ERDC TN-EMRRP-EBA-4
July 2010

Metric Development for Environmental Benefits Analysis

by S. Kyle McKay[1], Bruce A. Pruitt[2], Mark Harberg[3], Alan P. Covich[4], Melissa A. Kenney[5], and J. Craig Fischenich[1]

OVERVIEW: This technical note seeks to clarify concepts of metric development and application for U.S. Army Corps of Engineers (USACE) ecosystem restoration projects. Metrics are herein defined as measurable system properties that quantify the degree of achieving objectives. Metric lexicon, types, and policies are reviewed to adequately provide readers with background on the subject. Because metrics measure objective achievement, the importance of objectives and techniques for setting objectives are discussed. An iterative three-step metric development process is then proposed, consisting of: 1) metric selection based on a logical hierarchy of metric types, 2) metric evaluation using desirable properties of metrics, and 3) documentation and archival of metric development and application.

BACKGROUND: A key component to environmental benefits analysis is development of metrics to evaluate achievement of restoration objectives from both ecological and societal perspectives. Metrics are herein defined as measurable system properties that quantify the degree of objective achievement (Reichert et al. 2007). Moreover, a metric should measure the level of performance, raise awareness and understanding, measure progress toward programmatic goals and objectives, and support decision making. Generally speaking, metrics can be quantitative (e.g., length), semi-quantitative (e.g., big, bigger, biggest), non-quantitative (e.g., color), or nominal (e.g., yes or no); however, USACE policy requires restoration projects use metrics that are "expressed quantitatively" (ER 1105-2-100 (USACE 2000)). To clarify language, Table 1 presents terms often used synonymously with "metric" and subtle differences in how they are commonly applied.

[1] Research Civil Engineer, U.S. Army Engineer Research and Development Center (ERDC), Environmental Laboratory (EL).
[2] Principal Senior Scientist, Nutter & Associates.
[3] Environmental Resources Specialist, U.S. Army Corps of Engineers, Fort Worth District.
[4] Professor, Odum School of Ecology, University of Georgia.
[5] Postdoctoral Research Fellow, Department of Geography and Environmental Engineering, Johns Hopkins University and National Center for Earth-surface Dynamics.

Table 1. Terms commonly applied relative to metrics.

Term	Definitions Applied Herein	Common Application(s) of Term	Example
Metric	Measurable system properties that quantify objective achievement (Reichert et al. 2007)	Measuring objectives regardless of application	Objective: Minimize cost Metric: Project cost ($)
Attribute	Used synonymously with metric	Often used in reference to multi-criteria decision making	Objective: Minimize impact to wetland ecosystems Attributes: wetland acres, number of species
Indicator	Measurement indicative of something that is not measured but is of interest (Suter 2001)	Monitoring status or trends (e.g., EPA's Environmental Monitoring & Assessment Program, EMAP)	Objective: Improve stream water quality Indicator: invertebrate abundance and diversity
Performance Measure	A means of assessing programs, projects, products, activities, or services	Monitoring results of a specific action (e.g., post-project monitoring)	Objective: Increase oyster abundance Performance Measures: Size/number of oysters
Output	Raw measures generated by models or other procedures to capture the state of the system historically, currently, or in the future	Use ranges widely from the raw parameter values used in calculating indices (e.g., acres for habitat units) to the type of benefit of a project or program (e.g., NED outputs)	Objective: Minimize cost Outputs: capital cost, monitoring cost, operational costs
Index	A term combining a suite of parameters into a single value representing an overall process (Andreasen et al. 2001)	Usually applied in reference to combination of multiple metrics into one "index" either through a specified algorithm (e.g., IBI) or general mathematical means (e.g., sum, average)	Objective: Maximize biodiversity Indices: Index of biotic integrity (IBI), Shannon diversity index
Criterion	A decision factor against which objects can be measured	An evaluation tool for objectives, metrics, alternatives, plans, and projects	P&G criteria: Completeness, efficiency, effectiveness, acceptability

Within the Corps, metrics are applied to quantify environmental benefits at multiple scales, including:

- *Project alternative comparison* – How do ecosystem restoration planners compare environmental benefits from alternative restoration measures and plans (e.g., fish bypasses, ladders or lifts at a given location)?
- *Project performance assessment and success criteria development* – Following implementation, is the project successfully accomplishing objectives?
- *Adaptive management of outcomes* – Are the monitoring plan and metrics appropriate for identifying problems and adaptively managing deficiencies?
- *Regional programmatic assessment and portfolio management* – How do managers prioritize and manage multiple, smaller projects to achieve objectives in large-scale, system-wide efforts (e.g., Comprehensive Everglades Restoration Plan)?
- *National programmatic assessment and portfolio management* – How do managers prioritize and budget projects and track results to achieve objectives at a national scale?

USACE METRIC POLICY: This section lists USACE policy addressing topics that often arise in metric development for restoration projects.

Monetization. USACE policy has consistently stressed the challenges associated with monetization of restoration outcomes (e.g., "These measures…should be viewed on the basis of non-monetary outputs," ER-1165-2-501 (USACE 1999a); "environmental outputs considered in the evaluation process are typically not monetized", ER-1105-2-100 (USACE 2000)). However, monetized outcomes are only explicitly forbidden in §2-2 of the Planning Guidance Notebook (ER 1105-2-100 (USACE 2000)), which states, "Single purpose ecosystem restoration plans shall be formulated and evaluated in terms of their net contributions to increases in ecosystem value (NER outputs), expressed in non-monetary units." Exceptions to this clearly stated policy do nevertheless exist. For single-purpose restoration projects, "Monetary gains (e.g., incidental recreation or flood damage reduction) and losses (e.g., flood damage reduction or hydropower) associated with the project shall also be identified" (ER-1105-2-100 (USACE 2000)). Furthermore, "restoration projects which accomplish water quality improvement, habitat restoration, recreation, flood damage reduction, etc., are most likely to possess both NED and environmental quality (EQ) benefits" (EP 1165-2-502 (USACE 1999b)), which highlights multi-purpose projects as sources of both monetary and non-monetary benefits.

