is critical to its success and utility. Poorly designed or executed data management can greatly diminish the potential value and contributions of a monitoring program. However, good data management need not be a complex task and is surprisingly easy to accomplish, even for large programs such as what a sustained UMR effort promises to be. The data analysis accomplished for the thresholds report is an example of coping with a data management system that was initially unfamiliar to the analysts. However, because the data was managed in a relational system, accessing, retrieving, and most important understanding the data was a relatively easy task. We suggest that such a data management system be utilized for a future UMR monitoring program.

Summary and Recommendations

In summary, we recommend that the states take the information provided in this guidance to:

1. Develop a UMR-wide CWA monitoring strategy that the follows the principles outlined herein.

2. Utilize a modification of the EMAP-GRE design as the baseline spatial sampling design, i.e., execute an intensive, longitudinal "pollution survey" design.
3. Examine programmatic and organizational options for implementing such a strategy outlining the costs of each and the technical pros and cons.
4. Use the biological assemblage, biological index, and biocriteria threshold recommendations included herein as the basis for an initial biological assessment of the UMR main channel and future assessments based on a new monitoring strategy.
5. Develop and utilize a data management system that is easy to use, easy to access, and which delivers sampling data and transformed data in a portable and relational format.

References

Angradi, T.R. (editor). 2006. Environmental monitoring and assessment program, Great River ecosystems field operations manual. EPA/620/R06/002. US Environmental Protection Agency, Office of Research and Development, Washington, DC. http://www.epa.gov/emap/greatriver/fom.html

Angradi, T. R., M. S. Pearson, T. M. Jicha, D. L. Taylor, D. W. Bolgrien, M. F. Moffett, K. A. Blocksom, B. H. Hill. 2009a. Using stressor gradients to determine reference expectations for great river fish assemblages. Ecological Indicators 9:748-764

Angradi, T. R., M. S. Pearson, D. W. Bolgrien, T. M. Jicha, D. L. Taylor, and B. H. Hill. 2009b. Multimetric macroinvertebrate indices for mid-continent US great rivers. Journal of the North American Benthological Society 28:785-804.

Armitage, B. J., R. Mueller and E. T. Rankin. 2009. An Assessment of Threats to the Biological reference expectations for great river fish assemblages. Ecological Indicators Condition of the Wabash River Aquatic Ecosystem of Indiana. In Two Parts. Prepared for The Indiana Nature Conservancy, Indianapolis, IN.

Blocksom, K. A. and B. R. Johnson. 2009 Development of a regional macroinvertebrate index for large river bioassessment ecological indicators 9:313–328.

Barko, V. A., Palmer, M. W., Herzog, D. P., and B. Ickes. 2004. Influential environmental gradients and spatiotemporal patterns of fish assemblages in the unimpounded Upper Mississippi River: American Midland Naturalist 151(4): 369-385.

Barko, V. A., and D.P. Herzog. 2003, Relationships among side channels, fish assemblages, and environmental gradients in the unimpounded Upper Mississippi River: Journal of Freshwater Ecology 18(3): 377-382.

Chick, J.H. and M.A. Pegg. 2001. Invasive carp in the Mississippi River Basin. Science 292: 2250–2251.

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

Emery, E., R. Tewes, R. Argo, and J. Thomas. 2007. Development of a probability based monitoring and assessment strategy for select large rivers within U.S. EPA Region V. Report in fulfillment of EPA Grant RM-83169201. ORSANCO. Cincinnati, OH. 100 pp.

Karr, J.R. and C.O. Yoder. 2004. Biological assessment and criteria improve TMDL planning and decision-making. Journal of Environmental Engineering 130(6): 594-604.

Lyons, John, R.R. Piette, and K.W. Niermeyer. 2001. Development, validation, and application of a fish-based index of biotic integrity for Wisconsin's large warmwater rivers. Transactions of the American Fisheries Society: Vol. 130, No. 6, pp. 1077–1094.

Midwest Biodiversity Institute (MBI). 2010. Improving Water Quality Standards and Assessment Approaches for the Upper Mississippi River: UMR Clean Water Act Biological Assessment Implementation Guidance. Scope of Work for Upper Mississippi River basin Association, St. Paul, MN. 8 pp.

Miltner, R.J., Yoder, C.O. and E.T. Rankin. 2011. Preliminary Analysis of Biological Assessment Thresholds for Determining Aquatic Life Use Attainment Status in the Upper Mississippi River Mainstem. A Report Submitted to: Upper Mississippi River Basin Association, 408 St. Peter Street, St. Paul, MN 55102, Dave Hokanson, Project Manager by Midwest Biodiversity Institute, P.O. Box 21561, Columbus, OH 43221-0561

Moore, M. J.C., H. A. Langrehr, and T. R. Angradi. 2011. A submersed macrophyte index of condition for the Upper Mississippi River. Ms submitted to Ecological Indicators. Wisconsin DNR, Lacrosse, WI.

National Research Council. 2001. Assessing the TMDL approach to water quality management. National Academy Press, Washington, DC. 109 pp.

Newton, T.J., S.J. Zigler, J.T. Rogala, B.R. Gray, and M. Davis. 2011. Population assessment and potential functional roles of native mussels in the Upper Mississippi River. Aquatic Conserv: Mar. Freshw. Ecosyt. 21: 122-131.

Pitlo, J.Jr. and J.L. Rasmussen. 2004. A compendium of fishery information on the Upper Mississippi River. 3rd edition. UMRCC Special Publication. 265pp.Conservation Committee.

Rankin, E.T. and C.O. Yoder. 2011. Improving Water Quality Standards and Assessment Approaches for the Upper Mississippi River: UMR Clean Water Act Biological Assessment Implementation Guidance: Development of a Biological Condition Gradient for Fish Assemblages of the Upper Mississippi River and Development of a “Synthetic” Historical Fish Community. MBI Technical Report/2011-5-2. Submitted to UMRBA WQTF. 24 pp.

