

Critical technical elements of state bioassessment programs: a process to evaluate program rigor and comparability

Chris O. Yoder · Michael T. Barbour

Received: 26 November 2007 / Accepted: 24 July 2008
© Springer Science + Business Media B.V. 2008

Abstract We developed a systematic process to evaluate state/tribal bioassessment programs to provide information about the rigor of the technical approach. This is accomplished via on-site interviews to produce an evaluation that assigns one of four levels of rigor as an outcome. Level 4 is the most rigorous and reflects a technical capacity to accurately determine incremental condition and support management programs. The remaining three levels are less able to assess incremental condition and are appropriate for only some management support needs. Accurately determining impairment and diagnosing pollution-specific stressors are fundamental tasks that states/tribes must accomplish to provide management support. This goal is fulfilled to varying degrees by most states/tribes. The evaluation employs a checklist and a sliding scale of rigor for 13 technical elements. Feedback is provided to each state/tribe via a technical memorandum that describes the technical components of the monitoring program,

highlights strengths, and recommends improvements for specific technical issues. This can be used to refine the bioassessment and monitoring programs to better support management programs. The results of 14 state/tribal evaluations are included here. The majority (nine states, one tribe) revealed that most operate at level 2 with developmental activities that will elevate the level of program rigor already underway. Two states operate level 4 programs and each have numeric biocriteria and refined designated uses in their water quality standards. This is the ultimate goal of the process of engaging states in the development of bioassessment programs in the U.S.

Keywords Aquatic life uses · Designated uses · Biological assessments · Biological criteria · Tiered uses · State/tribal programs

Introduction

Biological assessments have been a part of state and tribal water quality monitoring and assessment programs for more than three decades. However, the technical rigor of each varies widely and among a variety of technical components and attributes that we term here as critical technical elements. Few U.S. states have adopted narrative biocriteria and only three have adopted numeric biological criteria and refined designated uses in

C. O. Yoder (✉)
Center for Applied Bioassessment & Biocriteria,
Midwest Biodiversity Institute, P.O. Box 21561,
Columbus, OH 43221-0561, USA
e-mail: yoder@rrohio.com

M. T. Barbour
Center for Ecological Sciences, Tetra Tech,
Owings Mills, MD, USA

their water quality standards (WQS). Their experiences, along with those of federal agencies (U.S. Environmental Protection Agency [U.S. EPA] and U.S. Geological Survey [USGS]) and national initiatives such as the National Water Quality Monitoring Council (NWQMC), and its predecessor the Intergovernmental Task Force on Monitoring Water Quality (ITFM), have underscored the need to better define the underlying technical and programmatic components of an adequate program (ITFM 1992, 1995; Yoder 1998).

It is an important national goal to have state and tribal programs adopt refined designated aquatic life uses and numeric biocriteria in their state water quality standards (WQS). This will not only lead to more accurate identifications and descriptions of impaired waters via the use of numeric biocriteria, but lead to better planning and management decisions via more specific and complete WQS. 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 each 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 as a baseline for determining what types of improvements would be needed to attain these objectives. Taken together the EPA guidance and the critical elements process provide detailed guidelines and milestones by which states, tribes, and U.S. EPA can evaluate and track progress in the development and implementation of more rigorous biological assessment programs in support of reaching the goal of having tiered aquatic life uses (TALU) and numeric biocriteria in state and tribal WQS.

This initiative also recognizes that biological assessment is an essential component of adequate state and tribal water quality monitoring and assessment programs that includes the integration of other chemical, physical, and environmental measures and indicators (ITFM 1992, 1995; Yoder 1998). The term “adequate” was deliber-

ately chosen to describe the type of monitoring that is needed to support a TALU-based approach¹ and to provide important context about what is expected of state and tribal monitoring and assessment programs, i.e., employing a cost-effective array of parameters and indicators following a systematic process for their inclusion depending on programmatic needs (Yoder 1998). It attempts to avoid a digression towards a rote set of “minimum” requirements where the inclusion of certain indicators can be viewed as a non-essential luxury or an open ended process suggested by “comprehensive” in which cost-effectiveness can easily become lost in the pursuit of more indicators and data points.

While the critical elements process does not directly assess the technical components of allied chemical and physical assessment, their integration with biological assessment is expected and is crucial to meeting important state, tribal, and national water quality management goals and objectives. The question of *how* biological data should be integrated with chemical and physical assessment tools and indicators is addressed as part of the overall state or tribal monitoring and assessment and WQS program review of which the critical elements evaluation is an important part (e.g., MBI 2004). Biological assessments and biocriteria are intended to serve as a direct measure of designated aquatic life use attainment and to aid in the diagnosis of the causes of biological impairment and their sources (U.S. EPA 2005). A sufficiently rigorous biological assessment program fosters a more effective diagnosis of causal associations and their targeting by management programs. This is but one incentive for states and tribes to develop the capacity for delivering a high level of biological assessment that is based on reasonably available methods and best practices.