Significance. "Because of the challenge of dealing with non-monetized benefits, the concept of significance of outputs plays an important role in ecosystem restoration evaluation" (ER 1105-2-100 (USACE 2000)). ER 1105-2-100 (USACE 2000) (§E-37) provides a framework for determining significance based on institutional, public, and technical importance. Institutional recognition is often found in federal, state, or local policy or law and may include concerns associated with the Endangered Species Act, state natural heritage programs, or zoning ordinances. Public significance is often evidenced by overt interest in a resource (e.g., attendance at public meetings), financial contributions, volunteering, participation through a local, state, regional, or international interest group (e.g., Trout Unlimited, The Nature Conservancy, or a Native American tribe), or other means of demonstrating an interest in a resource. Technical significance is based on scientific knowledge of the ecosystem and its status relative to threats and protection. ER 1105-2-100 (USACE 2000) provides six criteria for assessing the technical significance of a resource as: scarcity, representativeness, status and trends, connectivity, critical habitat, and biodiversity.

Metric Type. Although "ecosystem restoration outputs must be clearly identified and quantified in appropriate units" (ER 1105-2-100 (USACE 2000)), selection of outputs, units, and techniques for quantification is extremely challenging because many types of metrics may be applied to a given problem. In general, policy does not emphasize a preferred metric or metric type, but instead that "all relevant ecosystem components need to be described and assessed" (EP 1165-2-502 (USACE 1999b)). The objective of the Corps Ecosystem Restoration program is "to restore degraded ecosystem structure, function, and dynamic processes to a less degraded, more natural condition" (ER 1105-2-100 (USACE 2000)). Thus, metrics that capture aspects of structure, function, and process are obvious candidates for use in analyses. "Ecosystem structure is the state and spatial distribution of material forms within the ecosystem at a specified time. It includes both microscopic and macroscopic material components in diverse living and non-living assemblages. Ecosystem functions are dynamic processes that can be characterized by rate and direction of change in material and energy flows through time and space. Ecosystem functions redistribute components of structure through abiotic (non-living) and biotic (living) processes" (EP 1165-2-502 (USACE 1999b)).

Traditionally, the quantity and quality of habitat have been used jointly in the form of habitat units to measure structure provided by ecosystem restoration projects. Although powerful tools for ecosystem management and restoration, Habitat Suitability Index (HSI) models are not regarded as universal indicators of ecosystem condition, may inadequately represent ecosystem function and process, and should be selected and evaluated carefully to ensure local applicability (Hubert and Rahel 1989; Leftwitch et al. 1997; Vadas and Orth 2001). Furthermore, policy also specifies that single species approaches should be discouraged as they narrow the breadth of objectives and benefits provided, and community-based approaches are preferred (ER 1105-2-100 (USACE 2000)). Guidance has traditionally alluded to habitat-based approaches (e.g., "habitat-based evaluation methodologies, supplemented with production, user-day, population census, and/or other appropriate information, shall be used to the extent possible to describe and evaluate ecological resources and impacts associated with alternative plans" ER 1105-2-100, §C-3 (USACE 2000)), but by no means eliminates other approaches which may include other structural metrics (e.g., spatial characteristics, community structure, or water quality conditions) or be more function- or process-based (e.g., bank retreat, organic matter breakdown). Although a thorough discussion of the merits of structural and functional metrics is beyond the scope of this paper, function- and process-based metrics of ecological condition are often highlighted as underutilized measurements of the state of an ecosystem (Young et al. 2008).

"The concepts of ecosystem function and structure are closely intertwined, and both include abiotic and biotic elements and processes" (EP 1165-2-502 (USACE 1999b)). Thus, a combination of biotic and abiotic metrics measuring structure, function, and process will likely lead to the most comprehensive accounting of environmental benefits. Although biological outcomes are often the objective of restoration projects, abiotic metrics that are clearly linked to these outcomes are also acceptable.

Forecasting. The Planning Guidance Notebook provides a four-step evaluation procedure for forecasting restoration project benefits: "(1) forecast the most likely with-project conditions expected under each alternative; (2) compare each with-project condition to the without-project condition and document differences between the two; (3) characterize the beneficial and adverse effects by magnitude, location, timing, and duration; and (4) qualify plans for further consideration" (ER 1105-2-100 (USACE 2000)). Thus, the ability to forecast a metric for futures with and without the project is critical to project planning and metric selection. Unlike NED projects, "ecosystem restoration outputs are not discounted, but should be computed on an average annual basis, taking into consideration that the outputs achieved are likely to vary over time" (ER 1105-2-100 (USACE 2000)). Additionally, model selection, application, and interpretation should follow sound science and apply appropriate levels of professional judgment to forecast outcomes. However, the state of the science in ecosystem modeling is relatively uncertain and policy stipulates, "When identifying the NER plan the associated risk and uncertainty of achieving the proposed level of outputs must be considered" (ER 1105-2-100 (USACE 2000)). Metric uncertainty should be considered relative to acceptable risks of a decision.

OBJECTIVE SETTING IN ECOSYSTEM RESTORATION: "A clear definition of objectives and constraints is essential to the success of the planning process" (ER 1105-2-100 (USACE 2000)) and influences almost all aspects of a project, including metric development, alternative formulation and evaluation, and plan comparison and selection (Yoe and Orth 1996). Metrics provide a direct translation between an objective and measurement of that objective. Therefore, objectives are critical for development of metrics (Keeney and Gregory 2005). Though this may seem like common sense, too often environmental managers fail to adequately specify their objectives (Gregory and Keeney 2002; Tear et al. 2005; Wohl et al. 2005); thus, there is no clear roadmap to help chart the course or determine whether the direction chosen was successful or not. For ecosystem restoration, objective setting is particularly important because there are likely many objectives and constraints associated with projects, ranging from maximizing biodiversity to maintaining cultural integrity to minimizing costs, and metrics for these potentially competing objectives are needed to trade off benefits and compare alternatives.

Minimally, a good list of objectives must be complete and clear (Gregory and Keeney 2002). A complete list includes all objectives, primary and secondary, relevant for making a decision. Clear objectives state exactly what is meant and the direction of preference to minimize multiple interpretations. Developing complete and clear objectives is not a trivial task and requires critical thinking and wide discussion. Because of the importance of objectives to the planning process and the wide array of values that objectives might address, developing objectives collaboratively in an environment inclusive of many values and perspectives is often vital to creating a shared vision of project success (Yoe and Orth 1996; Tear et al. 2005). Gregory and Keeney (2002) provide the following objective setting framework to facilitate this process. Additional guidance on setting restoration-specific objectives can be found in reviews by Slocombe (1998), Palmer et al. (2005), Tear et al. (2005), Fischenich (2006), Reichert et al. (2007), Beechie et al. (2008), and Covich et al. (in preparation).