Steuck, M.J., S. Yess, J. Pitlo, A. Van Vooren, and J. Rasmussen. 2010. Distribution and relative abundance of Upper Mississippi River Fishes. Upper Mississippi River Conservation Committee, Onalaska, WI.

Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K., and R.H. Norris. 2006. Setting expectations for ecological condition of running waters: the concept of reference condition. Ecological Applications 16: 1267–1276.

Sullivan, J.S., D. Stoltenberg, S. Manoyan, J. Huang, R. Zdanowicz, and W. Redmon. 2002 Upper Mississippi River Water Quality Assessment. Summary Report, Chicago: Upper Mississippi River Conservation Committee/U.S. EPA, Chicago, IL.

Taylor, D. L., Bolgrien, D.W., Angradi, T.R., Pearson, M.S., and B. H. Hill. 2011. Habitat and hydrology condition indices for the Upper Mississippi, Missouri, and Ohio Rivers. U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, 6201 Congdon Blvd., Duluth, MN 55804, USA.

Thorp, J. E. 1992. Linkage between islands and benthos in the Ohio River, with implications for riverine management. Canadian Journal of Fisheries and Aquatic Sciences 49: 1873–1882.

Upper Mississippi River Basin Association (UMRBA). 2011. Upper Mississippi River Aquatic Life Designated Uses: Improving Protection under the Clean Water Act. Upper Mississippi River Basin Association, Water Quality Task Force. St. Paul, MN.

Upper Mississippi River Basin Association (UMRBA). 2009. Examining Biological Indicators for the Upper Mississippi River: Applications in Clean Water Act and Ecosystem Restoration Programs. Final Workshop Report. Upper Mississippi River Basin Association, St. Paul, MN.

U.S. EPA. 2005. Use of biological information to better define designated aquatic life uses in state and tribal water quality standards. Office of Water, Washington, DC. EPA 822-R-05-001. 188 pp.

Weigel, B.M. Lyons, J. and P. W. Rasmussen. 2006. Fish assemblages and biotic integrity of a highly modified floodplain river, the Upper Mississippi, and a large, relatively unimpacted tributary, the Lower Wisconsin. River Research and Applications 22(8): 923-936.

Yoder, C.O., V.L. Gordon, N.B. Kale, and D.K. Hokanson. 2010. Improving Water Quality Standards and Assessment Approaches for the Upper Mississippi River: UMR Clean Water Act Biological Assessment Implementation Guidance: Background and Scoping Report. Upper Mississippi River basin Association, St. Paul, MN. 25 pp. + Appendices.

Yoder, C.O. and M.T. Barbour. 2009. Critical technical elements of state bioassessment programs: a process to evaluate program rigor and comparability. Environ. Mon. Assess. DOI 10.1007/s10661-008-0671-1: 31-42.

Yoder, C.O. and E.T. Rankin. 1998. The role of biological indicators in a state water quality management process. J. Env. Mon. Assess. 51(1-2): 61-88.

Yoder, C.O. and 9 others. 2005. Changes in fish assemblage status in Ohio's nonwadeable rivers and streams over two decades, pp. 399-429. *in* R. Hughes and J. Rinne (eds.). Historical changes in fish assemblages of large rivers in the America's. American Fisheries Society Symposium Series.

Appendix A:
2008 and 2010 Impaired Waters Listings and Approved TMDLs on the Upper Mississippi River

Appendix Table A-1. 2008 and 2010 Impaired Waters Listings and Approved TMDLs on the Upper Mississippi River (impaired designated uses indicated in superscript).

***MINNESOTA*[1]**			***WISCONSIN*[2]**	
2008	**2010**	**St. Croix River**	**2010**	**2008**
PCBs (Fish Tissue)[FC] PFOS (Fish Tissue)[FC] Turbidity[AL] Nutrients (L. Pepin)[AR] *TMDLs approved:* Mercury (Fish Tissue)[FC] Mercury (Water)[FC]	PCBs (Fish Tissue)[FC] PFOS (Fish Tissue)[FC] Turbidity[AL] Nutrients (L. Pepin)[AR] *TMDLs approved:* Mercury (Fish Tissue)[FC] Mercury (Water)[FC]	***Reach 1 (48 mi)***	PCBs (Fish Tissue)[FC] Mercury (Water)[FC] Suspended Solids[AL] PFOS (Fish Tissue)[FC]	PCBs (Water)[FC] PCBs (Fish Tissue)[FC] Mercury (Water)[FC] Mercury (Fish Tissue)[FC] Suspended Solids[AL] PFOS (Fish Tissue)[FC]
PCBs (Fish Tissue)[FC] *TMDLs approved*: Mercury (Fish Tissue)[FC]	PCBs (Fish Tissue)[FC] *TMDLs approved*: Mercury (Fish Tissue)[FC]	**Chippewa River** ***Reach 2 (49 mi)***	PCBs (Fish Tissue)[FC] Mercury (Water)[FC] PFOS (Fish Tissue)[FC]	PCBs (Water)[FC] PCBs (Fish Tissue)[FC] Mercury (Water)[FC] Mercury (Fish Tissue)[FC]
PCBs (Fish Tissue)[FC] *TMDLs approved*: Mercury (Fish Tissue)[FC]	PCBs (Fish Tissue)[FC] *TMDLs approved*: Mercury (Fish Tissue)[FC]	**Lock & Dam 6** ***Reach 3 (21 mi)***	PCBs (Fish Tissue)[FC] Mercury (Water)[FC]	PCBs (Water)[FC] PCBs (Fish Tissue)[FC] Mercury (Water)[FC] Mercury (Fish Tissue)[FC]
PCBs (Fish Tissue)[FC] *TMDLs approved*: Mercury (Fish Tissue)[FC]	PCBs (Fish Tissue)[FC] *TMDLs approved*: Mercury (Fish Tissue)[FC]	**Root River** ***Reach 4 (63 mi)***	PCBs (Fish Tissue)[FC] Mercury (Water)[FC] - Pool 8 and Pool 10 Mercury (Fish Tissue)[FC] - Pool 9	PCBs (Water)[FC] PCBs (Fish Tissue)[FC] Mercury (Water)[FC] Mercury (Fish Tissue)[FC]
IOWA[3]				
No listing	No listing			
No listing	Aluminum[AL]	**Wisconsin River** ***Reach 5 (48 mi)***	PCBs (Fish Tissue)[FC] Mercury (Water)[FC]	PCBs (Water)[FC] PCBs (Fish Tissue)[FC] Mercury (Water)[FC] Mercury (Fish Tissue)[FC]
Mercury[FC] (Pool 12)		**Lock & Dam 11** ***Reach 6 (61 mi)***	PCBs (Fish Tissue)[FC] Mercury (Water)[FC]	PCBs (Water)[FC] PCBs (Fish Tissue)[FC] Mercury (Water)[FC] Mercury (Fish Tissue)[FC]

