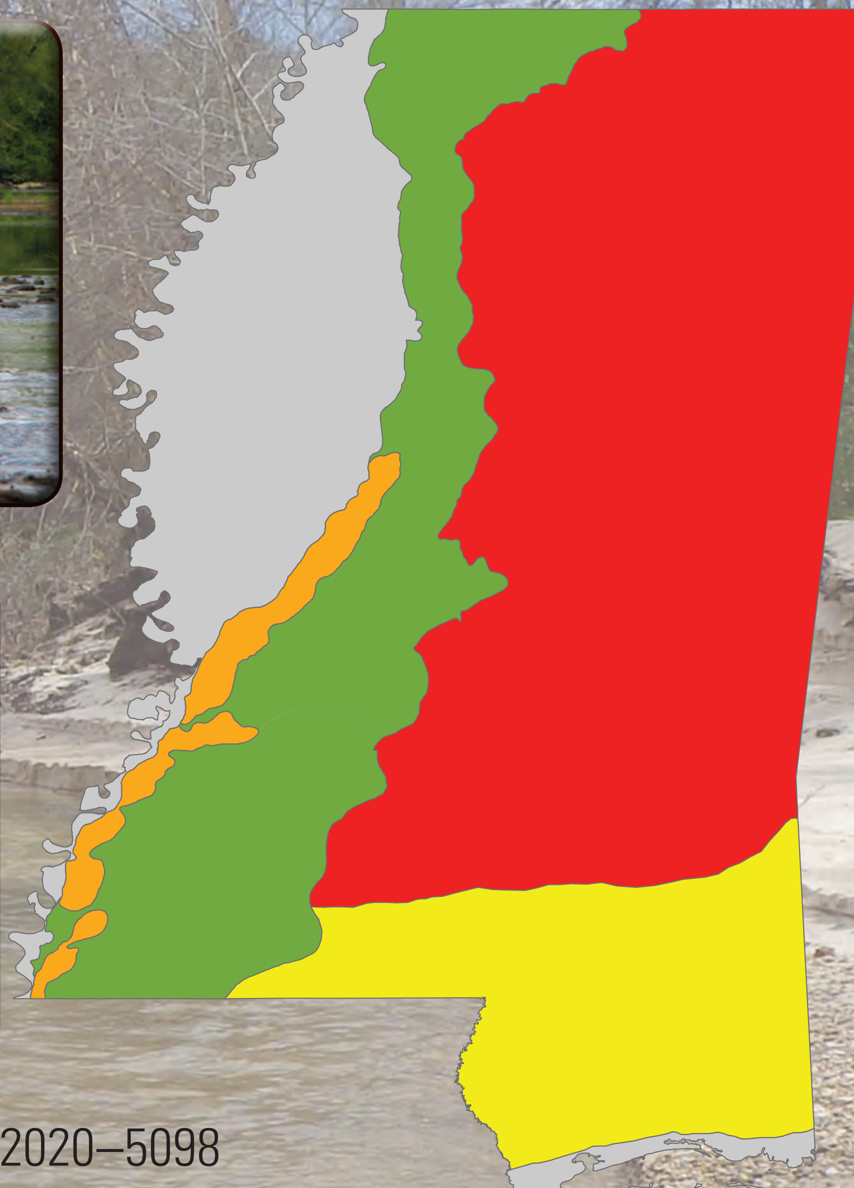
Prepared in cooperation with the Mississippi Department of Environmental Quality

Stressor Identification Framework of Biological Impairment in Mississippi Streams to Support Watershed Restoration and TMDL Development



Scientific Investigations Report 2020–5098

U.S. Department of the Interior
U.S. Geological Survey

Cover.

Background, Foster Creek, Amite County, Mississippi.

Top left, Chunky River, Lauderdale County, Mississippi.

Bottom left, Riverdale Creek, Grenada County, Mississippi.

Map showing Mississippi Benthic Index of Stream Quality bioregions (see fig. 7).

Photographs by Matt Hicks, U.S. Geological Survey.

Stressor Identification Framework of Biological Impairment in Mississippi Streams to Support Watershed Restoration and TMDL Development

By Matthew B. Hicks and Jennifer M. Cartwright

Prepared in cooperation with the Mississippi Department of Environmental Quality

Scientific Investigations Report 2020–5098

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
DAVID BERNHARDT, Secretary

U.S. Geological Survey
James F. Reilly II, Director

U.S. Geological Survey, Reston, Virginia: 2020

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit https://www.usgs.gov or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit https://store.usgs.gov.

Suggested citation:
Hicks, M.B., and Cartwright, J.M., 2020, Stressor identification framework of biological impairment in Mississippi streams to support watershed restoration and TMDL development: U.S. Geological Survey Scientific Investigations Report 2020–5098, 42 p., https://doi.org/10.3133/sir20205098.

ISSN 2328-0328 (online)

Contents

Figures

Tables

Conversion Factors

International System of Units to U.S. customary units

Multiply	By	To obtain
	Length	
kilometer (km)	0.6214	mile (mi)
kilometer (km)	0.5400	mile, nautical (nmi)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:
°F = (1.8 × °C) + 32.

Datum

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).

Supplemental Information

Specific conductance is given in micromhos per centimeter (µmho/cm).