Step 1: Write down the concerns you want to address. Objective setting begins by listing all of the potential elements that may influence the decision. In this process, allow ideas to flow freely because limiting oneself to feasible, previously stated, or eloquently worded objectives may hinder creativity or comprehensiveness. In group decision-making, it is often helpful for each individual to undertake this exercise and then combine ideas in a later step (Gregory and Keeney 2002). The following statements are often applied to spur creativity in this process: list problems and opportunities associated with this project, examine the study authority, consult available watershed management plans, compose a wish list, think about the best and worst possible outcomes, consider the pros and cons of good and bad alternatives, think about how to explain a chosen alternative to someone else, contemplate what others in similar situations have considered when making their decisions, and reflect on the ultimate goals of pursuing this effort (Yoe and Orth 1996; Keeney 2007).

Step 2: Convert general concerns into succinct objectives. The comprehensive listing from the previous step must now be converted to a set of pointed, concise objective statements. Objective statements are most clearly communicated through a verb-object format such as minimize cost or maximize habitat for species X, but care should be taken to avoid being too vague or too specific in this process. Specificity associated with the direction, magnitude, location, timing, and duration of an objective often provides benchmarks of success (Yoe and

Orth 1996; USACE 2000; Tear et al. 2005). For instance, the objective to "maximize floodplain habitat for species X from river miles 10 to 20" captures key elements of direction of preference, project boundaries, and target species. However, over-specificity such as "restore 100 percent of floodplain habitat for species X from river miles 10 to 20" may be inappropriate because this objective may be impossible to achieve given land ownership constraints, clear definition of what constitutes the floodplain, or uncertainty regarding historical or future home ranges of the species. Thus, caution should be used in application of absolute targets or benchmarks in objectives (e.g., 100 percent or 2000 acres) and acknowledgement of uncertainties and natural dynamism should be explicitly incorporated. Furthermore, objective setting should be transparent, open to peer review, scientifically credible (e.g., capturing known ecological thresholds), well-documented, and mindful of relevant uncertainties (Tear et al. 2005).

Step 3: Structure objectives. Objectives may also be clarified by considering their underlying motivation and structuring them accordingly. This process of separating ends (fundamental goals) from means (waypoints to achieving goals) is often needed to adequately summarize a given problem. For instance, "maximize fish passage at hydrologic barriers" is a means to the end "improve aquatic ecosystem health." One technique for distinguishing means from ends is to simply ask "Why?" until the fundamental objective (or end) is identified (Gregory and Keeney 2002). After fundamental objectives have been identified, it is often beneficial to structure objectives as a nested set of fundamental and means objectives (Box 1). In Corps planning, the highest order objectives are transferable between projects and are dictated by policy. For instance, the Planning Guidance Notebook (ER 1105-2-100 (USACE 2000)) specifies "The objective of ecosystem restoration is to restore degraded ecosystem structure, function, and dynamic processes to a less degraded, more natural condition." Moreover, structuring of objectives provides an opportunity to separate objectives used in the decision (e.g., maximize habitat) and procedural or process objectives for how the decision will be made (e.g., engage public, develop alternatives collaboratively; Slocombe 1998; Keeney 2007).

Step 4: Clarify what is meant by each objective. Assess, refine, collaborate, and iterate. Once objectives have been developed it is often beneficial to impartially examine the objective set. Desirable objective sets are: complete, nonredundant, concise, specific, understandable, flexible, measurable, attainable, congruent, and acceptable (Yoe and Orth 1996; USACE 2000; Keeney 2007). Many of these properties may be in conflict with each other, thus the goal is not to meet all of these, but instead to achieve balance among them. Furthermore, objective setting does not stop once the first set is developed, and objectives may need to be refined based on interaction with stakeholders, sponsors, and the public, increased familiarity with the project, consideration of Corps authorities and missions, or newly available scientific knowledge (Tear et al. 2005). Though many objectives are required to fully address ecosystem restoration projects, depending on the project authority, funding, timing, and complexity, not all objectives may be associated with Corps project decision-making. Careful thought should be considered in developing the list of objectives that the federal portion of the project will address and what will be left to others. For instance, if a watershed plan identifies improving aquatic, riparian, and upland habitat connectivity as key components to recovering a given species, then the Corps project may emphasize the former ecosystems while deferring to other local or federal partners for upland restoration efforts.

Multi-purpose stream projects often identify similar project goals (or fundamental objectives), such as enhancing ecological health, protecting infrastructure, improving water quality, and minimizing costs. However, concrete objectives are specified based on preferences and circumstances surrounding a given project (e.g., funding agency mission; location, time and funding constraints; level of public involvement). For instance, if a hypothetical town is planning on restoring a 2-mile reach of river with excessively eroding banks, an initial list of objectives may include reducing erosion, enhancing ecological health, and protecting infrastructure. After reviewing local issues and consulting an existing watershed plan, objectives are augmented to include reducing fine sediment loading (a problem for the town's water treatment plant) and providing recreation (to support local kayaking tourism). Additionally, the team may want to improve the decision-making process by engaging the public. These large-scale concerns represent the fundamental objectives of the team, but they must be redefined in greater detail for use in alternative comparison. The team can convert these concerns into succinct objectives, structure them, and eliminate irrelevant or redundant objectives to summarize the problem in the following objectives: four fundamental, nine component, and two process (with associated metrics in parentheses). A number of these objectives address multiple values (e.g., protection of citizens addresses public safety concerns as well as flood damage cost), some are measured by the same metrics (e.g., fine sediment loading and riparian habitat), and some are measured by multiple metrics (e.g., fine sediment loading). These do not represent every aspect of alternative selection for stream restoration projects, but instead those aspects relevant to *this* multi-purpose stream project.

- Provide socio-economic benefit.
  - Protect citizens (# of homes in 100-yr floodplain)
  - Protect public works (# of impacted bridges, pipelines, and roads in reach)
  - Increase recreational boating (boater survey addressing kayaking terrain offered by plans)
- Minimize costs.
  - Maximize cost efficiency (linear ft / $)
  - Minimize capital cost (total planning, design, and construction cost in $)
  - Minimize O&M cost (average annual O&M cost in $)
- Protect water quality.
  - Reduce fine sediment loading (qualitative bank erosion hazard index score, Rosgen 2001; length of excessively eroding bank in feet; acres of riparian zone)
- Enhance biological integrity.
  - Enhance coldwater fisheries (community-based habitat units[1])
  - Increase riparian habitat quantity (acres of riparian zone)
- Process Objectives
  - Minimize disturbance to intact habitat (length of reach impacted by construction)
  - Engage public in decision making process (# of people participating in public meetings)

[1] Box 2 provides the logic for selection of the fisheries metric.