			ILLINOIS[4]	
	Mercury[FC] (Pool 12)		PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC]
2008	**2010**	**Lock & Dam 13**	**2010**	**2008**
Arsenic[DW] Nutrients (localized)[AL] Aluminum[AL]	Arsenic[DW] Aluminum[AL] Cadmium[AL] *TMDLs approved:* Nutrients (localized)[AL]	***Reach 7 (89 mi)***	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW]
		Iowa River		
Arsenic[DW] Indicator Bacteria[AR] Aluminum[AL]	Arsenic[DW] Bacteria[AR] Aluminum[AL] Cadmium[AL]	***Reach 8 (73 mi)***	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR] Total Dissolved Solids[DW]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR]
MISSOURI[5]		**Des Moines River**		
		Reach 9 (37 mi) **Lock & Dam 21**	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW]
No listing	No listing	***Reach 10 (88 mi)*** **Cuivre River**	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC]

TMDLs approved: PCBs[FC] Chlordane[FC]	*TMDLs approved:* PCBs[FC] Chlordane[FC]	**Reach 11 (41 mi)**	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR]
2008	**2010**	**Missouri River** **Missouri River**	**2010**	**2008**
		Reach 12 (78 mi)	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR]
Lead (localized)[AL] Zinc (localized)[AL] *TMDLs approved:* PCBs[FC] Chlordane[FC]	No listing *TMDLs approved:* PCBs[FC] Chlordane[FC] Lead (localized)[AL] Zinc (localized)[AL]	**Kaskaskia River** **Reach 13 (118 mi)** **Ohio River**	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR] Iron[AL] Dissolved Oxygen[AL] pH[AL] Total Suspended Solids[AL]	PCBs (Fish Tissue)[FC] Mercury (Fish Tissue)[FC] Manganese[DW] Fecal coliform[AR] Sulfates[DW]

Key to Designated Uses:

FC = Fish consumption
AL = Aquatic Life
AR = Aquatic Recreation/Swimming/Primary Contact
DW = Drinking Water

Note that these are generalized designated use descriptions and may vary somewhat from the specific language used by States to describe designated uses.

---

[1] 2008 Minnesota listings are from the final 2008 list as approved by U.S. EPA on June 10, 2008. 2010 Minnesota listings are from the draft list submitted to U.S. EPA in March 2010.

[2] 2008 Wisconsin listings are from the draft 2008 list submitted to U.S.EPA in July 2008. 2010 Wisconsin listings are from the draft list submitted to U.S.EPA in April 2010.

[3] 2008 Iowa listings are from the final 2008 list as approved by U.S. EPA on August 4, 2010. 2010 Iowa listings are from the draft list made available for public comment in January 2011.

[4] 2008 Illinois listings are from the final 2008 list as approved by U.S. EPA on October 22, 2008. 2010 Illinois listings are from the draft list produced by the state in April 2010.

[5] 2008 Missouri listings are from the final 2008 list as approved by U.S. EPA on December 16, 2009. 2010 Missouri listings are from the draft list approved by the U.S. EPA on April 29, 2011.

Appendix B: Biological Monitoring Gear Comparison

Table 1: Electrofishing method/ gear comparison table for main channel border fish sampling by different entities on the Upper Mississippi River.

METHOD/GEAR CATEGORY	LTRMP	U.S. EPA-GRE	U.S. EPA - NRSA	WIDNR - NRSP
PLATFORM:	5.5 meter aluminum john boat	5.5 meter aluminum john boat	John boat – size not specified	5 m aluminum john boat
POWER SOURCE:	Gasoline powered generator 5 Kw AC	Gasoline powered generator 5 Kw AC	Gasoline powered generator	Gasoline powered generator 5 Kw AC
CURRENT TYPE:	Pulsed DC	Pulsed DC	Pulsed DC	Pulsed DC
WATTAGE (AC POWER SOURCE):	3000 W constant	3000 W constant	Not specified	3000 W constant
VOLTS (DC OUTPUT):	Variable @60 Hz 25% duty cycle	?	?	Variable @60 Hz 25% duty cycle
AMPERAGE (OUTPUT):	?	?	?	?
ANODE TYPE/ LOCATION:	Boat hull	Boat hull	Not specified	Droppers, two sets off side
CATHODE TYPE/LOCATION:	2 booms, circular ring dropper array	2 booms, umbrella dropper array	2 booms, umbrella or straight dropper array	1 boom, circular ring dropper array
NUMBER OF NETTERS/ MESH SIZE:	2 netters; 3 mm (1/8”) mesh	2 netters; 3 mm (1/8”) mesh	1 netter; 6 mm (1/4”) mesh	1 netter (seated); 17 mm mesh
DISTANCE SAMPLED (Km)/ BANK(S):	0.2 Km (random bank)	1.0 Km (random bank)	40 x mean width; max 4 Km[1] 10 transects (alternating bank)	1.6 Km (either bank BPJ)
TIME SPECIFICATIONS:	900 seconds	3600 seconds	4000-9000 seconds	2400 seconds
SAMPLING DIRECTION:	Down current	Down current	Down current	Down current
CPUE BASIS:	No. or Kg/15 minutes	No. or Kg/?	No./? (no biomass)	No. or Kg/1.6 Km
SAMPLING INDEX PERIOD:	June 15 – October 30[2]	Summer-early fall	Summer-early fall	mid-June to mid-October
DAY/NIGHT:	Both, time specified	Daytime (after 1000 h)	Daytime	Daytime
FISH IDENTIFICATION	Species (AFS nomenclature)	Species (AFS nomenclature)	Species (AFS nomenclature)	Species (AFS nomenclature)
FISH ANOMALIES & DISEASE	External anomalies; 6 distinctions	DELTs + other abnormalities	DELTs + other anomalies	External anomalies