Abbreviations

ρ	Spearman's rank correlation coefficient
ANOVA	analysis of variance
BRT	boosted regression trees
CADDIS	Causal Analysis/Diagnosis Decision Information System (EPA)
CART	classification and regression tree
CWA	Clean Water Act
DA	discriminant analysis
DO	dissolved oxygen
EPA	U.S. Environmental Protection Agency
EPT	Ephemeroptera, Plecoptera, and Trichoptera
HUC	hydrologic unit code
IBI	index of biological integrity
LD	least disturbed
M-BISQ	Mississippi Benthic Index of Stream Quality
MDEQ	Mississippi Department of Environmental Quality
NI	nonimpaired
RF	random forests
SI	stressor identification
SSC	site-specific comparator
TMDL	total maximum daily load
USGS	U.S. Geological Survey
WOE	weight of evidence

Stressor Identification Framework of Biological Impairment in Mississippi Streams to Support Watershed Restoration and TMDL Development

By Matthew B. Hicks and Jennifer M. Cartwright

Abstract

The Clean Water Act (CWA) requires States to identify waters that are impaired for designated uses. These waters are published through a State's §303(d) list. The CWA also requires that a total maximum daily load (TMDL) be completed for each water body to calculate the maximum amount of contaminants that can be present in that water body and still meet water-quality standards. The Mississippi Department of Environmental Quality (MDEQ) uses a statewide monitoring and assessment strategy to collect benthic macroinvertebrate community data to assess the health of streams and rivers and to identify impaired waters. Waters that are found to be impaired based on the macroinvertebrate community data are listed on the Mississippi §303(d) list, and the cause of impairment is listed as "biological impairment." Although the CWA requires TMDLs to be developed for applicable contaminants identified in the §303(d) list, TMDLs cannot be computed for stream reaches in Mississippi listed for biological impairment because the actual stressors causing the impairment have not yet been determined. The MDEQ and other water-resource managers in Mississippi require a framework for stressor identification in biologically impaired streams and rivers. This report is organized to (1) provide a general overview of biological impairment and stressor identification in stream ecosystems and (2) provide a detailed framework for stressor identification of Mississippi streams that are biologically impaired. The intent is for the framework to reduce subjectivity, provide consistency, and allow for adaptation as the science evolves. The stressor identification framework for Mississippi involves six key steps:

1. Define the impairment,
2. List the candidate causes of impairment and develop a conceptual model,
3. Compile all relevant data,
4. Evaluate the data,
5. Identify probable causes of impairment by using a weight-of-evidence approach, and
6. Generate a report of results.

Introduction

The Clean Water Act (CWA), §305(b), requires each State to describe the quality of their surface-water and groundwater resources in a report for the U.S. Environmental Protection Agency (EPA), Congress, and the public on a biennial basis. The Mississippi Department of Environmental Quality (MDEQ), as the lead agency for environmental protection in Mississippi, is the State agency responsible for generating this report. To support this reporting process, statewide data are routinely collected through several monitoring activities associated with surface-water-quality programs. To assess the health of waters, MDEQ uses a modified process from EPA guidance (EPA, 1997, 2002). Waters assessed as not having the quality to attain their designated uses in the §305(b) assessment process become candidates for listing on Mississippi's §303(d) list (Stribling and others, 2016).

Most of the data used to assess the overall health of streams and rivers in Mississippi and generate the §305(b) report and the §303(d) list are from statewide monitoring efforts of biological assemblages. Biological assemblage data, based on taxonomic identification and enumeration of benthic macroinvertebrates, are used to calculate a regionally calibrated index of biological integrity (IBI), the Mississippi Benthic Index of Stream Quality (M-BISQ). For a given site, the M-BISQ score is compared to a least-disturbed condition score for the region of the State in which the site is located. If the M-BISQ score is below the least disturbed condition score, the stream segment is placed on the §303(d) list, with the cause listed as "biological impairment." The CWA requires action to be taken to restore the health of these listed impaired waters. Efforts to address biological impairment in streams and rivers—such as watershed restoration projects or development of total maximum daily loads (TMDLs)—require accurate identification of the specific stressors causing the impairment.

In many cases, stressor identification (SI) is complicated because multiple stressors may interact to produce biological impairment. Since the early 2000s, several scientific studies have focused on characterizing stressor-response relations in the context of natural and human-induced landscape variables and changes (Downes, 2010; Wagenhoff and others, 2011; Allan and others, 2013). However, these studies have been conducted in Australia, Europe, and in the upper Midwest and Northeast regions of the United States, and are either broad in scope, or are focused on a single stressor or a single biological response indicator. The need exists for improved understanding of connections and responses of biological communities to agents of stress at smaller, regional scales and to specific stressors in ecosystems of varying geology, topography, and land use.

In 2003, MDEQ developed and implemented an SI process to identify specific causes of impairment for waters that are on the §303(d) list for biological impairment, so that actions could be taken to improve water quality. It has been more than a decade since the original SI process was developed, and it has not been updated to reflect recent scientific advancements in characterization of stressor-response relations. Several analytical and statistical frameworks have recently been applied to the problem of SI in other States (Griggs and Buchanan, 2012). Although each State has approached the problem somewhat differently, these frameworks commonly rely on a weight-of-evidence (WOE) approach to characterize and compare candidate stressors for biological impairment.

Goal and Objectives of This Document

The goal of this document is to provide a framework suitable for conducting identification of stressors causing impairment in biologically impaired streams in Mississippi. The intent is for the framework to (1) reduce subjectivity and provide consistency in the SI process, (2) allow for well-documented and clearly understood steps taken in the process, and (3) result in a sustainable approach for future use, allowing for adaptation as the science evolves. To accomplish this goal, the following objectives are described in detail in this document:

- Provide a general overview of biological impairment and SI in stream ecosystems generally and in Mississippi specifically,
- Describe the general framework that was developed,
- Describe common complications and shortcomings encountered in the SI process and provide options for addressing them, and
- Provide options for future improvement of the SI framework.

General Overview of Stressor Identification in Stream Ecosystems

Biological Impairment

Biological impairment of streams is a common ecological problem, especially in watersheds that are heavily affected by intensive land uses such as agriculture, development, and resource extraction. Improvements in biological integrity, water quality, and physical habitat are key goals of the CWA and State-level water-resource management activities. Efforts to address biological impairment in streams—such as watershed restoration projects or development of TMDLs for contaminants—require accurate identification of the specific stressors causing the impairment. In this context, a stressor can be defined as “a variable that, as a result of human activity, exceeds its range of normal variation and adversely affects individual taxa or community composition” (Townsend and others, 2008).