Box 1. Objective setting hierarchy for an example multi-purpose stream project.

METRIC DEVELOPMENT: Following development of a complete and clear set of objectives, metrics may be identified to evaluate those objectives and inform decisions. A single metric measuring one objective may be sufficient to distinguish between alternatives, or a set of metrics measuring the full suite of objectives may be required. Considering the variability and complexity of ecosystems, any set of metrics should be considered incomplete and at best only representative of a myriad of decision factors related to project prioritization and funding. The following sections outline an iterative three-step process for metric set development that seeks to provide structure to this problem. This framework (Figure 1) extends the metric selection and evaluation process proposed by Keeney and Gregory (2005) by including documentation and organized archival, significant shortcomings often encountered in restoration projects.

As was the case with objectives, metrics should be developed iteratively with all parties pertinent to the decision. Development of metrics by one party or discipline may skew metrics to a particular field of study or value system. For instance, given the objective “increase river-floodplain connectivity,” a metric may be proposed by a civil engineer as a percentage of time stage is greater than 12 ft, by a biogeochemist as the ratio of river to floodplain nitrate uptake, or by an aquatic ecologist as acreage of floodplain spawning habitat provided, all of which may be valid metrics. Metric development often requires bridging of knowledge from multiple fields of study and value systems; thus, a facilitated discussion amongst experts, professionals, and stakeholders may be required to reach consensus on metrics.

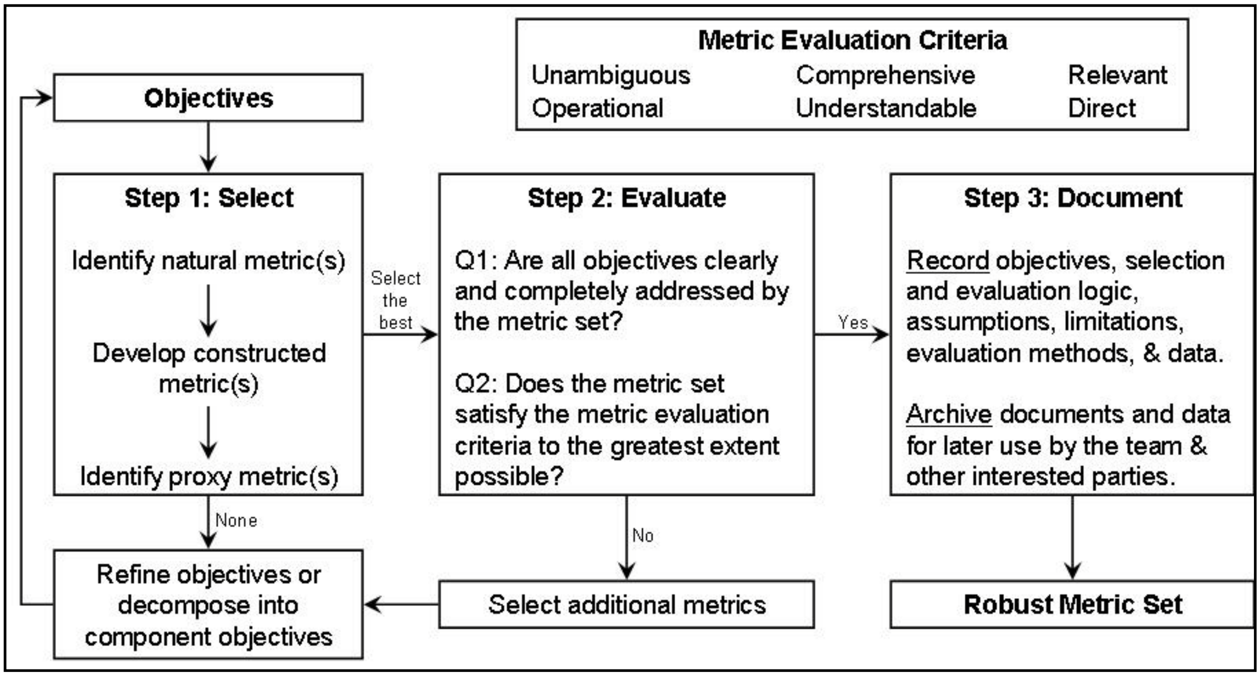


Figure 1. Metric development process (adapted from Keeney and Gregory 2005). Please refer to the text as well as Tables 2 and 3 for additional information on metric selection, evaluation, and documentation.

Step 1: Metric Selection. Metric selection must be based on and mapped to specific project objectives. As such, no complete list of universal metrics can exist; however, one need not reinvent the wheel for each project. Exhaustive lists of metrics for various ecosystems exist and provide excellent starting points for metric selection (Orians et al. 2000; Somerville and Pruitt 2004; Thayer et al. 2003, 2005; Clark et al. 2008; Faber-Langendoen et al. 2008; USEPA 2009; U.S. Forest Service (USFS) 2009). Past projects with similar (albeit not identical) objectives are also good sources of information. Subject matter and local experts may also provide direction on state-of-the-science or state-of-the-practice metrics. When selecting metrics, it is often beneficial to use one, or all, of these resources to identify many metrics measuring the same objective (example above) and later remove metrics based on the evaluation conducted below (Box 2).

Keeney and Gregory (2005) classify metrics into three categories (natural, constructed, and proxy) based on their ability to measure an objective directly and suggest that the degree to which a metric measures an objective should be considered in metric selection. Natural metrics are those that directly and unequivocally measure an objective. Constructed metrics are developed to directly measure an objective when no obvious natural metric exists. Proxy metrics are indirect measures of a given variable or process. Table 2 summarizes and gives examples of these metric types. In accordance with Keeney and Gregory (2005), the authors suggest a hierarchy of metric types with preference given first to natural, then to constructed, and finally to proxy metrics. For instance, if reducing stream temperature is an objective of a riparian restoration project, how does one select metrics? First, stream temperature in degrees Celsius is the obvious, natural metric of this objective; however, resource or modeling limitations may prohibit detailed forecasting of this metric. Second, a constructed scale of temperature could be developed and scored by a panel of subject matter experts (e.g., Poor = >40°C, Fair = 30-40°C, Good = 20-30°C, Excellent = <20°C). Finally, if resources or time prohibit use of a panel, the team may apply forested streamside acreage from aerial photos as a proxy for stream temperature that assumes that temperature goes down with increased acreage. Also of note is that a single metric may represent more than one type. For instance, quantity and quality of habitat is a proxy for a given population's health or viability, which is a function of environmental tolerances as well as available food, mates, genetic diversity, and other factors. Habitat Suitability Index (HSI) curves seek to capture as many of these parameters as practicable through a constructed scale representing habitat quality. Thus, habitat units are both constructed and proxy metrics.