Taken together the overarching goal of what has become known as the EPA TALU-based

¹The 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.

approach is to achieve the better integration of monitoring and assessment and WQS (U.S. EPA 2005). Presently, each program can be implemented in an independent manner and be viewed as an acceptable structure for water quality management under the U.S. Clean Water Act. However, the synergy that is fostered by the TALU-based process in these two major program areas provides the underlying support for producing better management outcomes. This was strongly implied by the NRC (2001) in their review of the EPA TMDL (total maximum daily load) process and it involved the same fundamentals that are addressed by the TALU-based approach to water quality management.

Critical technical elements of state bioassessment programs

To fulfill the need for a systematic, baseline evaluation of state and tribal bioassessment programs we developed the *Critical Technical Elements of Bioassessment Program* (Barbour and Yoder 2007). While 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 program. While we recognize here that different technical approaches can achieve similar levels of rigor, it is likely that some will produce differences in the level of rigor that are

revealed by the technical elements. The critical elements consist of 13 technical attributes of a biological assessment program and these can be grouped into three distinct areas; program design, methods, and data interpretation (Table 1). 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. Written feedback is provided to the state or tribe via a technical memorandum and critical elements checklist that describes the status of each technical element and what is needed to elevate 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.

In developing the methodology, we consulted several sources of information and experience. Key among these are the U.S. EPA *Comprehensive Assessment and Listing Methodology (CALM)* process guidelines (U.S. EPA 2002b), *Important Elements and Concepts of an Adequate Watershed Monitoring and Assessment Program* (Yoder 1998), the most recent national survey of state and tribal biological assessment programs (U.S. EPA 2002a), the *Region V State Bioassessment and Ambient Monitoring Programs: Initial Evaluation and Review* (MBI 2004), the NWQMC

Table 1 The critical technical elements of a bioassessment program or protocol showing a sliding scale of associated resolution and degree of development for each

		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. Reference conditions	1	2	3	4
	5. Criteria for reference sites	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
Interpretation	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

Data Collection Methods Board (Federal Register No. 66(52): 15273–15275), and contemporary technical literature on bioassessments and biocriteria. The critical elements process also builds upon existing technical guidance documents for streams (U.S. EPA 1996), lakes and reservoirs (U.S. EPA 1998), and estuaries (U.S. EPA 2000), the Rapid Bioassessment Protocols (U.S. EPA 1999), the U.S. EPA Environmental Monitoring and Assessment Program (EMAP), and leading state programs (U.S. EPA 2005). Thus far we have focused on how states and tribes assess the biota of flowing waters, wadeable perennial streams in particular as these are the most in common waterbody type that is routinely sampled by all states and tribes. However, we have piloted evaluations of other waterbody types (e.g., non-wadeable rivers, wetlands) when these are routinely included in a state or tribal program. As such the technical evaluation is specific to a waterbody ecotype (e.g., wadeable streams, non-wadeable large rivers, wetlands, lakes, estuaries, etc.).

The guiding principles of the critical elements approach are intended to help state and tribal monitoring and assessment programs achieve levels of standardization, rigor, reliability, and reproducibility that are *reasonably attainable* under current technology and reasonable levels of funding. In turn, this will produce an accurate, comparable, comprehensive, and cost-effective monitoring and assessment program that is capable of meeting the broad goal of supporting all relevant water quality management programs (Fig. 1). An important goal of this process, in concert with the other activities of the national biological criteria program, is an adherence to the following principles:

Accuracy biological assessments should produce sufficiently accurate delineations of condition so that type I and II assessment errors are minimized;

Comparability bioassessment programs that utilize different technical approaches should produce comparable assessments in terms of biological condition ratings, detection of impairments, and diagnostic properties (Houston et al. 2002; Bonada et al. 2006);

Adequate Monitoring & Assessment Supports Multiple Water Quality Management Programs

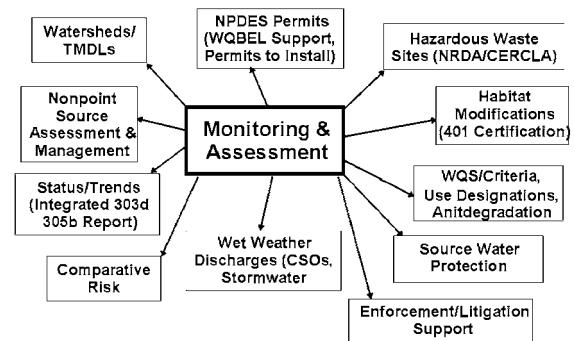


Fig. 1 The multiple uses of monitoring and assessment information in support of specific water quality management programs that are expected as an outcome of the critical elements and state/tribal review process

Comprehensiveness biological assessments should be integrated with chemical, physical, and other stressor and exposure indicators, each used in their respective indicator roles (Yoder and Rankin 1998) to demonstrate the relationship between human caused impacts and biological response; and,

Cost-effectiveness the term as used here means that the benefits of having a rigorous and reliable biological assessment program to support making better management decisions outweighs the intrinsic costs of program development and implementation (NRC 2001).