Biological responses to stressors can occur at various taxonomic levels and can manifest in a variety of ways. Biological data indicating impairment can be derived from individual organisms (for example, deformities, tumors, or mortality), from populations of a given species (for example, reductions in local abundance or changing demographic structure), or from assemblages or communities of multiple species (for example, changes in species richness or shifts in community structure based on taxon-level characteristics) (Suter and others, 2002). Common responses of biological communities to stressors include (1) decreased species richness, (2) loss of sensitive species with increased dominance of a small number of tolerant species, (3) increase in generalists and decrease in specialists, (4) reduced overall biomass, and (5) smaller populations of each species, even with no discernible change in community structure (Klein, 1979; Karr and others, 1986; Onorato and others, 1998; Cormier and others, 2000).

Causal Assessment and Stressor Identification

A formal method of SI has several advantages over informal or ad hoc approaches to determining causes of biological impairment. In many cases, definitive causes of biological impairment in streams will not be immediately obvious, and a structured process of SI will be required to accurately identify likely causes. In addition, informal methods of causal inference are inherently subjective and prone to problems (for example, bias or lapses of logic) that can only be revealed by a formal method (Suter and others, 2002). For these reasons, informal methods may fail to convince skeptical stakeholders and may be vulnerable to legal challenge. In cases where remediation of identified stressors has the potential to be costly, time-consuming, politically challenging, or economically disruptive, it is essential that remediation efforts be appropriately targeted to achieve optimal results, thus strengthening the need for a structured and formal SI process. In environmental law, general causation concerns the question of whether an environmental cause (in this case, a candidate stressor) is *capable* of producing a deleterious result (in this case, biological impairment), whereas specific causation concerns the narrower question of whether, at a given site, the stressor in question actually *did* produce biological impairment (Schleiter, 2009). In a formal SI process, general causation may be a necessary but insufficient basis for establishing specific causation. In many cases, specific causation cannot be decisively proven; thus, the SI process involves the compilation and consideration of evidence for or against specific causation.

The EPA issued official guidance on the SI process in 2000 to provide a consistent and actionable framework for causal inference in biologically impaired stream ecosystems (EPA, 2000). This SI framework has since been adapted and applied in multiple State and regional contexts (Griggs and Buchanan, 2012; Minnesota Pollution Control Agency, 2013; Virginia Department of Environmental Quality, 2014). The EPA also has provided an online resource to support SI efforts: the Causal Analysis/Diagnosis Decision Information System (CADDIS) (EPA, 2010a). CADDIS provides background information on many common types of stressors, illustrative examples of the SI process from case studies, guidance on the use of analytical methods to support causal inference, and a literature database from which to derive evidence for causal assessments.

Once an impairment is detected or suspected, the SI process involves three main steps: (1) listing candidate causes of impairment, (2) analysis of available evidence, and (3) causal characterization to draw conclusions (EPA, 2000; Suter and others, 2002; Cormier and others, 2003). As illustrated in figure 1, the SI process is conceptually separate from, but also directly informs, two subsequent steps in ecological remediation: apportionment of sources of the identified stressors and management actions to directly address the identified stressors. The SI process is designed to be iterative if needed, such that results of management actions informed by a preliminary SI process can provide

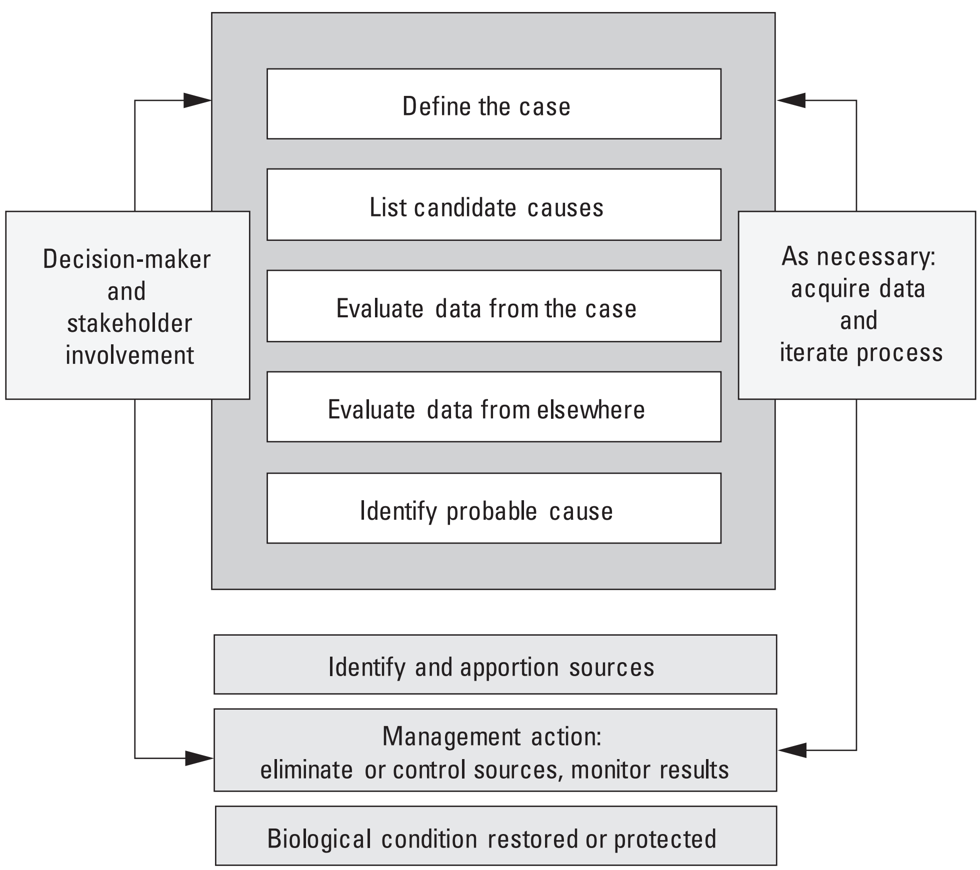


Figure 1. Overview of the U.S. Environmental Protection Agency (EPA) stressor identification process (from EPA, 2000).

new evidence for subsequent rounds of SI (EPA, 2000; Suter and others, 2002; Cormier and others, 2003). Decision-maker and stakeholder engagement is encouraged at each step in the SI process (EPA, 2000; Suter and others, 2002; Cormier and others, 2003).