[1] Ten discrete transects delineated; minimum of five are sampled if >500 fish collected.
[2] Three discrete periods within the overall index period are specified.

Table 1. continued

METHOD/GEAR CATEGORY	MDNR	MPCA
PLATFORM:	17' alum. johnboat	17' aluminum johnboat
POWER SOURCE:	5 Kw generator; Coeffelt VVP-15 electrofisher	5 Kw generator; Smith-Root 5.0 GPP electrofisher
CURRENT TYPE:	pulsed DC	pulsed DC
WATTAGE (AC POWER SOURCE):	5000W	5000W
VOLTS (DC OUTPUT):	300V	0-500V (low); 0-1000V (high)
AMPERAGE (OUTPUT):	5-7 A	0-20 A
ANODE TYPE/ LOCATION:	Wisconsin Ring/ boom	umbrella-type droppers/ 2 booms
CATHODE TYPE/LOCATION:	boat hull; 6 - 5/16-inch stainless steel cables	boat hull
NUMBER OF NETTERS/ MESH SIZE:	2 netters; 3 mm (1/8") mesh	2 netters; 3 mm (1/8") mesh
DISTANCE SAMPLED (Km)/ BANK(S):	1.6 Km; single bank	1.5 Km: 0.5 Km each; right bank, left bank, mid-channel
TIME SPECIFICATIONS:	?	?
SAMPLING DIRECTION:	Down current	Down current
CPUE BASIS:	No./1.6 Km	No. or Kg/?
SAMPLING INDEX PERIOD:	mid-May to late September	mid-June to mid-September
DAY/NIGHT:	Daytime	Daytime

Table 2: Macroinvertebrate method/ gear comparison table for main channel invertebrate sampling by different entities on the Upper Mississippi River.

METHOD CATEGORY	LTRMP	U.S. EPA-GRE	U.S. EPA - NRSA
COLLECTION DEVICE(S):	Ponar Grab[3]; Petite Ponar	Rectangular frame (335 x 508-mm frame; 500 µm mesh) dip net	D Frame dip net (500 µm)
COLLECTION LOCATION(S):	Open channel	Shoreline margin at 11 transect locations over 0.5 Km; two habitat types 5.7 m^2 total area	Shoreline margin at 11 transect locations over 4 Km (composited); 10 x 15 m area;
EFFORT:	523 cm^2 (Ponar); 232 cm^2 (Petite Ponar) – 1 grab	5.7 m^2 total area among 11 transects	2 sweeps dominant habitat type; 1 sweep secondary habitat type
CPUE BASIS:	No. individuals/m^2	NA	NA
SUBSAMPLE SIZE:	10% of samples retained for lab analysis; whole sample is analyzed	400 organisms	500 organisms
SUBSAMPLE EXTRACTION:	1.18 mm sieve (field); hand pick entire sample	500 µm sieve (field); gridded tray (lab)	500 µm sieve (field); gridded tray (lab)
SAMPLE PRESERVATION:	70% ethanol + rose bengal	10% formalin (field); transferred 70% ethanol (lab)	95% ethanol
LABORATORY PREP:	U.S. #30 sieve	Pick sample to target 400 organism subsample size; no scan for rare/large taxa	Pick sample to target 500 organism subsample size; no scan for rare/large taxa
TAXOMOMIC RESOLUTION:	Family, Genus (fingernail clams & zebra mussel to species)	"Lowest feasible" (usually Genus)	Genus for all groups
SAMPLING INDEX PERIOD(S):	May 1 – June 14 (Pools 4, 8, 13); April 1 – June 1 (Pool 26 & LaGrange Pool); after Spring flood pulse in open channel	Summer – early fall	Summer-early fall

[3] Primary sampling device.

Appendix C: Critical technical Elements Matrices for U.S. EPA-GRE and U.S. ACE LTRMP

Appendix Table C-1. A checklist for evaluating the degree of development for each technical element of a bioassessment program and associated comments on the elements for the U.S. EPA-GRE bioassessment program. The point scale for each element ranges from lowest to highest resolution.

Element 1	(Lowest) 1.5 / 2.0	2.5 / 3.0	3.5 / 4.0	4.5 (Highest)	Comments
Index Period	Collection times are variable throughout the year, and sampling is performed without regard to seasonal influences.	An index period is conceptually recognized, but sampling may take place outside of this period for convenience or to match existing programs; sampling outside of the index is not adjusted for seasonal influences.	A well-documented seasonal index period(s) is calibrated with data for reference conditions, but sampling may take place outside of this period for convenience or to match existing programs; sampling outside of the index is adjusted for seasonal influences. Index periods are selected based on known ecology to minimize natural variability, maximize gear efficiency, and maximize the information gained about the assemblage.	Same as Level 3, but administrative needs and index periods fully reconciled. Scientific basis of temporal sampling influences management decision framework.	Formal index period of July 1 – September 30. **Points** 4.5