This paper presents a framework for metric development, but application of this framework is perhaps most clearly shown through demonstration. This example discusses and documents metric development for a single objective from Box 1, enhancing coldwater fisheries.

Selection: Metrics were selected and evaluated by a panel of subject matter experts from the project development team, resource agencies, academia, and consulting firms (A list of panel qualifications and expertise is often beneficial to confirm that they represent the "best" professional judgment available.). Quantifying fisheries enhancement is challenging due to contributing factors varying from hydraulic conditions to sufficient water quality to existing biotic community structure. Three metrics were proposed by the panel for consideration: 1) population size of three key species (with varying life histories which represent the community at large) as predicted by a calibrated population model, 2) habitat units as predicted by a community habitat suitability index model, and 3) a qualitative habitat score based on the Rapid Bioassessment Protocols for wadeable streams as modified for local conditions (Barbour et al. 1999; hereafter referred to as metrics 1, 2, & 3, respectively). Seasonal biological survey data spanning five years are available for calibrating or modifying models as needed.

Evaluation: Although each metric could be used to quantitatively assess the relative merits of alternatives, metrics were evaluated in light of the time and resource constraints of this project.

- Metric 1 provides a relevant, direct, understandable, and somewhat comprehensive measure of the state of coldwater fisheries; however, time required to develop the model for each species rendered it operationally infeasible. Additionally, the panel had concerns about ambiguity of the metric due to a wide range of naturally varying conditions (e.g. discharge, water temperature) inducing significant uncertainties in forecasted populations.
- Metric 2 is an indirect and non-comprehensive measure of coldwater fisheries in that it assesses habitat quantity and quality rather than the biological community of interest. This metric does, however, provide a measure of fish community health relevant to the management levers applied, with unambiguous interpretation, and in a manner that is understandable to project stakeholders. Additionally, the metric may be operationally applied within time and resource constraints.
- Metric 3 is highly operational due to the limited data inputs and rapidly applied methods; however, its heavy reliance on professional judgment and indirect assessment of habitat rather than the biological community raised concerns with the expert panel.
- Based on this discussion, metric 2 was recommended for use in this analysis. This metric does not represent the best measure of coldwater fish population integrity, but given its operationality and close relationship to management levers, a sacrifice of comprehensiveness and directness was deemed satisfactory for this report.

Box 2. Documenting metric development.

Table 2. Types of metrics.

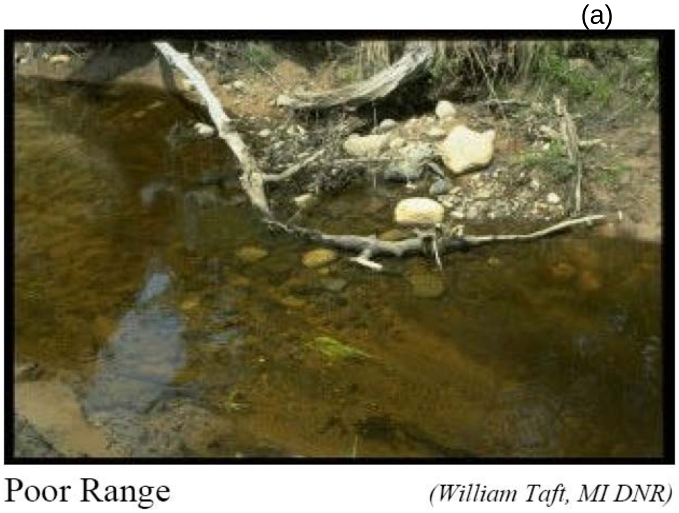
Metric Type	Definition and Description	Examples (O=Objective, M=Metric)
Natural	Metric that directly measures an objective, is in general use, and represents the common unit of measure of a given variable. This type of metric is often straightforward and obviously corresponds to the objective.	O: Minimize cost M: Project cost ($) O: Increase California Condor population size M: Condor breeding pairs
Constructed	Metric developed to directly measure an objective when no natural metric exists. Keeney and Gregory (2005) identify five types of constructed metrics: defined levels (e.g., stakeholder surveys), quality-quantity scales (e.g., HSI models), value models incorporating two natural metrics (e.g., toxicity as the product of dose and exposure), weighted scales (e.g., Grade Point Average), and pictures (e.g., pictorial guides for selecting Manning's *n*). Regardless of the type of constructed metric, experienced team members familiar with the project, alternatives and resulting consequences will construct more robust metrics for measuring an objective.	O: Minimize substrate embeddedness M: Embeddedness scale scored by experts (Figure 2a & 2b) O: Maximize slider turtle habitat M: Slider turtle habitat units using acreage for quantity and predetermined HSI curves for quality (Figure 2c)
Proxy	Metrics indirectly measuring a given variable or process that are often used because of relative ease of measurement or understanding. In aquatic ecosystem restoration, perhaps the most commonly applied proxy metrics are indicator species that are assumed to integrate multiple aspects of ecosystem integrity (e.g. water quality, water quantity, and community structure).	O: Improve stream water quality. M: Macroinvertebrate community structure (e.g., EPT ratio) O: Increase duck habitat M: Quantity of ducks * duration of habitat utilization ("Duck-days")

Step 2: Metric Evaluation. Once a metric or metric set has been selected, metrics may be evaluated based on whether the metric set adequately addresses project objectives and meets desirable qualities of metrics. Many criteria may be used for metric evaluation (Poole et al. 1997; Harwell et al. 1999; Dale and Beyeler 2001; Orians et al. 2000; Andreasen et al. 2001; Kurtz et al. 2001; Yoe 2002; Keeney and Gregory 2005); however, the following six properties are proposed as fundamental qualities of a "good" metric set inclusive of all others: relevant, unambiguous, comprehensive, direct, operational, and understandable. Table 3 summarizes each metric evaluation criterion and related topics identified in the literature.