General Steps of a Stressor Identification Process

Getting Started

Two important tasks are required before initiating the SI process. First, the scope of the SI process needs to be clearly defined, which involves delineating a geographic area for analysis, defining a project timeframe, and itemizing available data types for analysis. Key considerations to inform these decisions include the regulatory context, resource availability (for example, data availability, personnel costs, and available budget), and stakeholder interests and concerns (EPA, 2000).

Second, the nature of the biological impairment should be described using specific quantitative characterizations, such as counts or continuous variables, when possible. Multiple forms of impairment should be described separately rather than aggregating into an overall metric (Cormier and others, 2003). For example, a decline in fish species richness should be described separately from a decrease in mean fish body size. This is especially important when biological data are combined into a multimetric index. Defining biological impairment should be described with individual traits and (or) taxa presence, absence, or relative abundance and not the overall multimetric index score. This helps ensure that stressors can be accurately identified for each type of biological impairment. Each impairment should be described in terms of its nature, magnitude, and spatial and temporal extents (Suter and others, 2002). In general, exceedance of regulatory criteria is not a sufficient description of biological impairment because (1) not all stressors have such criteria, (2) numeric criteria may not accurately represent true thresholds that cause impairment, and (3) criteria for individual contaminants do not allow for synergistic effects or effects of sporadic spikes in concentration (EPA, 2000; Morris and others, 2006). As opposed to relying on a single point estimate, such as criterion, the original species sensitivity distribution models used to develop the criteria may be more useful.

Step 1. List Candidate Causes

The goal of the first step is to develop a comprehensive list of candidate stressors within the spatial and temporal scope of the SI project. Initially, an exhaustive list of all possible causes of impairment (stressors) is ideal. Identification of candidate stressors may derive from expert opinion, existing data, qualitative observations, and stakeholder input. Whenever possible, characterization of likely stressor sources can be helpful. This includes, for example, differentiating point-source from nonpoint-source contaminants and georeferencing possible point sources, such as wastewater effluent discharge sites. Stressors that are deemed highly improbable may be removed from the final list of candidates, accompanied by clear written explanations for the reasons justifying removal (EPA, 2000).

Conceptual models linking candidate stressors to observed metrics of impairment are an important component of the listing process. Figure 2 gives an example conceptual model from EPA (2010b) that shows sources, stressors, and causes with linkages in the Coal River watershed in West Virginia. Furthermore, including sources or activities associated with the proximate stressor is helpful because it sets the stage for load allocation in the process of TMDL development. Typically, a causal chain is required to link a candidate stressor to an observed impairment. For a given candidate stressor, the entire causal scenario should be articulated rather than listing each link in the chain as a separate stressor. For example, a causal scenario could involve agricultural runoff entering a stream, which increases nutrient levels, which promotes algal growth, which depletes dissolved oxygen (DO), which produces a fish kill (the observed biological impairment). In this example, algal growth itself is not a stressor; it is an intermediate link in a causal chain for which agricultural runoff is the ultimate stressor. Commonly, multiple competing hypotheses may exist for the pathways linking a stressor to the observed impairment. In such cases, each hypothetical conceptual model should be outlined separately. Ideally, these hypotheses should be refutable based on existing quantitative data. Additionally, each hypothetical causal pathway should be linked to a prediction that is unique to that pathway (Downes, 2010).