Element 2	(Lowest) 1.5 / 2.0	2.5 / 3.0	3.5 / 4.0	4.5 (Highest)	Comments
Spatial Coverage	An individual site is used for assessment of watershed condition; simple upstream/ downstream and fixed station designs prevail; assessments at local scale.	Multiple sites are used for watershed assessment; spatial coverage only for questions of general status or locally specific problem areas; synoptic (non-random) design at coarse scale (e.g., 8-digit HUC common); spatial extrapolation is based on "rules of thumb"; may be supplemented by simple upstream/downstream assessments.	Spatial network suitable for status assessments; system-wide spatial coverage with single purpose design at system-wide scale; may be supplemented by occasional intensive surveys.	Comprehensive spatial network suitable for reliable mainstem assessment in support of multiple water quality management programs at a more detailed scale (e.g., intensive surveys); system-wide approach or similar scheme to complete system-wide monitoring in a specified period of time; multiple spatial designs appropriate for multiple issues.	Generalized random tessellated design applied to main channel border; state-specific needs were built in; intensification was accomplished in 2006 increase reference analogs. **Points** 4.0

Appendix Table C-1. (continued)

Element 3	(Lowest) 2.0 2.5	3.0 3.5	4.0	4.5 5.0 (Highest)	Comments
Natural Classification	No partitioning of natural variability in aquatic ecosystems. Minimal classification limited to individual watersheds or basins with generalized stratification on a regional basis; does not incorporate differences in stream characteristics such as size, gradient.	Classification recognizes one stratum, usually a geographical or other similar organization such as fishery based cold or warmwater, and is applied statewide; lacks other intra-regional strata such as watershed size, gradient, elevation, temperature, etc.	Classification is based on a combination of landscape features and physical habitat structure (inter-regional); achieves highest level of classification possible by considering all relevant intra-regional strata and subcategories of specific river types.	Fully partitioned and stratified classification scheme based on a true regional approach that transcends jurisdictional (i.e., State) boundaries to strengthen inter-regional classification and recognizes zoogeographical aspects of assemblages.	Identified upper and lower mainstem as needing different metrics; fish validated; inverts. In need of more sites. **Points** 4.5

Element 4	(Lowest) 2.0 2.5	3.0 3.5	4.0	4.5 5.0 (Highest)	Comments
Criteria for Reference Sites	No criteria, except informal BPJ selection of control sites. May be little documentation and supporting rationale.	Based on "best biology", i.e., BPJ on what the best biology is in the best waterbody; minimal non-biological data used.	Non-biological criteria supported by narrative descriptors only; combine BPJ with narrative description of land use and site characteristics; may use chemical and physical data thresholds as primary filters.	Quantitative descriptors used to support non-biological criteria; characteristics of sites are such that the best biological organization expected to be supported; chemical and physical characteristics of sites used only as secondary and tertiary filters to avoid circularity in other criteria.	Developed a quantitative stressor gradient for the UMR; mix of different chemical and physical measures and attributes. **Points** 5.0

Appendix Table C-1. (continued)

Element 5	(Lowest) 1.0	1.5 2.0	2.5 3.0	3.5 4.0 (Highest)	Comments
Reference Conditions	No reference condition; presence and absence of key taxa or best professional judgment. rather than established reference conditions may constitute the basis for assessment.	Reference condition based on biology of a 'best' site or waterbody; a site-specific control or paired watershed approach may be used for assessment; regional reference sites lacking.	Reference conditions based on site-specific data, but are used in watershed scale assessments; regional reference sites are conceptually recognized, but are too few in number and/or spatial density to support the derivation of biocriteria.	Applicable regional reference conditions are established within the applicable waterbody ecotypes and aquatic resource classes; consist of multiple sites that either represent reference or are along the BCG in such a manner to allow extrapolation of expected conditions for assessing and monitoring within waterbody ecotype. Re-sampling of reference sites done systematically over a period of years.	Employed a stressor based process as an analog to reference site based approach; being a onetime effort it does not include a recalculation of the stressor gradient. **Points** 3.5

Element 6	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Taxonomic Resolution	Gross observation of biota; single assemblage only; very low taxonomic resolution (e.g., order/family level for macro-invertebrates.; family for fish by non-biologists).	Single assemblage (usually macroinvertebrates); low taxonomic resolution (e.g., family level) by experienced biologists.	Single assemblage with high taxonomic resolution (e.g., "lowest practical" i.e., genus/species); if multiple assemblages, others are lower resolution or infrequently used.	Two or more assemblages with high taxonomic resolution (e.g., "lowest practical" i.e., genus/species); capacity to use each assemblage concurrently is maintained; practitioners are certified in accordance with available offerings (e.g., NABS, state credible data provisions).	Four assemblages were included and all at lowest practicable level; practitioners are certified where such is available. **Points** 5.0

Appendix Table C-1. (continued)

Element 7	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Sample Collection	Approach is cursory and relies on operator skill and BPJ, producing highly variable and less comparable results; Training limited to that which is conducted annually for non-biologists who compose the majority of the sampling crew. Documentation of methods more as an overview.	Textbook methods are used rather than in-house development of detail of SOPs to specify methods; a QA/QC document may have been prepared; training consists of short courses (1-2 days) and is provided for new staff and periodically for all staff.	Methods are evaluated and refined (if needed) for State purposes; detailed and well documented; SOPs are updated periodically and supported by in-house testing and development; a formal QA/QC program is in place with field replication taken; rigorous training is for all professional staff, regardless of skill mix to raise skill levels and enhance interaction and consistency.	Same as Level 3, but methods cover multiple assemblages.	Methods are very well documented in SOPs; rigorous QA/QC; rigorous training & professional development. **Points** 5.0

Element 8	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Sample Processing	Biological samples are processed in the field using visual guides; sorting and identification are dependent on operator skill and effort.	Organisms are identified and enumerated primarily in the field prohibiting ample QC but done by trained staff; for fish cursory examination of presence and absence only; no in-house development of SOPs.	Laboratory processing of all samples (except for fish); A formal QA/QC program is in place; rigorous training is provided; vouchering of organisms done for ID verification.	Same as Level 3, but is applicable to multiple assemblages; subsampling level tested. Notations made on fish as to diseased, erosion, lesion, tumors.	Laboratory processing of applicable assemblages; SOPs and QA/QC in place; training provided. **Points** 5.0