The critical elements process is intended primarily for use by state and tribal program managers and staff who are responsible for the monitoring and assessment and WQS programs and in cooperation with their regional EPA counterparts. First, states and tribes can use the guidelines to determine where they are in the biological assessment and criteria development process and how to develop, structure, and, if necessary, modify their programs. Examples of this already exist in selected states and are described in the critical elements documentation (Barbour and Yoder 2007). U.S. EPA envisions a process by which the critical elements process will provide an opportunity to better engage states and tribes regarding methods and approaches, design and implementation issues, and resource needs (U.S. EPA 2005).

The critical elements process should also provide useful information for communicating the quality of the bioassessment and biocriteria program to EPA and state and tribal program managers, and for identifying attainable milestones, allocating resources, and tracking progress.

Methods

The review of a state or tribal monitoring and assessment program requires an understanding of the current program elements—both technical and programmatic—and a 3-day site visit to discuss and evaluate the monitoring and assessment and WQS programs with managers and staff. The review is structured in accordance with a general outline that includes the baseline technical and management programs of a state or tribal water quality program. A detailed agenda is developed in advance of each visit with the appropriate state/tribe and EPA regional participants. All of the guidance and methods documents and supporting materials are supplied to the participants in advance of each review.

Purpose of program evaluation

The 3 day evaluation is intended to gather baseline information about the state or tribal monitoring and assessment and WQS programs. It includes an assessment of the technical approaches and how each is used to support and water quality management programs and assess their outcomes. As a result, recommendations for enhancements dealing with the design, methodology, and execution of the monitoring and assessment and WQS programs are made to the state or tribe. An important goal is for the state or tribe to produce credible data as a basis for making informed decisions regarding the ecological condition of the aquatic resources based on the development and implementation of a TALU-based process.

Attendance at the program review

Managers and staff who can speak to the operation and management of the WQS, report-

ing and listing (305b/303d integrated report and listings, TMDL development and implementation), watershed planning, nonpoint source assessment and management, dredge and fill (CWA 404/401), and NPDES permitting programs are asked to participate. The evaluation process consists of direct interactions with the respective program management and staff to evaluate the status of their bioassessment, monitoring and assessment, and WQS programs and to describe how each is presently used to support water quality management.

Basis for program evaluation

Since 2000, EPA has been developing and implementing a process for incorporating tiered aquatic life uses (TALUs) and numeric biocriteria in state and tribal monitoring and assessment and WQS programs. Successfully implementing a TALU-based approach is directly dependent on the rigor, comprehensiveness, and integration of the monitoring and assessment and WQS programs. TALUs play a key role in determining not only the WQS that are applied in a given management scenario, but also in determining the extent and severity of impaired waters through the application of numeric biocriteria via adequate monitoring and assessment. Hence, the development and implementation of TALUs may alter prior determinations and actions that were based on general uses and less than adequate monitoring and assessment.

An important task for each state and tribe and EPA in general will be to manage the transition to a TALU-based approach to water quality management while fulfilling programmatic obligations that are already in place. This is especially true in states with TMDL consent decrees in which the present determination of impaired waters was based on WQS consisting of general or non-TALU use designations and less rigorous biological assessments or stand alone chemical monitoring and indicators. Nevertheless, EPA believes that the long term benefits of developing and implementing TALU-based WQS and monitoring and assessment will result in more accurate decisions about impaired waters and better water quality management programs in general.

Overview and summary of the state's or tribe's programs

To put the evaluation into perspective, a summary of the current program is prepared that addresses aspects of; (1) the state or tribal surface water monitoring and assessment program, (2) the structure of the existing WQS (emphasizing designated aquatic life uses), (3) the development of bioassessment tools (indicators, assemblages, indices, models), (4) how biological data is presently used to delineate impaired waters, determine associated stressor effects, and foster the integration of WQS and monitoring and assessment, and, (5) the relationship of biological data to other chemical/physical data and indicators and water quality management program uses of this information including regulatory (e.g., NPDES permits) and other management applications (e.g., TMDLs, watershed management). This summary includes appropriate references to the state or tribal regulations and administrative codes and a description of the structure and content of aquatic life designated uses. If applicable, the structure and content of state or tribal biological criteria are also described as are any provisions relating to the use of ambient monitoring and assessment data and results. It also includes a description of any guidance or similar documentation for conducting use attainability analyses if applicable. Any aspects of the WQS that deal with data quality objectives including references to standard operating procedures are also described.