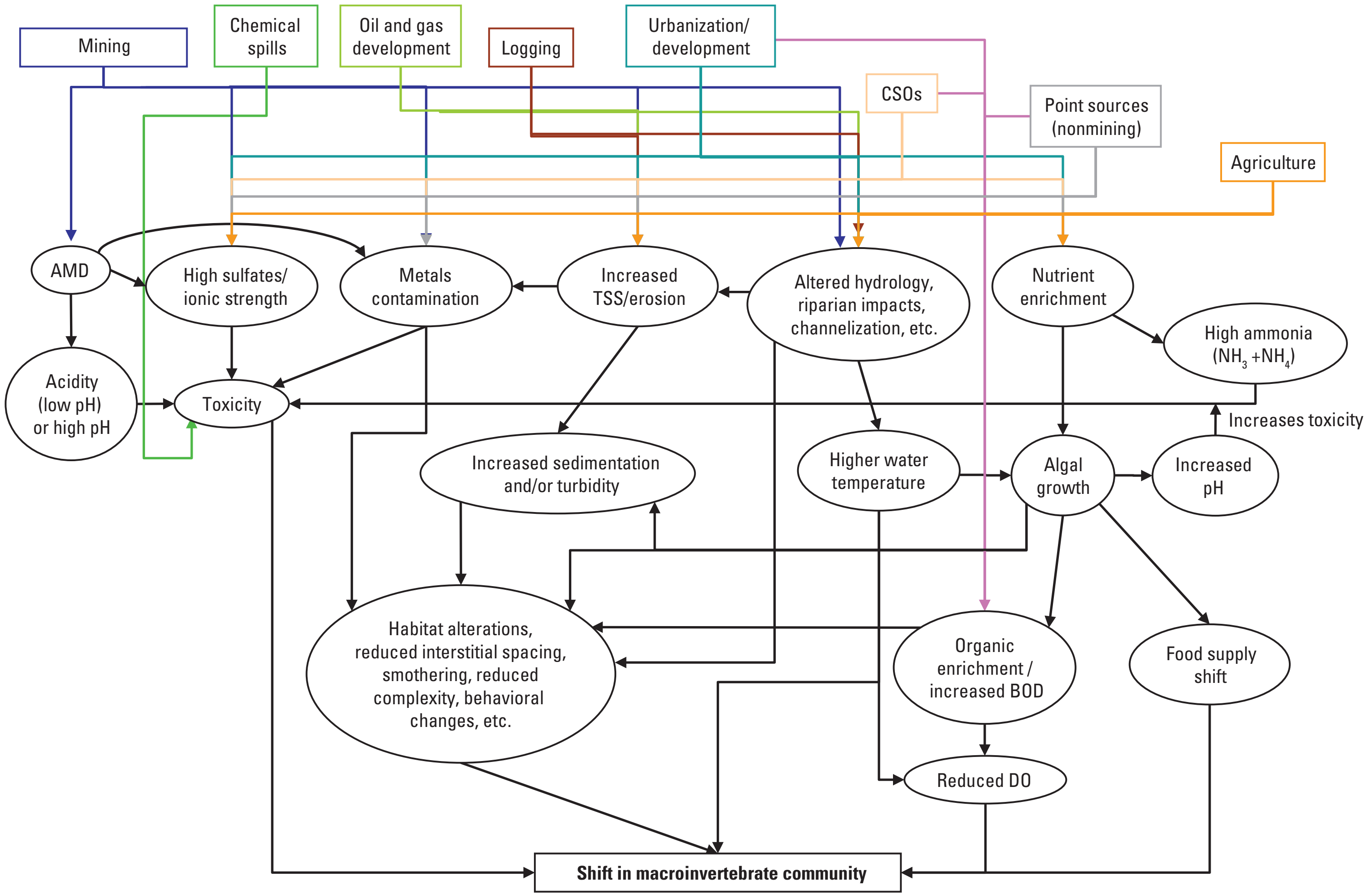


Figure 2. Conceptual model of sources and stressors in the Coal River watershed, West Virginia (from U.S. Environmental Protection Agency, 2010b). [AMD, acid and metalliferous drainage; CSOs, combined sewer overflows; TSS, total suspended solids; BOD, biological oxygen demand; DO, dissolved oxygen]

Step 2. Analyze Evidence by Using Four Categories of Association

The goal of the second step is to perform quantitative and qualitative analyses to determine strength of association between candidate stressors and observed biological impairments. EPA guidance on SI processes suggests four categories of association by which evidence may be evaluated (EPA, 2000; Suter and others, 2002; Cormier and others, 2003):

1. Associations between measurements of candidate causes and effects. This type of association can be demonstrated by co-occurrence of a candidate cause and an observed effect at the same time and place. For this co-occurrence to be valid, the effect should be observed at the same time and place as the candidate cause, and the effect should not be observed when and where the candidate cause is not present. An example would be a case in which algal, chlorophyll *a* measurements and nutrient concentrations at a study site were greater than those at the reference site. Where spatial or temporal gradients exist in candidate causes and biological impairments, the intensity of the candidate cause and its proposed effect should be positively associated.

2. Association between candidate causes and effects at the site with causes and observed effects elsewhere. This type of association combines measures of exposure to a stressor at a study site with lab results or field studies from other sites which demonstrate effects of exposure. If lab tests, field tests, or observational studies from other sites indicate exposure-response relations between a stressor and a biological outcome, an important consideration is whether the stressor is present at the study site in sufficient quantity and frequency for the anticipated result to occur. An additional consideration is that lab conditions and organisms may not accurately represent field conditions. If no lab or field data exist for a candidate cause, analogous or surrogate data may be used (see Insufficient Data section), although this will increase the uncertainty of the association. An example of this association would be a situation in which 19 out of 20 streams in the same ecoregion with similar levels of organic enrichment had depressed DO.

3. Measurements at the site associated with the candidate causal mechanism. This type of association can increase confidence in the proposed causal mechanism by documenting intermediate links in a causal chain. For example, a causal scenario by which agricultural runoff produces fish kills by means of nutrient enrichment and hypoxia is more plausible if increases in algal biomass are measured. This type of association is particularly helpful in evaluating multiple competing causal mechanistic explanations that produce similar biological outcomes (Suter and others, 2002).

4. Association of effects of mitigation and manipulation of causes. This type of association may involve reduction or elimination of candidate stressors from a study site by manipulating key sources (for example, cattle exclusion from streams or effluent absence during shutdowns at a water-treatment plant). These “natural experiments” provide opportunities to obtain strong causal evidence if biological impairment is diminished following elimination of a candidate stressor. Lab experiments may also provide causal evidence by artificially eliminating (under controlled lab conditions) a stressor that exists in field conditions at a study site. Field experiments that control exposure can also provide causal evidence, such as caging previously unexposed organisms at contaminated locations (Crane and others, 2007). An example of this association would be a case in which stocked trout populations failed to survive over a 2-year period when fields were in alfalfa rotation with no pesticide application.

In summary, this step seeks to use a process that eliminates candidate causes early in the SI process if there is evidence of an interrupted causal pathway or evidence of lack of co-occurrence of a response with the candidate cause. If a candidate cause cannot be eliminated, then the process entails using strength of associations by using thresholds of effects from either the laboratory or the field (fig. 3).