Appendix Table C-1. (continued)

Element 9	(Lowest) 2.0	2.5	3.0	3.5	4.0	4.5	5.0 (Highest)	Comments
Data Management	Sampling event data organized in a series of spreadsheets e.g., (by year, by data-type, etc); QC cursory and mostly for transcription errors.		Separate quasi-databases for physical-chemical and biological data (Excel, Access, dBase, etc) with separate GIS shape files of monitoring stations; data-handling methods manuals available; QC for data entry, value ranges, and site locations.		True relational database containing biological and sampled site info (Oracle, etc); fully documented and implemented data QAPP; structure allows for data export and analysis and biocriteria development; includes dedicated database management.		Relational database of bioassessment data (including indices and biocriteria) with real-time connection to spatial data coverage showing monitored sites in relation to other relevant spatial data layers (population density; impervious surfaces; vegetation coverage, low-flight photos, nutrient concentrations, ecoregion, etc); fully documented and implemented data QAPP; data available from multiple assemblages to enable integrated analysis.	SWIM system developed by EPA; relational database that can be accessed for analysis; data entry QA/QC; applies to all assemblages. Points 5.0

Element 10	(Lowest) 1.5	2.0	2.5	3.0	3.5	4.0	4.5 (Highest)	Comments
Ecological Attributes	Linkage to the BCG or adherence to the basic ecological attributes as a foundation is lacking; simple measures of presence/absence.		Only inferences can be made for a few of the comparatively simple ecological attributes, e.g., sensitive/tolerant taxa of a ubiquitous nature; single dimension measures used.		Ecological attributes used as a foundation for bioassessment, but may not be fully developed, or may be lacking. BCG incorporated into conceptual underpinnings.		The ecological attributes of the BCG form the conceptual foundation; level of rigor represents or extends to all underpinnings of the ecological attributes.	GRE applied "IBI type" of approach which incorporates most of the underpinnings of the BCG although a formal BCG exercise was not included. Points 4.5

Appendix Table C-1. (continued)

Element 11	(Lowest) 1.0	1.5 / 2.0	2.5 / 3.0	3.5 / 4.0 (Highest)	Comments
Biological Endpoints and Thresholds	Assessment may be based only on presence or absence of targeted or key species; (Some citizen monitoring groups use this level); attainment thresholds not specified; this approach may be sufficient for Coarse problem identification. Coarse method (low signal) and detects only high and low values.	A biological index or endpoint is established for specific water bodies, but is likely not calibrated to waterbody classes or statewide application; index is probably relevant only to a single assemblage; presence/absence based on all taxa; BPJ thresholds based on single dimension attributes. Limited to pass/fail determinations of attainment status that does not reflect incremental measurement along the BCG.	A biological index, or model, has been developed and calibrated for use throughout the State or region for the various classes of a given waterbody type; the index is relevant to a single assemblage; attainment thresholds are based on discriminant model or distribution of candidate reference sites, or some means of quantifying reference condition. Can distinguish 3-4 increments along the BCG; supports narrative evaluations based on multimetric or multivariate analysis that are relevant to the BCG.	Biological index(es), or model(s) for multiple assemblages is (are) developed and calibrated for use throughout the mainstem and corresponds to the BCG; integrated assessments using the multiple assemblages are possible, thus improving both the assessment and diagnostic aspects of the process; multiple parameters for evaluation, based on integrated data calibrated to regional reference condition. Able to detect status (integrated signal) on a continuous scale along the BCG; power to detect at least 5-6 categories of condition.	Biological indices are well developed and calibrated for entire mainstem and applies to multiple assemblages; used disturbance score as analog to reference condition; uncertainty remains about ability to distinguish increments along the BCG. Points 3.5

Element 12	(Lowest) 1.0	1.5 / 2.0	2.5 / 3.0	3.5 / 4.0 (Highest)	Comments
Diagnostic Capability	Diagnostic capability lacking.	Coarse indications of response via assemblage attributes at gross level, i.e., general indicator groups (e.g., EPT taxa); Supporting analysis across spatial and temporal scales limited.	More detailed development of indicator guilds and other aggregations to distinguish and support causal associations; usually involves refined taxonomy (at least genus level); supported by analysis of larger datasets and/or extensive case studies; patterns repeatable across different sources; developed for a single assemblage only.	Response patterns are most fully developed and supported by organized and extensive research and case studies across spatial and temporal scales; results are actively used in biological assessment and in assigning associated causes and sources for program support purposes; involves refined taxonomy; accomplished for two assemblage groups.	Diatom inferred trophic index approach is being used; urban signatures have been developed for macroinvertebrates. Points 3.0

Appendix Table C-1. (continued)

Element 13	(Lowest) 1.5	2.0 2.5	3.0 3.5	4.0 4.5 (Highest)	Comments
Professional Review and Documentation	Review limited to editorial aspects.	Internal scientific review only, Outside review for objectivity left for higher levels.	Outside review of documentation and reports conducted. However, selection of peer review can be subjective.	Formal process for technical review to include multiple reference and documented system for reconciliation of comments and issues. Process results in methods and reporting improvements. Can include peer-reviewed journal publications.	Subject to ORD review process; extensive record of peer review publications.
					Points 4.5

CE Score = 57.5
CE % = 95.8%
Level = Level 4 (>95.0%)

Appendix Table C-2. A checklist for evaluating the degree of development for each technical element of a bioassessment program and associated comments on the elements for the U.S. ACE-LTRMP bioassessment program. The point scale for each element ranges from lowest to highest resolution.