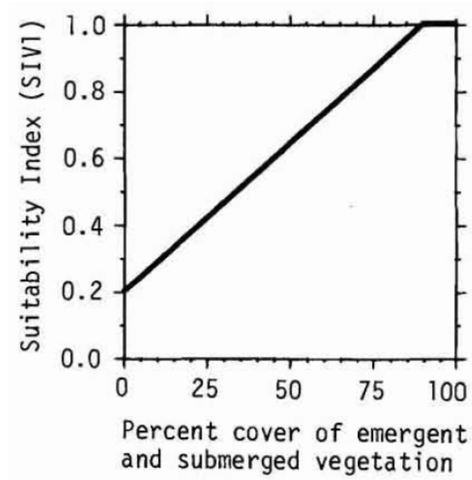
This is not to imply that each metric will meet all six evaluation criteria, but that every metric set should address each criterion to the greatest extent practicable. Using these criteria, the restoration team should evaluate the selected metric set, which often results in trading off one criterion for another. In restoration practice, metrics often sacrifice comprehensiveness and are based heavily on operational criteria, and although operationality is a critical concern, it must be emphasized that metrics should not be chosen on this factor alone (Keeney and Gregory 2005).

Habitat Parameter	Condition Category			
	Optimal	Suboptimal	Marginal	Poor
2.a Embeddedness (high gradient)	Gravel, cobble, and boulder particles are 0-25% surrounded by fine sediment. Layering of cobble provides diversity of niche space.	Gravel, cobble, and boulder particles are 25-50% surrounded by fine sediment.	Gravel, cobble, and boulder particles are 50-75% surrounded by fine sediment.	Gravel, cobble, and boulder particles are more than 75% surrounded by fine sediment.
SCORE	20 19 18 17 16	15 14 13 12 11	10 9 8 7 6	5 4 3 2 1 0

(a)



(b)



(c)

Figure 2. Sample constructed metrics: (a) defined-level scale, (b) pictorial scale for stream substrate embeddedness (Barbour et al. 1999), and (c) suitability index for slider turtle (Morreale and Gibbons 1986).

Step 3: Metric Documentation. The final step in metric development is an obvious, but often overlooked issue, documentation and archival. This is because in restoration planning, metrics are often developed "on the fly" and may not be documented well, if at all, particularly if project monitoring and follow-up is of low priority. Thus, with time, it becomes difficult to understand why a metric was used, what objective(s) it addressed, and how successful it was at measuring those objectives. Appropriate documentation of metric development and application should minimally address the following points (which may also be applied in review of metrics; See Box 2 for sample documentation of a single metric):

- Complete and clear statement of objectives
- Metric development process applied (e.g., committee, literature review)
- Logic for metric selection and which objective(s) it measures
- Literature, expert, or past-project support for use of the metric
- Assumptions and limitations associated with metrics
- Application of professional judgment in metric development or assessment
- Ability of the metric set to address the metric evaluation criteria (Table 3)
- Techniques for assessing and forecasting the metric (e.g., numerical models, expert judgment, monitoring plans, data collection protocols)
- Any review the metric set has undergone (e.g., interagency project team)

- In long-term or high-cost efforts, it may be valuable to document initial metrics as well as any changes in metrics because of new information or modification of the objectives.

Table 3. Metric evaluation criteria.

Criteria	Description	Related Properties
Relevant	Relevant metrics account for specified objectives, scientific knowledge base and priorities of decision-makers (e.g., resource significance, project authority) at appropriate spatial and temporal scales. Moreover, if metrics are scientifically irrelevant, non-repeatable or non-verifiable, then the scientific basis for decisions may be compromised.	Scientifically valid, differentiates between alternatives, appropriate spatio-temporal scale and resolution
Unambiguous	Unambiguous metrics clearly measure consequences of alternatives and are not obscured by direction, magnitude, scale, threshold, or outcome or measurement uncertainty.	Clear direction of preference, clear interpretation, transparent, repeatable, low signal-to-noise ratio, minimal natural uncertainty
Comprehensive	Comprehensive metric sets address the suite of objectives and cover the potential range of consequences. In terms of implementation, comprehensiveness is often captured through multiple metrics and well-designed monitoring and forecasting programs.	Completeness, wholeness
Direct	Direct metrics address objectives as clearly as possible. This underscores the importance of measuring what can be controlled by a given action since restoration is often reliant upon many variables beyond the control of the restoration team. Although multimetric indices appear to integrate numerous metrics or objectives into a single score, great care should be taken in their development to ensure they remain direct by maintaining sensitivity to outcomes, clarity, and conceptual meaning.	Scientifically accurate and precise, rely on well-tested theory, clarity in data quality objectives, respond to management actions, sensitivity to outcomes, clearly mapped to objectives
Operational	Operationality is critical because if a metric cannot be assessed, forecasted, or monitored within budgetary, time, or labor constraints, then it cannot feasibly inform decisions.	Cost-effective, feasible, sufficient data for analyses, manageable data storage and processing, efficiency
Understandable	Understandable metrics clearly communicate decisions to those interested in the analysis.	Clarity, communicable to scientists, practitioners, decision makers, and stakeholders, appropriate scale/units for communication

An appropriate archival plan and infrastructure is also important so that information may be referenced and applied to improve objective setting, metric development, and decision-making in future projects. This is partially accomplished via peer-reviewed journal papers (articles); however, in practice, results are rarely published beyond “grey literature” project reports or government documents, which may or may not complete post-project evaluations. Efforts should be made to encourage documentation of metric development in these reports and the archival of said reports in common repositories such as corporate offices, libraries, project websites, or online databases. Improved documentation and better archival will result in better communication of successes and failures associated with metric development and will help avoid repeating mistakes.

OBSTACLES TO METRIC DEVELOPMENT: This technical note presents a framework for metric development that may be applied to ecosystem management and restoration from project to programmatic scales. By applying this metric development process, metrics for measuring environmental benefits may be identified; however, it must be emphasized that the following obstacles can arise in metric development for environmental benefits analysis:

- Objectives and metrics should be considered throughout the analysis of any decision and may require adaptation as a project progresses; however, setting objectives and developing metrics should be particularly emphasized at the beginning of the decision process so that alternatives and forecasting can be developed around them.