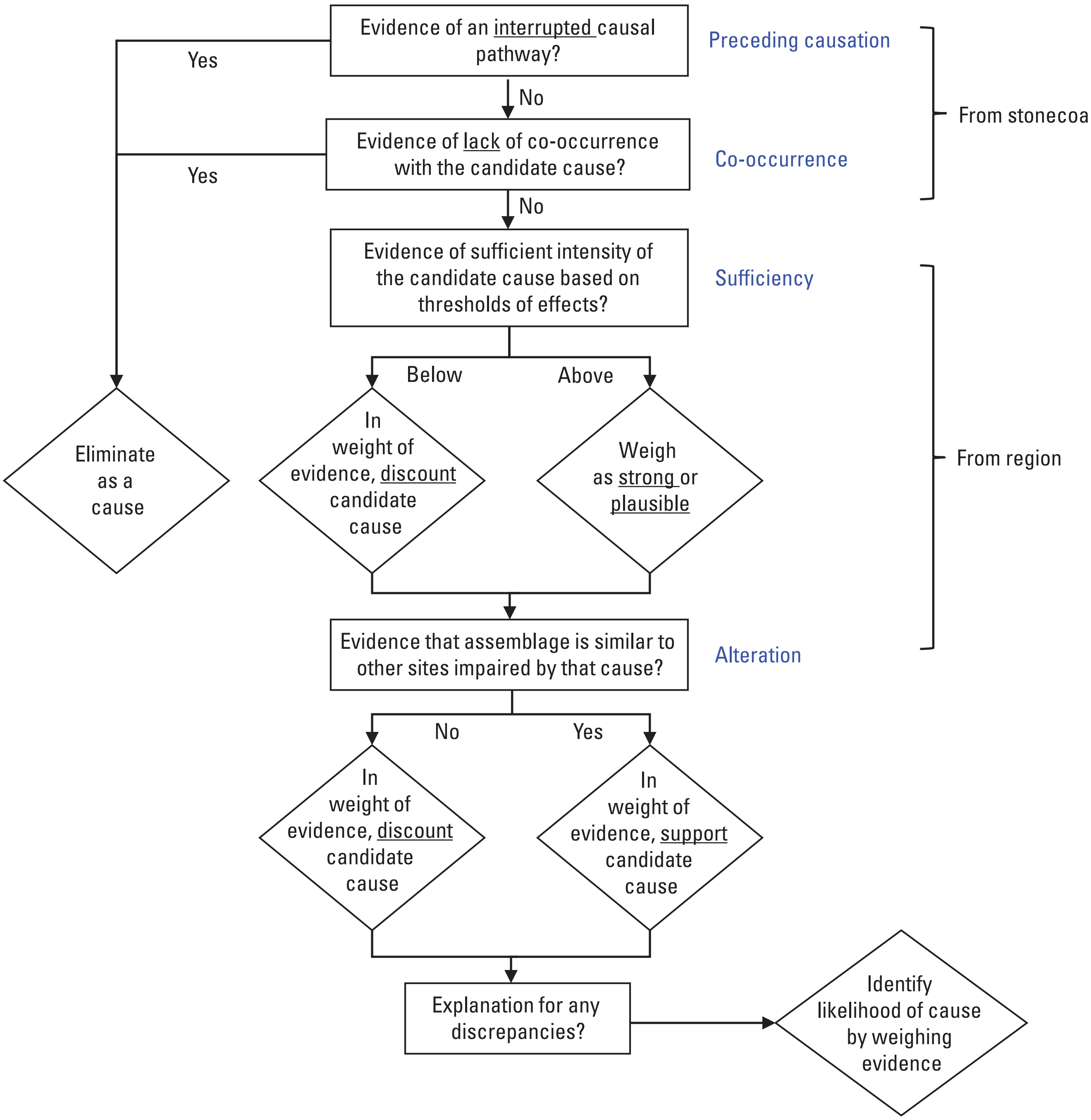


Figure 3. Diagram illustrating the process used to identify a candidate cause during the stressor identification process (from U.S. Environmental Protection Agency, 2010b).

Step 3. Characterize Causes by Using Elimination, Diagnosis, and Weight of Evidence

The goal of the third step is to characterize candidate stressors to draw conclusions about probable cause(s) of biological impairment at a study site. This step may involve elimination of candidate causes, diagnosis of causes, or causal evaluation by using a WOE approach. It is important to remember that the SI process involves building a case for and against alternative candidate cause. Evidence against a candidate cause is just as important as evidence for a candidate cause.

Where available evidence contradicts a candidate causal scenario, the candidate cause may be removed from further consideration. In general, only evidence from the study site in question (not from other related study sites) should be used for elimination. Elimination of candidate causes may be subject to several problems depending on the relations between stressors at a site. For example, ancillary causes may be masked by primary causes, thus leading ancillary causes to be erroneously eliminated. For more examples of complications by which one type of stressor may mask another, see the Causal Masking section in this report and table 4-1 in EPA (2000).

For certain stressors, causal diagnosis may be possible following techniques adapted from medical practice. For example, some pathogens and toxins produce unique and characteristic symptoms that may enable diagnosis, and similarly, hypoxia may be diagnosable based on characteristic fish behavior (gasping for air at the water surface). For many other stressors and for invertebrate impairment, definitive diagnostics are generally not available.

Following elimination or diagnosis of candidate causes (if applicable), any remaining candidate causes may be evaluated by using a WOE approach. EPA (2010b) present six characteristics of causation to organize arguments that directly show how types of evidence support the inference, which may simplify the organization of evidence and the reason for identifying some causes over others. The six characteristics of causation are the following:

1. Co-occurrence—Cause and effect co-occur at the same place in a study site,
2. Sufficiency—Cause is at levels known to cause observed effect,
3. Time order—Cause temporally precedes effect,
4. Alteration—The entity is changed by the interaction with the cause,
5. Antecedence—Each causal relation is a result of a larger web of cause and effect relations, and
6. Interaction—The cause physically interacts with the entity in a way that induces the effect.

By using these causal characteristics that are relevant to the study site, type of biological impairment observed, and datasets available, the SI process concludes with an identification of probable causes of impairment, accompanied by a detailed description of the conceptual, analytic, and logical bases for the conclusions drawn. In addition, an evaluation of uncertainty relating to the SI process should be documented.

Common Complications and Options for Addressing Them

SI efforts are commonly made more challenging by one or more complicating factors. These factors include insufficient data, presence of confounding environmental factors, and the possibility for one or more causes to mask another cause.

Insufficient Data

Limitations on data quality and spatial and temporal completeness are common in SI projects. Where resources permit acquisition of additional data, a listing of candidate causes (the first step in the SI process) may support prioritization of locations and data types for future data collection efforts.