Element 1	(Lowest) 1.5 2.0 2.5 3.0 3.5 4.0 4.5 (Highest)				Comments
Index Period	Collection times are variable throughout the year, and sampling is performed without regard to seasonal influences.	An index period is conceptually recognized, but sampling may take place outside of this period for convenience or to match existing programs; sampling outside of the index is not adjusted for seasonal influences.	A well-documented seasonal index period(s) is calibrated with data for reference conditions, but sampling may take place outside of this period for convenience or to match existing programs; sampling outside of the index is adjusted for seasonal influences. Index periods are selected based on known ecology to minimize natural variability, maximize gear efficiency, and maximize the information gained about the assemblage.	Same as Level 3, but administrative needs and index periods fully reconciled. Scientific basis of temporal sampling influences management decision framework.	Fish: June 15-October 31 with 3 subperiods to maintain continuity and support randomized design. SAV: June 15-August 15 based on peak biomass production. Points 4.5

Element 2	(Lowest) 1.5 2.0 2.5 3.0 3.5 4.0 4.5 (Highest)				Comments
Spatial Coverage	An individual site is used for assessment of watershed condition; simple upstream/ downstream and fixed station designs prevail; assessments at local scale.	Multiple sites are used for river reach assessment; spatial coverage only for questions of general status or locally specific problem areas; synoptic (non-random) design at coarse scale; spatial extrapolation is based on "rules of thumb"; may be supplemented by simple upstream/downstream assessments.	Spatial network suitable for status assessments; system-wide spatial design using river reaches with single purpose design at coarse scale; may be supplemented by occasional intensive surveys.	Comprehensive spatial network suitable for reliable watershed assessments in support of multiple water quality management programs at more detailed scale; system-wide river reach approach or similar scheme to complete system-wide monitoring in a specified period of time; multiple spatial designs appropriate for multiple issues.	(see Brian's notes) Points 3.0

Appendix Table C-2. (continued)

Element 3	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Natural Classification	No partitioning of natural variability in aquatic ecosystems. Minimal classification limited to individual watersheds or basins with generalized stratification on a regional basis; does not incorporate differences in stream characteristics such as size, gradient.	Classification recognizes one stratum, usually a geographical or other similar organization such as fishery based cold or warmwater, and is applied statewide; lacks other intra-regional strata such as watershed size, gradient, elevation, temperature, etc.	Classification is based on a combination of landscape features and physical habitat structure (inter-regional); achieves highest level of classification possible by considering all relevant intra-regional strata and subcategories of riverine habitat types.	Fully partitioned and stratified classification scheme based on a true regional approach that transcends jurisdictional (i.e., State) boundaries to strengthen inter-regional classification and recognizes zoogeographical aspects of assemblages.	Points 5.0

Element 4	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Criteria for Reference Sites	No criteria, except informal BPJ selection of control sites. May be little documentation and supporting rationale.	Based on "best biology", i.e., BPJ on what the best biology is in the best waterbody; minimal non-biological data used.	Non-biological criteria supported by narrative descriptors only; combine BPJ with narrative description of land use and site characteristics; may use chemical and physical data thresholds as primary filters.	Quantitative descriptors used to support non-biological criteria; characteristics of sites are such that the best biological organization expected to be supported; chemical and physical characteristics of sites used only as secondary and tertiary filters to avoid circularity in other criteria.	LTRMP has not selected reference sites, but is well equipped to do so if tasked; LTRMP selected reference sites for EMAP-GRE. Points 5.0

Appendix Table C-2. (continued)

Element 5	(Lowest) 1.0	1.5 / 2.0	2.5 / 3.0	3.5 / 4.0 (Highest)	Comments
Reference Conditions	No reference condition; presence and absence of key taxa or best professional judgment. rather than established reference conditions may constitute the basis for assessment.	Reference condition based on biology of a 'best' site or waterbody; a site-specific control or paired watershed approach may be used for assessment; regional reference sites lacking.	Reference conditions based on site-specific data, but are used in watershed scale assessments; regional reference sites are conceptually recognized, but are too few in number and/or spatial density to support the derivation of biocriteria.	Applicable regional reference conditions are established within the applicable waterbody ecotypes and aquatic resource classes; consist of multiple sites that either represent reference or are along the BCG in such a manner to allow extrapolation of expected conditions for assessing and monitoring within waterbody ecotype. Re-sampling of reference sites done systematically over a period of years.	Long term documentation of a quality gradient and determination of non-random patterns supports this being sufficient. Points 4.0

Element 6	(Lowest) 2.0	2.5 / 3.0	3.5 / 4.0	4.5 / 5.0 (Highest)	Comments
Taxonomic Resolution	Gross observation of biota; single assemblage only; very low taxonomic resolution (e.g., order/family level for macro-invertebrates.; family for fish by non-biologists).	Single assemblage (usually macroinvertebrates); low taxonomic resolution (e.g., family level) by experienced biologists.	Single assemblage with high taxonomic resolution (e.g., "lowest practical" i.e., genus/species); if multiple assemblages, others are lower resolution or infrequently used.	Two or more assemblages with high taxonomic resolution (e.g., "lowest practical" i.e., genus/species); capacity to use each assemblage concurrently is maintained; practitioners are certified in accordance with available offerings (e.g., NABS, state credible data provisions).	Fish and SAV are done to lowest practicable levels; no certification available, but skill levels are equivalent. Points 5.0

Appendix Table C-2. (continued)

Element 7	(Lowest) 2.0	2.5 / 3.0	3.5 / 4.0	4.5 / 5.0 (Highest)	Comments
Sample Collection	Approach is cursory and relies on operator skill and BPJ, producing highly variable and less comparable results; Training limited to that which is conducted annually for non-biologists who compose the majority of the sampling crew. Documentation of methods more as an overview.	Textbook methods are used rather than in-house development of detail of SOPs to specify methods; a QA/QC document may have been prepared; training consists of short courses (1-2 days) and is provided for new staff and periodically for all staff.	Methods are evaluated and refined (if needed) for State purposes; detailed and well documented; SOPs are updated periodically and supported by in-house testing and development; a formal QA/QC program is in place with field replication taken; rigorous training is for all professional staff, regardless of skill mix to raise skill levels and enhance interaction and consistency.	Same as Level 3, but methods cover multiple assemblages.	Points 5.0