Bioassessment program evaluation

The technical components of the bioassessment program are reviewed and evaluated as part of the critical elements process. An evaluation of the 13 technical elements is completed using a standardized checklist (Appendix Table 1) and scoring methodology of which the results are reported as part of the technical memorandum. This is accomplished by proceeding through the critical elements checklist with the state or tribal technical staff and in accordance with the methodology (Barbour and Yoder 2007). The process attempts to reach a consensus among all participants about the score awarded to each technical

element by finding which level of rigor description best fits the current program. It is important here to exclude developmental activities that have not been fully implemented, but also recognize how these will advance the program in the technical memorandum. The evaluation yields a raw score that is converted to a percentage, which is then used to derive the overall level of rigor of the state or tribal program based on established scoring ranges. The percent score is the quantitative output of the evaluation process and serves as a score for making comparisons between different points in time based on subsequent evaluations. As such the critical elements process can serve as a monitoring and progress documentation tool for the states and tribes and EPA.

Recommendations

A summary of recommendations for elevating the level of rigor by each of the 13 critical technical elements is provided in the technical memorandum. The overall review process is aimed at the ultimate goal of the state or tribe adopting a TALU-based approach to monitoring and assessment and WQS (including numeric biocriteria) based on the development and maintenance of a level 4 bioassessment program (U.S. EPA 2005; Barbour and Yoder 2007). Many states have ongoing developmental activities that will directly affect some or all of the critical elements. Based on information garnered from the state and tribal review process these are described about how each will eventually change the critical element score. This is where the critical elements process serves as a tool for predicting and monitoring state and tribal progress. It also serves as a place to highlight where such developmental activities are needed to fill gaps in the existing program.

Results

Critical technical elements reviews of 14 states and one tribe were conducted between January 2004 and July 2007. Generally these were accomplished as part of the same 3-day review of the overall monitoring and assessment and WQS programs that was described in "Methods" section.

Table 2 Critical technical elements review results and specific recommendations for elevating scores to the highest level of rigor for an individual state

Element	Comment
<p>Element 1: Index period Maximum score = 4.5 Score assigned = 3.0</p>	<p>The score of 3.0 reflects a varied adherence to a seasonal index period. Logistical bottlenecks seem to be the principal reason for deviations that can extend into the following spring of each year. Elevating the score for this element will require a strict adherence to the prescribed index period.</p>
<p>Element 2: Spatial coverage Maximum score = 4.5 Score assigned = 3.0</p>	<p>The current score of 3.0 conservatively reflects the synoptic design and spatial density of sampling sites that is currently employed. Elevating the current score to the maximum of 4.5 will require increased spatial density within watershed assessment units particularly going beyond the “pour point” as the primary sampling site on a particular river or stream.</p>
<p>Element 3: Natural classification Maximum score = 5.0 Score assigned = 3.5</p>	<p>The CE score of 3.5 should be elevated to 5.0 with the developments that are already underway including the addition of new regional reference sites and the fuller inclusion of the other bioregions.</p>
<p>Element 4: Criteria for reference sites Maximum score = 5.0 Score assigned = 4.5</p>	<p>As the criteria (site-scoring process) for reference site screening is refined, the CE score of 4.5 should improve to 5.0 as it is being employed in the selection of new regional reference sites.</p>
<p>Element 5: Reference conditions Maximum score = 4.0 Score assigned = 3.0</p>	<p>The CE score of 3.0 should improve to 4.0 with the additional regional reference sites that are being established as part of the ongoing improvements described for elements 3 and 4.</p>
<p>Element 6: Taxonomic resolution Maximum score = 5.0 Score assigned = 4.5</p>	<p>The CE score of 4.5 reflects the full development of the macro-invertebrate assemblage and the in progress development of second and third assemblages. Reaching the CE score of 5.0 is contingent on the full development and use of a second assemblage.</p>
<p>Element 7: Sample collection Maximum score = 5.0 Score assigned = 4.5</p>	<p>The CE score of 4.5 reflects the full development of the macro-invertebrate assemblage (i.e., for one bioregion only) and the in-progress development of a second and third assemblage. Reaching the CE score of 5.0 is contingent on the full development and use of a second assemblage and for all applicable bioregions.</p>
<p>Element 8: Sample processing Maximum score = 5.0 Score assigned = 4.5</p>	<p>The CE score of 4.5 reflects the full development of the macroinvertebrate assemblage for a single bioregion and the in-progress development of the other bioregions and a second and third assemblage. Reaching the CE score of 5.0 is contingent on the full development and use of a second assemblage.</p>
<p>Element 9: Data management Maximum score = 5.0 Score assigned = 4.0</p>	<p>The CE score of 4.0 can be improved to 5.0 once the current data management system includes all data (i.e., habitat and fish) and is readily accessible.</p>
<p>Element 10: Ecological attributes Maximum score = 4.5 Score assigned = 3.0</p>	<p>The CE score of 3.0 should increase with the development of the macroinvertebrate index for all bioregions. A descriptive analysis of the biological condition gradient (BCG) for each representative bioregion and application of these concepts to the full development of the biological indicators and assemblages will improve the score to 4.5.</p>
<p>Element 11: Biological endpoints & thresholds Maximum score = 4.0 Score assigned = 2.5</p>	<p>The CE score of 2.5 will be increased to at least 3.0 with the full development of the macroinvertebrate index and the derivation of appropriately detailed numeric biocriteria. Achieving a score of 4.0 will require that this also be accomplished for a second biological assemblage.</p>