If resource limitations preclude additional data collection, use of surrogate variables may be considered. For example, electrical conductivity can be used as a surrogate for total dissolved solids, and turbidity can be used as a surrogate for total suspended solids (Miguntanna and others, 2010; Vander Laan, 2012). Water-column chlorophyll *a* concentration can be a surrogate for algal biomass (Dodds and Welch, 2000; EPA, 2000). Additional surrogates may be identified by searching the CADDIS literature database. In some cases, information on locations and attributes of stressor sources may be used as surrogates. For example, locations and times of effluent release for wastewater treatment plants might be used as surrogates for chemical analysis of effluent (EPA, 2000; Suter and others, 2002). Similarly, Vander Laan and others (2013) used total upstream reservoir volume and volume of the largest upstream reservoir from the National Inventory of Dams (U.S. Army Corps of Engineers, 2009) as surrogates for hydrologic alteration resulting from impoundments upstream from a study site.

Notably, use of surrogate data increases uncertainty associated with causal identification (EPA, 2000; Suter and others, 2002). Thus, it is important that surrogate variables be clearly identified as such in descriptions of the SI process. Use of surrogate variables involves assumptions that one variable (for which data exist) effectively represents another variable (for which data are absent or of insufficient quality) (EPA, 2000; Suter and others, 2002). The degree to which the use of surrogates is justified depends upon the establishment of a clear, direct, and ideally quantitative relation between the surrogate variable and the variable it is intended to represent.

Confounding Environmental Factors

One of the most common challenges in the SI process involves differentiation of stressor effects on biological outcomes from natural variability in biological states and processes. Several confounding factors have been identified in the course of SI projects and have been addressed by using various analytical approaches (EPA, 2000; Suter and others, 2002). For example, natural landscape gradients commonly covary with anthropogenic influences (Allan and others, 2012). Clements (1994) used multivariate regression to address natural seasonal and longitudinal (upstream to downstream) changes in community composition, which complicate the discernment of community shifts in response to metal contamination by comparison of upstream reference sites to sites downstream of effluent discharge points. Norton and others (2000) accounted for correlations among stream size and stressor variables and biological response variables by performing regression analysis on variables with drainage area and using the regression residuals in subsequent analysis. Similarly, Vander Laan and others (2013) addressed the issue of natural gradients in stressor variables by modeling expected values and calculating differences between observed and expected values. Morris and others (2006) used site-selection criteria to address the issue of low fish abundance in intermittent headwater streams. As these examples illustrate, analytical approaches to deal with confounding variables should be chosen based on knowledge of the types of correlations present in the data at hand, as well as correlations with unmeasured environmental factors (Downes, 2010).

Typically, multiple stressors interact at any given study site to produce an observed biological outcome (EPA, 2000; Morris and others, 2006; Downes, 2010). For example, nutrient enrichment may interact with increased stream temperature to promote hypoxia. The interactive effects of multiple stressors can be difficult to predict based only on knowledge of individual stressor effects (Townsend and others, 2008; Matthaei and others, 2010). However, modeling these interactions can be accomplished by using a variety of analytical and statistical approaches. For example, Townsend and others (2008) used interaction terms in multiple linear regression models (for example sediment × nutrient concentrations). Matthaei and others (2010) addressed stressor interactions by experimentally manipulating three common stressors (nutrient loads, sediment additions, and water withdrawals) to quantify their individual and interactive effects. Several machine-learning approaches and variance-partitioning models provide robust modeling options in situations where multiple explanatory variables are correlated (Cutler and others, 2007; Elith and others, 2008; Smucker and others, 2013).

Rather than relying solely on statistical methods to model complex interactions, it can be helpful to incorporate possible interactions into the causal pathways identified in the SI process because the nature of stressor interactions depends in part on modes of action of the individual stressors. For example, multiple toxicants have similar modes of action in crustaceans such that their effects are multiplicative, whereas functional synergy occurs between copper toxicity and low DO, whereby increased gill movement increases copper uptake in mayfly nymphs (Townsend and others, 2008). The nature and magnitude of stressor interaction may also depend on whether the stressors involved directly affect physiology or behavior, or indirectly affect food webs.

Causal Masking

Because streams are complex systems, one or more causes of impairment may mask yet another cause for impairment, thus making its detection difficult or impossible. For example, causes of impairment located upstream may mask downstream causes, or severe causes of impairment (principal causes) may mask more subtle ones (ancillary causes). Table 4-1 in EPA (2000) presents several other examples of masking considerations. When the goals of SI include ecosystem restoration, the objective is to identify all causes that are able to impair the stream biology. To do this, each candidate cause can be eliminated iteratively with continuous monitoring of stream biological indicators. Each candidate cause eliminated during the first round of ecosystem restoration should be reexamined to determine whether it had been masking secondary causes (Suter and others, 2002). Notably, if a principal cause persists, removal of ancillary causes will produce little to no improvement in stream biological indicators. In addition, unresolved ancillary causes of biological impairment can restrict improvement even after the principal cause is removed (EPA, 2000). Any action (for example, regulatory, remediation, or ecosystem restoration) should be followed by monitoring to assess effects (Suter and others, 2002), with a goal of identifying any previously unexamined ancillary causes.

In addition, geographic movement of organisms may mask impairment and complicate the detection of causes. For example, local reproductive impairment may be masked if organisms can migrate from adjacent habitat to fill a niche left vacant because of the reproductive impairment (Cormier and others, 2000). This form of impairment is more likely to be an issue for mobile species, such as fish, than for sessile species, such as mussels.

Other Important Considerations

Definitions of Impairment and Candidate Causes

The outcome of an SI process may be sensitive to the manner in which impairment and candidate causes are defined. In general, biological impairments need to be defined precisely, which often requires defining multiple interrelated impairments (Cormier and others, 2003). To accurately target causes of impairment, aggregate metrics—such as an IBI—should be avoided because they conflate many distinct biological variables and hamper the process of identifying mechanistic causes of individual biological impairments. Instead, aggregate metrics can be broken down into their component variables. In addition, although taxonomic aggregation (for example, from family or order) can lower data collection costs, it can mask impairment of sensitive species within taxonomic groups (Clements and others, 2000), indicating that taxonomic detail can be important for accurate SI results (Downes, 2010).