Element 8	(Lowest) 2.0	2.5 / 3.0	3.5 / 4.0	4.5 / 5.0 (Highest)	Comments
Sample Processing	Biological samples are processed in the field using visual guides; sorting and identification are dependent on operator skill and effort.	Organisms are identified and enumerated primarily in the field prohibiting ample QC but done by trained staff; for fish cursory examination of presence and absence only; no in-house development of SOPs.	Laboratory processing of all samples (except for fish and SAV); A formal QA/QC program is in place; rigorous training is provided; vouchering of organisms done for ID verification.	Same as Level 3, but is applicable to multiple assemblages; subsampling level tested. Notations made on fish as to diseased, erosion, lesion, tumors.	Need to more formally standardize fish DELT procedures. Points 4.5

Appendix Table C-2. (continued)

Element 9	(Lowest) 2.0 / 2.5	3.0 / 3.5	4.0 / 4.5	5.0 (Highest)	Comments
Data Management	Sampling event data organized in a series of spreadsheets e.g., (by year, by data-type, etc); QC cursory and mostly for transcription errors.	Separate quasi-databases for physical-chemical and biological data (Excel, Access, dBase, etc) with separate GIS shape files of monitoring stations; data-handling methods manuals available; QC for data entry, value ranges, and site locations.	True relational database containing biological and sampled site info (Oracle, etc); fully documented and implemented data QAPP; structure allows for data export and analysis and biocriteria development; includes dedicated database management.	Relational database of bioassessment data (including indices and biocriteria) with real-time connection to spatial data coverage showing monitored sites in relation to other relevant spatial data layers (population density; impervious surfaces; vegetation coverage, low-flight photos, nutrient concentrations, ecoregion, etc); fully documented and implemented data QAPP; data available from multiple assemblages to enable integrated analysis.	Probably exceed 5.0 score. Points 5.0

Element 10	(Lowest) 1.5 / 2.0	2.5 / 3.0	3.5 / 4.0	4.5 (Highest)	Comments
Ecological Attributes	Linkage to the BCG or adherence to the basic ecological attributes as a foundation is lacking; simple measures of presence/absence.	Only inferences can be made for a few of the comparatively simple ecological attributes, e.g., sensitive/tolerant taxa of a ubiquitous nature; single dimension measures used.	Ecological attributes used as a foundation for bioassessment, but may not be fully developed, or may be lacking. BCG incorporated into conceptual underpinnings.	The ecological attributes of the BCG form the conceptual foundation; level of rigor represents or extends to all underpinnings of the ecological attributes.	Fish: no formal process, but program capacity exists. SAV: working index done in typical MMI fashion. Points 4.0

Appendix Table C-2. (continued)

Element 11	(Lowest) 1.0	1.5	2.0	2.5	3.0	3.5	4.0 (Highest)	Comments
Biological Endpoints and Thresholds	Assessment may be based only on presence or absence of targeted or key species; (Some citizen monitoring groups use this level); attainment thresholds not specified; this approach may be sufficient for Coarse problem identification. Coarse method (low signal) and detects only high and low values.		A biological index or endpoint is established for specific water bodies, but is likely not calibrated to waterbody classes or statewide application; index is probably relevant only to a single assemblage; presence/absence based on all taxa; BPJ thresholds based on single dimension attributes. Limited to pass/fail determinations of attainment status that does not reflect incremental measurement along the BCG.		A biological index, or model, has been developed and calibrated for use throughout the system for the various classes of a given waterbody type; the index is relevant to a single assemblage; attainment thresholds are based on discriminant model or distribution of candidate reference sites, or some means of quantifying reference condition. Can distinguish 3-4 increments along the BCG; supports narrative evaluations based on multimetric or multivariate analysis that are relevant to the BCG.		Biological index(es), or model(s) for multiple assemblages is (are) developed and calibrated for use throughout the State or region and corresponds to the BCG; integrated assessments using the multiple assemblages are possible, thus improving both the assessment and diagnostic aspects of the process; multiple parameters for evaluation, based on integrated data calibrated to regional reference condition. Able to detect status (integrated signal) on a continuous scale along the BCG; power to detect at least 5-6 categories of condition.	Fish: current LTRMP can distinguish 4 categories with Lyons index; need to test GRFin. SAV index just developed, not tested. **Points** 3.0

Element 12	(Lowest) 1.0	1.5	2.0	2.5	3.0	3.5	4.0 (Highest)	Comments
Diagnostic Capability	Diagnostic capability lacking.		Coarse indications of response via assemblage attributes at gross level, i.e., general indicator groups (e.g., EPT taxa); Supporting analysis across spatial and temporal scales limited.		More detailed development of indicator guilds and other aggregations to distinguish and support causal associations; usually involves refined taxonomy (at least genus level); supported by analysis of larger datasets and/or extensive case studies; patterns repeatable across different sources; developed for a single assemblage only.		Response patterns are most fully developed and supported by organized and extensive research and case studies across spatial and temporal scales; results are actively used in biological assessment and in assigning associated causes and sources for program support purposes; involves refined taxonomy; accomplished for two assemblage groups.	**Points** 2.5

Appendix Table C-2. (continued)

Element 13	(Lowest) 1.5	2.0 2.5	3.0 3.5	4.0 4.5 (Highest)	Comments
Professional Review and Documentation	Review limited to editorial aspects.	Internal scientific review only, Outside review for objectivity left for higher levels.	Outside review of documentation and reports conducted. However, selection of peer review can be subjective.	Formal process for technical review to include multiple reference and documented system for reconciliation of comments and issues. Process results in methods and reporting improvements. Can include peer-reviewed journal publications.	Points 4.5

CE Score = 55.0
CE % = 91.7%
Level = Level 3+ (85-94.9%)