Table 2 (continued)

Element	Comment
Element 12: Diagnostic capability Maximum score = 4.0 Score assigned = 2.5	The comparatively low CE score of 2.5 is a common characteristic of bioassessment programs that are in development and/or which have singularly been focused on status assessments. Improving the score for this element will occur as a result of addressing preceding elements 2, 3, 6, 10, and 11 and gaining a familiarity with how diagnostic capacity is developed. This will require some dedication to exploratory analyses in which the response of the biological assemblages is evaluated along the stressor axis of the BCG.
Element 13: Professional review Maximum score = 4.5 Score assigned = 2.5	The CE score of 2.5 can be elevated to 4.5 by instituting a more formal peer review process and also by publishing some of the ongoing developments in peer reviewed journals.

All of 14 states and one tribe were visited at least once between 2004 and 2007 with six being visited on multiple occasions. The review process was initiated in 2002 for pilot testing in EPA Region V and included all six of the states in that region (MBI 2004). For those states that were reviewed more than once, the most recent CE evaluation results are reported here. The other states and one tribe were evaluated either as part of a similar regional process, in response to direct requests by a state, tribe, or the EPA regional office. As such, the collection of reviewed states and single tribe does not represent a random sample, but it does include a geographical representation of the Midwestern, Eastern, and Western U.S. One general observation that we can make is that the specific technical and management issues faced by each state and tribe is strongly influenced by specific regional issues and phenomena.

Supporting documentation that is reviewed about these programs includes the state or tribal monitoring strategy, technical procedures, technical reports, and WQS, and these are utilized prior to, during, and following each review. A technical memorandum is then produced within 10–15 weeks by a two person review team and provided to the state or tribe and the EPA regional staff for review and comment. A summary of the critical technical elements review results of a single state appears in Table 2.

The results of the 14 case example reviews indicate that nine states and one tribe function at a technical rigor consistent with a level 2 program (Fig. 2; identities are anonymous). Of the remain-

ing states two are consistent with level 4, two are at level 3, and one at level 1. There were no strong geographic or jurisdictional patterns evident in the results. One observation that we made during these evaluations is that the higher level programs have their monitoring and assessment and WQS programs either functionally or administratively merged. In the case of the two level 4 states, the WQS and monitoring and assessment programs were overseen by the same senior level manager.

A frequency plot of the scores for each of the 13 technical elements reveals a wide range of results among the states particularly in the design and interpretation elements (Fig. 3). The median and upper quartile (75th percentile) scores were the

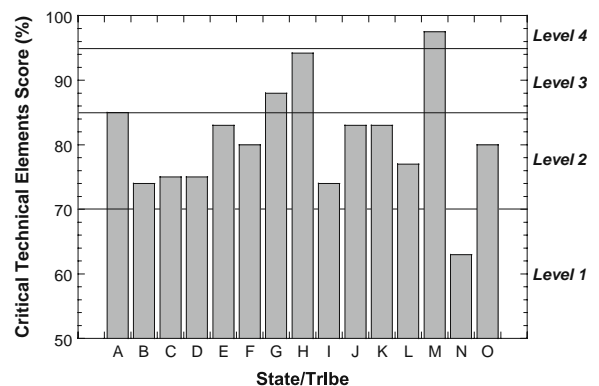


Fig. 2 Results of the most recent critical technical elements reviews of 14 state and one tribal bioassessment program conducted between January 2004 and July 2007. The results are expressed as a percentile of the total possible score and indexed to the four levels of rigor

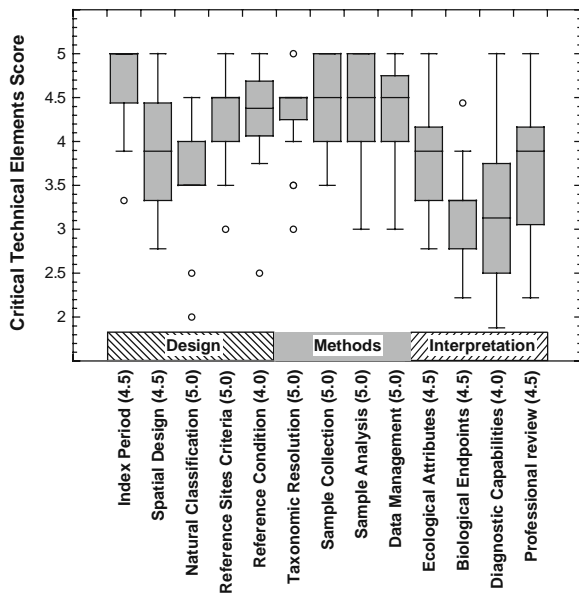


Fig. 3 Box-and-whisker plot of the range of scores assigned to 14 states and one tribe for each of 13 critical technical elements listed on the x-axis (maximum score for each element is indicated) and further aggregated by the three functional categories (design, methods, or interpretation) that each element occupies. The scores for each element are pro-rated to yield equivalent scores for comparison purposes