- When selecting objectives and metrics, make sure all components of an ecosystem have been adequately considered. Developing a conceptual model of an ecosystem can be useful for identifying these key processes and drivers (Fischenich 2008). Covich et al. (in preparation) also provide direction on which ecosystem components to consider in restoration.

- Although objectives for ecosystem restoration often center on biological endpoints (e.g., re-establish a commercially viable population of oysters), population viability analyses are challenging for even the most well-studied taxa. Therefore, restoration metrics are often centered on abiotic elements supporting a given biological target (i.e., habitat restoration). The current state of the science requires the use of these abiotic metrics, but a restoration team should carefully consider the linkage between physical processes and biological outcomes as well as implicit assumptions in analyses (e.g., sufficient source populations) as abiotic measures are often used as proxies for biological targets.

- Special attention should be given to the spatial and temporal scale of metrics. Metrics may respond with different restoration trajectories or thresholds. External factors outside the control of the restoration team may also be important, such as long-term shifts in baseline conditions (e.g., species invasion or sea level rise) or natural levels of stochastic ecosystem response (e.g., periodic drought). Furthermore, environmental benefits may be calculated and reported at multiple scales, and there may be a need to account for cumulative benefits and impacts outside of the project area. Because of these challenges, multiple metrics may be required that act across multiple spatial and temporal scales to adequately inform decision-making (e.g., fish passage improvement at a given location and the effect of that improvement for the larger watershed).

- Objectives and metrics often have varying levels of importance to the overall execution of a project. As such, determining the relative importance of objectives and metrics to decision-making may be critical for identifying the appropriate alternative. For instance, the recovery of a threatened species may carry more significance in project decision-making than the maintenance of a currently ubiquitous taxa. The incorporation of scientific as well as societal importance, significance, value, or utility into metric development is difficult, and techniques such as Multi-Criteria Decision Analysis (MCDA) may assist teams in overcoming these challenges.

- If multiple metrics for a project are identified, metrics may be traded off under different alternatives (e.g., habitat quantity for species A may be sacrificed for habitat for species B). Trade-off analysis is a useful tool for understanding how benefits of a project change under different alternatives, but Yoe (2002) recommends that no more than six or seven metrics be used in a trade-off analysis due to the difficulty in comparing multiple sources of information simultaneously. Comparison and combination of dissimilar metrics is a challenging issue that may be confronted through a number of techniques that include but are not limited to: Multi-Criteria Decision Analysis, normalization of metrics to a similar scale (e.g., 0 to 1), or conversion of all metrics to a similar unit (e.g., dollars, energy).

- Given the change in resolution of objectives from the project to the programmatic scale, different metrics may be needed to report benefits at these varying scales. For instance, different metrics may be required to differentiate between alternatives for a project within the Louisiana Coastal Area than are used to report regional benefits of that project.

- Last, but not least, although selection of appropriate metrics is the end product of the framework presented, "the foundation for any decision is a clear statement of objectives" (Keeney and Gregory 2005). Thus, metrics cannot be considered without first addressing objectives, so the only way to develop good metrics is to first develop good objectives!

ADDITIONAL INFORMATION: Research presented in this technical note was developed under the Environmental Benefits Analysis (EBA) Research Program. The USACE Proponent for the EBA Program is Rennie Sherman and the Technical Director is Dr. Al Cofrancesco. Technical reviews by Dr. Andy Casper, Sarah Miller, Antisa Webb (ERDC Environmental Laboratory), and Dr. Dick Cole (USACE Institute for Water Resources) greatly improved prior drafts of this document.

For additional information, contact the author, S. Kyle McKay (601-415-7160, *Kyle.McKay@usace.army.mil*), or the manager of the Environmental Benefits Analysis Research Program, Glenn Rhett (601-634-3717, *Glenn.G.Rhett@usace.army.mil*). This technical note should be cited as follows:

> McKay, S. K., B. A. Pruitt, M. Harberg, A. P. Covich, M. A. Kenney, and J. C. Fischenich. 2010. *Metric development for environmental benefits analysis*. EBA Technical Notes Collection. ERDC TN-EMRRP-EBA-4. Vicksburg, MS: U.S. Army Engineer Research and Development Center. *http://cw-environment.usace.army.mil/eba/*

REFERENCES

Andreasen, J. K., R. V. O'Neil, R. Noss, and N. C. Slosser. 2001. Considerations for the development of a terrestrial index of ecological integrity. In *Ecological Indicators*. Elsevier. 1(1): 21-35.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. *Rapid bioassessment protocols for use in streams and wadeable rivers, Second edition*. EPA 841-B-99-002. Washington, DC: USEPA.

Beechie, T., G. Pess, P. Roni, and G. Giannico. 2008. Setting river restoration priorities: A review of approaches and a general protocol for identifying and prioritizing actions. *North American Journal of Fisheries Management*. 28 (4): 891-905.

Clark, W. C., and others. 2008. *The state of the Nation's ecosystems, 2008: Measuring the lands, waters, and living resources of the United States*. The H. John Heinz Center for Science, Economics, and the Environment. Washington, DC: Island Press.

Covich, A., S. K. McKay, M. A. Kenney, B. A. Pruitt, M. Harberg, J. C. Fischenich, and J. J. Boland. Enhancing aquatic ecosystem integrity with effective objectives and metrics for restoration. In preparation, *Ecological Restoration*.

Dale, V. H., and S. C. Beyeler. 2001. Challenges in the development and use of ecological indicators. In *Ecological Indicators*, 1:3-10. Elsevier.

Faber-Langendoen, D., G. Kundray, C. Nordman, L. Sneddon, L. Vance, E. Byers, J. Rocchio, S. Gawler, G. Kittel, S. Menard, P. Comer, E. Muldavin, M. Schafale, T. Foti, C. Josse, and J. Christy. 2008. *Ecological performance standards for wetland mitigation: An approach based on ecological integrity assessments*. Arlington, VA: NatureServe.

Fischenich, J. C. 2006. *Functional objectives for stream restoration*. ERDC TN-EMRRP-SR-52. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Fischenich, J. C. 2008. *The application of conceptual models to ecosystem restoration*. ERDC TN-EBA-TN-08-1. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Gregory, R. S., and R. L. Keeney. 2002. Making smarter environmental management decisions. *Journal of the American Water Resources Association* 38 (6).