highest for the methods elements and the index period element and lowest for the interpretation elements, especially for the biological endpoints and diagnostic capacity elements. These results are not surprising given that a program would be the most developed for methods first and then followed by the design and interpretation elements. The latter is a good example of the dependency of the interpretation elements on having the methods and design elements addressed and in place. Diagnostic capacity used here means the ability to determine categorical stressors based on the “signatures” of response found in the biological data (Simon 2003). While some working examples of applying such tools exist (Eagleson et al. 1990; Yoder and Rankin 1998; Riva-Murray et al. 2002; Yoder and DeShon 2003) and are the basis of this element there is scant recognition or use of these principles in most state and tribal programs. The differential responses to categories of stressors (i.e., toxicity, habitat, nutrient enrichment, low D.O., etc.) are best exhibited at the most

refined levels of taxonomy (e.g., genus/species for macroinvertebrates) and observation (e.g., anomalies on fish) and by monitoring programs that result in the development of long term databases that include sufficiently detailed gradients of stressors and biological responses across sufficiently detailed spatial and temporal scales. This directly depends on the technical elements dealing with spatial sampling design and coverage, sampling methods, level of taxonomy, ecological attributes, biological endpoints and thresholds, and data management. Better development of diagnostic capacity is dependent on these other basic elements and it ultimately affects how effectively biological assessments are used in supporting water quality management. As such, the interdependency and order of the critical technical elements is amply demonstrated in the diagnostic capacity of a bioassessment program.

Discussion

The review and evaluation of the 14 state and single tribe bioassessment programs reveals that most are in various stages of development towards the type of program that is envisioned by U.S. EPA (2005). Achieving and maintaining a level 4 program is the desired status for states and tribes to have the capacity for employing tiered aquatic life uses (TALU) and numeric biocriteria in support of water quality management programs at relevant spatial and temporal scales. Given that standard, the state/tribal program review and critical elements process indicates that most have significant tasks to accomplish. It is encouraging to note that most have developmental programs underway that if sustained and completed should result in elevated program rigor in the next 5–10 years. Most are using their present bioassessment capabilities to accomplish doable tasks such as initial status assessments and making waterbody specific assessments. This illustrates that it is possible to accomplish a reasonable use of bioassessment as the program is being developed. Key to success, however, is in recognizing the inherent limitations of the current program and the types of decisions that may need to be deferred until the program rigor is improved.

The process and requirements for achieving a TALU-based approach to monitoring and assessment and WQS was recently described by U.S. EPA (2005). The critical technical elements review described herein is a critical part of that process by providing a baseline status assessment of a state or tribal bioassessment program that can be used to ascertain current status and support developmental efforts (Barbour and Yoder 2007). The feedback provided by the evaluation process and as memorialized in specific recommendations (e.g., Table 2) made via a technical memorandum can be used to develop a detailed plan for attaining a level 4 program. Once a sufficient technical infrastructure is in place the development of TALU-based WQS can proceed culminating in more refined tools, criteria, and broader support for all relevant water quality management programs. While the attainment of a level 4 bioassessment program does not guarantee similar improvements in WQS and other management programs, there is thus far a co-occurrence of attaining level 4 and also having numeric biocriteria and TALU codified in the state WQS.

Our experience gained during these evaluations indicates that while no two states are exactly alike, the challenges facing each in the improvement of their bioassessment programs are largely of a similar origin. While resource constraints in terms of personnel and funding are an in common theme in each state, we believe that how Clean Water Act management programs have been implemented and managed is an equally important and largely unrecognized factor. It consists of how individual programs have been independently developed by different EPA offices that collectively shape state priorities and which are almost exclusively measured by their administrative *outputs*. Ideally, the effectiveness or “success” of a water quality management program would be based on measures of the environmental *outcomes* that are produced by the aggregate of program activities i.e., *outputs*. In fact this is one of the most important and fundamental roles of ambient monitoring and assessment. However, successfully measuring environmental outcomes requires that sufficient indicators be developed that accurately represent not only the attributes and quality of the resource that is affected by these programs, but that also

link back to the stressors that these programs are attempting to manage. However, what we find most frequently is that the “success” of these management programs is wholly based on *outputs*, i.e., what each program produces as measured by the administrative activities within a particular management program. For example, within the permits program the programmatic outputs are measured by the number and types of permits, their rate of issuance and renewal, parameters included, backlogs eliminated, compliance rates, and similar measures. Using environmental outcomes would entail conducting ambient monitoring and assessment before and after the issuance of a permit to determine the extent of any changes in the affected water body. However, this is seldom accomplished in a formal sense as the permit program was designed to be measured by outputs, with outcomes being an assumed by-product of the outputs driven process. This has been well documented by a number of government reviews (e.g., GAO 2003) and is illustrated in Fig. 4 by the administrative outputs based column in which the operational results of each management program are used to judge its effectiveness. What we and others (U.S. EPA 1995a, b; NRC 2001; Karr and Yoder 2004) are proposing is that the improved monitoring and assessment and refined WQS of a TALU-based approach be more fully employed to support an environmental

Administrative Output vs. Resource Outcomes Based Management

	ADMINISTRATIVE OUTPUTS BASED	RESOURCE END OUTCOMES BASED
Goal:	Program Performance (Program execution)	Environmental Performance (Attain designated uses)
Measures:	Administrative Actions (Lists, Permits, Funding, Rules)	Indicator End-points (Biological , Chemical, Physical)
Results:	Improve Programs (Reduce backlogs, improve timeliness)	Programs are Tools to Improve the Environment (Admin. outputs evaluated by environmental end outcomes)

Fig. 4 The key goals, measures, and results of an administrative outputs based approach compared to an environmental outcomes-based approach to the management of state and tribal water quality management programs

outcomes based approach (Fig. 4, right column) in which the outputs of the management programs and their effectiveness are ultimately judged by the environmental outcomes that each produces. Under the administrative outputs approach, the character and quality of watersheds are simply the sum of the programs that affect them (Courtemanch 2007). This potentially results in a sub-optimization of watershed quality, especially if the programs are driven by incomplete information, inaccurate delineations of quality, failing to target limiting stressors, and the use of inaccurate standards and criteria, the latter of which can result from a reliance on general uses and univariate WQS. The opportunity to gain new insights from the feedback that a sufficiently rigorous assessment of the environmental indicator responses is oftentimes missed because the outputs are not dependent on such monitoring and assessment information. A TALU-based approach ensures that such information is available in time to affect the planning and permitting that is accomplished early on in the management process, thus providing a higher likelihood that positive outcomes will be produced. Hence, there is a natural sequence of information that drives the process towards an outcomes based approach.

A TALU-based approach delivers support at the front end of the management process by assuring that WQS are both appropriate and attainable prior to their use in developing abatement strategies and responses and it ensures that management efforts are targeting the places of greatest need and value. The essential and mostly missing ingredients are adequate monitoring and assessment that produces the types of data and waterbody specific assessments by which outcome assessments are made possible. The problem is that in most states neither the monitoring and assessment nor WQS programs are of a sufficient rigor to deliver these types of data, indicators, and criteria on a more widespread and day-to-day basis. However, examples of this do exist and are once again exemplified in the two level 4 state programs; these programs are described in more detail in U.S. EPA (2005).

The critical technical elements and the monitoring and assessment and WQS program evaluation process of which it is a part is a good start

on getting states and tribes to the point where their management programs can be developed, managed, and evaluated by true environmental outcomes and at the same scale at which management is being applied. At present this is an evolving area of practice and policy within EPA, but several states are taking steps to follow the guidance in U.S. EPA (2005) to achieve the better use of bioassessments through a TALU-based approach to water quality management.

References

- Barbour, M. T., & Yoder, C. O. (2007). *Critical technical elements of a bioassessment program* (p. 71). Washington, DC: U.S. EPA, Office of Water.
- Bonada, N., Prat, N., Resh, V. H., Statzner, B. (2006). Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. *Annual Review of Entomology*, 51, 495–523.
- Courtemanch, D. C. (2007). *A tale of two rivers. Use of biological information to more precisely define aquatic life uses: Tiered Aquatic Life Uses (TALU)*. Baltimore, MD: U.S. EPA Office of Water Workshop Presentations.
- General Accounting Office (GAO) (2003). *Major management challenges and program risks: Environmental protection agency. GAO-03-112, performance and accountability series*. Washington, D.C: U.S. General Accounting Office.
- Eagleson, K., Lenat, D., Ausley, L., & Winborne, F. (1990). Comparison of measured instream biological responses and responses predicted using the *Ceriodaphnia dubia* chronic toxicity test. *Environmental Toxicology and Chemistry*, 9(8), 1019–1028.
- Karr, J. R., & Yoder, C. O. (2004). Biological assessment and criteria improve TMDL planning and decision-making. *Journal of Environmental Engineering*, 130(6), 594–604.
- Houston, L., Barbour, M. T., Lenat, D., & Penrose, D. (2002). A multi-agency comparison of aquatic macroinvertebrate-based stream bioassessment methodologies. *Ecological Indicators*, 4, 279–292.
- Intergovernmental Task Force on Monitoring Water Quality (ITFM) (1992). *Ambient water quality monitoring in the United States: First year review, evaluation, and recommendations* (p. 51). Washington, D.C.: Inter-agency Advisory Committee on Water Data.
- Intergovernmental Task Force on Monitoring Water Quality (ITFM) (1995). *The strategy for improving water-quality monitoring in the United States: Final report of the intergovernmental task force on monitoring water quality*. Reston, Virginia: U.S. Geological Survey.
- Midwest Biodiversity Institute (MBI) (2004). Region V state bioassessment and ambient monitoring