Harwell, M. A., V. Myers, T. Young, A. Bartuska, N. Gassman, J. H. Gentile, C. C. Harwell, S. Appelbaum, J. Barko, B. Causey, C. Johnson, A. McLean, R. Smola, P. Templet, and S. Tosini. 1999. A framework for an ecosystem integrity report card. *Bioscience* 49 (7): 543-556.

Hubert, W. A., and F. J. Rahel. 1989. Relations of physical habitat to abundance of four nongame fishes in high-plains streams: A test of habitat suitability index models. *North American Journal of Fisheries Management* 9:332-340.

Keeney, R. L. 2007. Developing objectives and attributes. In *Advances in decision analysis: From foundations to applications*. ed. W. Edwards, R. F. Miles, Jr., and D. von Winterfeldt, 104-128. New York: Cambridge University Press.

Keeney, R. L., and R. S. Gregory. 2005. Selecting attributes to measure the achievement of objectives. Operations research. *Informs* 53 (1): 1-11.

Kurtz, J. C., L. E. Jackson, and W. S. Fisher. 2001. Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development. In *Ecological Indicators.* Elsevier. 1:49-60.

Leftwich, K. N., P. L. Angermeier, and C. A. Dolloff. 1997. Factors influencing behavior and transferability of habitat models for a benthic stream fish. *Transactions of the American Fisheries Society* 126:725-734.

Morreale, S. J., and J. W. Gibbons. 1986. *Habitat suitability index models: Slider turtle*. Biological Report 82(10.125). U.S. Fish and Wildlife Service.

Orians, G. H., M. Alexander, P. L. Brezonik, G. Brush, E. Gorham, A. Janetos, A. H. Johnson, D. L. Markowitz, S. W. Pacala, J. Pastor, G. W. Petersen, J. R. Pratt, T. Root, M. L. Rosenzweig, M. Russell, and S. Stafford. 2000. *Ecological indicators for the Nation*. Washington, DC: National Academies Press.

Palmer, M. A., E. S. Bernhardt, J. D. Allan, P. S. Lake, G. Alexander, S. Brooks, J. Carr, C. N. Dahm, J. Follstad Shah, D. L. Galat, S. G. Loss, P. Goodwin, D. D. Hart, B. Hassett, R. Jenkinson, G. M. Kondolf, R. Lave, J. L. Meyer, T. K. O'Donnell, L. Pagano, and E. Sudduth. 2005. Standards for ecologically successful river restoration. *Journal of Applied Ecology* 42:208-217.

Poole, G. C., C. A. Frissell, and S. C. Ralph. 1997. In-stream habitat unit classification: Inadequacies for monitoring and some consequences for management. *Journal of the American Water Resources Association* 33(4):879-896.

Reichert, P., M. Borsuk, M. Hostmann, S. Schweizer, C. Sporri, K. Tockner, and B. Truffer. 2007. Concepts of decision support for river rehabilitation. *Environmental Modeling and Software.* Elsevier. 22:188-201.

Rosgen, D. L. 2001. A practical method of computing streambank erosion rate. *Proceedings of the Seventh Federal Interagency Sedimentation Conference, Reno, Nevada, March 25-29, 2001*, 2:9-15.

Slocombe, D. S. 1998. Defining goals and criteria for ecosystem-based management. *Environmental Management* 22(4):483-493.

Somerville, D. E., and B. A. Pruitt. 2004. *Physical stream assessment: A review of selected protocols for use in the Clean Water Act Section 404 Program*. Washington, DC: EPA, Office of Wetlands, Oceans, and Watersheds, Wetlands Division.

Suter, G. W. 2001. Applicability of indicator monitoring to ecological risk assessment. *Ecological Indicators*. Elsevier 1:101-112.

Tear, T. H., P. Karieva, P. L. Angermeier, P. Comer, B. Czech, R. Kautz, L. Landon, D. Mehlman, K. Murphy, M. Ruckelshaus, J. M. Scott, and G. Wilhere. 2005. How much is enough? The recurrent problem of setting measurable objectives in conservation. *Bioscience* 55(10):835-849.

Thayer, G. W., A. D. Nickens, T. A. McTigue, S. J. Lozano, R. J. Bellmer, P. F. Gayaldo, F. M. Burrows, P. J. Plmateer, D. H. Merkey, and P. T. Pinit. 2003. Science based restoration monitoring of coastal habitats. NOAA *Decision Analysis Series* 1(23).

Thayer, G. W., D. H. Merkey, T. A. McTigue, F. M. Burrows, R. J. Salz, and P. F. Gayaldo. 2005. Science based restoration monitoring of coastal habitats. *NOAA Decision Analysis Series* 2(23).

U.S. Army Corps of Engineers (USACE). 1999a. *Civil Works Ecosystem Restoration Policy*. ER-1165-2-501. Washington, DC.

U.S. Army Corps of Engineers (USACE). 1999b. *Ecosystem Restoration – Supporting Policy Information*. EP-1165-2-502. Washington, DC.

U.S. Army Corps of Engineers (USACE). 2000. *Planning Guidance Notebook*. ER-1105-2-100. Washington, DC.

U.S. Environmental Protection Agency (USEPA). 2009. Environmental Monitoring and Assessment Program (EMAP). *http://www.epa.gov/emap/*. Accessed 19 August 2009.

U.S. Forest Service (USFS). 2009. National Environmental Status and Trends Indicator Project (NEST). *http://www.fs.fed.us/NEST/*. Accessed 19 August 2009.

Vadas, R. L., and D. J. Orth. 2001. Formulation of habitat suitability models for stream fish guilds: Do the standard methods work? *Transactions of the American Fisheries Society* 130:217-235.

Wohl, E., P. L. Angermeier, B. Bledsoe, G. M. Kondolf, L. MacDonell, D. M. Merritt, M. A. Palmer, N. L. Poff, and D. Tarboton. 2005. River restoration. *Water Resources Research* 41.

Yoe, C. 2002. *Trade-off analysis planning and procedures guidebook*. IWR Report 02-R-2. Alexandria, VA: U.S. Army Institute for Water Resources.

Yoe, C. E., and K. D. Orth. 1996. *Planning Manual*. IWR Report 96-R-21. Alexandria, VA: Institute for Water Resources.

Young, R. G., C. D. Matthaei, and C. R. Townsend. 2008. Organic matter breakdown and ecosystem metabolism: Functional indicators for assessing river ecosystem health. *Journal of the North American Benthological Society* 27(3):605-625.