- programs: Initial evaluation and review. Midwest Biodiversity Institute Tech. Rept. MBI/01-03-1. 36 pp. + appendices. http://www.epa.gov/region5/water/wqb/wqb_r5mon.htm.
- National Research Council (NRC) (2001). *Assessing the TMDL approach to water quality management*. ISBN 0-309-07579-3 (p. 109). Washington, D.C.: The National Academy Press.
- Riva-Murray, K., Bode, R. W., Phillips, P. J. (2002) Impact source determination with biomonitoring data in New York State: Concordance with environmental data. *Northeastern Naturalist*, 9(2), 127–162.
- Simon, T. P. (2003). *Biological response signatures: Patterns in biological integrity for assessment of freshwater aquatic assemblages* (p. 576). Boca Raton, FL: Lewis.
- U.S. EPA (1995a). *Environmental indicators of water quality in the United States*. EPA 841-R-96-002. Washington, D.C.: Office of Water, U.S. Environmental Protection Agency.
- U.S. EPA (1995b). *A conceptual framework to support development and use of environmental information in decision-making*. EPA 239 R-95-012. Washington, D.C.: Office of Policy, Planning, and Evaluation, U.S. Environmental Protection Agency.
- U.S. EPA (1996). *Biological criteria: Technical guidance for streams and rivers*. Authors: G. A. Gibson, M. T. Barbour, J. B. Stribling, J. Gerritsen, and J. R. Karr. EPA 822-B-96-001. Washington, D.C.: Office of Science and Technology, U.S. Environmental Protection Agency.
- U.S. EPA (1998). *Lake and reservoir bioassessment and biocriteria. Technical guidance document*. Authors: J. Gerritsen, R. Carlson, D.L. Charles, D. Dycus, C. Faulkner, G.R. Gibson, R.H. Kennedy, and S.A. Markowitz. EPA 841-B-98-007. Washington, D.C.: Office of Water, U.S. Environmental Protection Agency.
- U.S. EPA (1999). *Rapid bioassessment protocols for use in streams and Wadeable rivers: Periphyton, benthic macroinvertebrates and fish* (2nd ed.) Office of Water, Washington, D.C. Authors: M. T. Barbour, J. Gerritsen, B. D. Snyder, and J. B. Stribling. EPA 841-B-99-002. Washington, D.C.: Office of Water.
- U.S. EPA (2000). *Estuarine and coastal marine waters: Bioassessment and biocriteria technical guidance*. Authors: G. R. Gibson, M. L. Bowman, J. Gerritsen, and B. D. Snyder. EPA 822-B-00-024. Washington, D.C.: Office of Water.
- U.S. EPA (2002a). *Summary of biological assessment programs and biocriteria development for states, tribes, territories, and interstate commissions: Streams and Wadeable rivers*. EPA-822-R-02-048. Washington, D.C.: Office of Environmental Information and Office of Water, U.S. Environmental Protection Agency.
- U.S. EPA (2002b). *Consolidated assessment and listing methodology: Toward a compendium of best practices* (1st ed.). Washington, D.C.: Office of Wetlands, Oceans, and Watersheds, Office of Water, U.S. Environmental Protection Agency. <http://www.epa.gov/owow/monitoring/calm.html>.
- U.S. EPA (2005). *Use of biological information to better define designated aquatic life uses in state and tribal water quality standards: Tiered aquatic life uses*. EPA-822-R-05-001. Washington, DC: Office of Water, U.S. Environmental Protection Agency.
- Yoder, C. O. (1998). Important concepts and elements of an adequate state watershed monitoring and assessment program. In *Proceedings of the NWQMC national conference monitoring: Critical foundations to protecting our waters* (pp. 615–628). Washington, DC: U.S. Environmental Protection Agency.
- Yoder, C. O., & DeShon, J. E. (2003). Using biological response signatures within a framework of multiple indicators to assess and diagnose causes and sources of impairments to aquatic assemblages in selected Ohio rivers and streams. In T. P. Simon (Ed.), *Biological response signatures: Patterns in biological integrity for assessment of freshwater aquatic assemblages* (pp. 23–81). Boca Raton, FL: Lewis.
- Yoder, C. O., & Rankin, E. T. (1998). The role of biological indicators in a state water quality management process. *Journal of Environmental Monitoring and Assessment*, 51(1–2), 61–88.