Similarly, candidate causes need to be defined as precisely as possible; for example, use “riffle siltation” rather than “poor habitat” (Suter and others, 2002). In general, aggregate causes may be difficult to address, and so, are less helpful to identify than their specific component causes. For example, Wang and others (2000) examined relations between urbanization (an aggregate cause of impairment) and fish community structure, which is of limited value in identifying which specific stressors were at work (for example, contaminants, suspended sediment, or impervious surface). Remediation may be possible for certain individual stressors or sets of stressors, but urbanization itself is unlikely to be reversible.

Regional Clustering of Impairment

Certain regions may have particular sets of stressors that are widespread and common to many streams, owing to particular land-use practices that are regionally common, such as metal contamination from mining in the southern Rocky Mountains in Colorado (Clements, 1994; Clements and others, 2000) and degraded channel morphology from dredging in agricultural streams in Ohio (Norton and others, 2000; Morris and others, 2006). Known stressors established in the literature for a given region are important candidate causes in new SI investigations but should not be assumed to be the only (or even most important) stressor at a given site without thorough investigation of other candidate stressors.

The Role of Statistical Hypothesis Testing

Statistical procedures are very useful in stressor identification to develop evidence and models of association and to assess strength and credibility of the evidence. Statistical analyses are also useful for assigning scores to the evidence based on qualities such as strength of association using the coefficient of determination (r^2 value), frequency of occurrence, number of occurrences, and so forth. Statistical hypothesis tests address the likelihood of the data given the null (no-effect) hypothesis. They do not test whether the association is causal. Statistical tests are more useful for determining either the logical implication (direction) or strength (for example, r^2 value) of a modeled causal relation. Using r-values rather than r^2 is better because it also defines direction of the relation.

Summary statistics and correlations or regressions are encouraged, with caveats not to draw causal conclusions from statistical tests such as statistically significant correlations because correlation does not necessarily indicate causation. Suter and others (2002) suggested linear or nonlinear regression and (for categorical data) categorical regression to analyze relations. Correlation coefficients, r-values, and probability values can be reported but should not be used to definitively select or eliminate a candidate cause. Rather, they are a measure of qualities of the evidence to infer the logical implication of the evidence by the direction (sign) of the association and the strength by the magnitude of the r-value.

Reporting Level of Confidence

Many analytical tests are expressed in a manner that emphasizes confidence in the results, such as using confidence limits, probabilities, and likelihoods, and can be helpful in providing a level of confidence to the overall assessment. Confidence is also negatively affected by multiple sources of uncertainty regarding analysis and evaluation of candidate stressors and should be documented.

Possible sources of uncertainty include the following:

1. Data quality: age or incompleteness of data, possible measurement errors or biases, specificity of methods, quality of metadata;
2. Use of surrogates, analogies, or extrapolation of data beyond what was originally measured;
3. Masking considerations, including possible ancillary causes that could not be identified because they were masked by a primary cause; and
4. Confounding factors that may cause correlation between two variables to be a poor indicator of causation.

Data Types for Stressor Identification in Stream Ecosystems

A variety of biological, hydrological, physical, and chemical data types can support SI efforts. Biological data include aggregate metrics for biological communities, as well as variables at the level of individuals, population demographics, community structure, and data representing ecological and biological stressors. Water chemistry includes variables representing concentrations of nutrients and DO, potentially toxic organic compounds or metals, water pH, specific conductance, and total dissolved solids, among other indicators. Water physical properties, such as temperature, suspended sediment, and turbidity, can produce direct effects as stressors on biological communities and may also interact with water chemical properties to generate stressful conditions. Physical habitat characteristics include the structure and diversity of habitat types (such as riffles and pools), degree of shading, presence of woody debris, substrate particle size, and temporal stability of habitat. Channel characteristics may include natural geomorphic properties of stream channels (for example, stream size, stream gradient, bed or bank erosion, and density of riparian vegetation) as well as human modifications to stream channels (for example, channelization, dredging, and infrastructure features). Streamflow characteristics include several hydrologic properties reflecting high-flow conditions, low- or base-flow conditions, temporal variability in flow, and frequency and seasonality of particular flow conditions. For a partial list of ecologically relevant streamflow characteristics that may be useful in SI efforts, see Table II in Knight and others (2014). Watershed characteristics, such as land-use type, impervious surface coverage, use of tile drainage, use of agricultural best-management practices, crop types, and locations and numbers of point sources (such as wastewater treatment plants or industrial effluent discharges) may produce direct effects and may regulate many of the previously described variables relevant to SI efforts.

Analytical Approaches for Stressor Identification in Stream Ecosystems

Analysis of State and regional monitoring data can help to determine whether environmental conditions or biological characteristics at a site differ from desired characteristics and to understand the relation between a stressor and a response in a particular region. Many analytical techniques have been used successfully for causal analysis (EPA, 2010a). Several are briefly described here, and case studies and references for each are listed in table 1.

Tests of Significant Difference

Tests of significant difference can be useful in determining whether the value(s) for a variable at a test site are significantly outside the range of values observed at reference sites. These tests evaluate a null hypothesis stating that the observed value(s) at the test site are not significantly different than those at the reference site against an alternative hypothesis that the test site value(s) are significantly different than those of the reference site. Traditionally, these tests use a 0.05 test of significance; however, for the use in stressor identification, it is recommended to consider limits of significance greater than 0.05 because they may indicate the potential for difference, albeit supported by less power of significance (EPA, 2010a).

Because of the assumptions underlying parametric tests (see the Constraints and Limitations subsection), several nonparametric approaches may be helpful:

- The Anderson-Darling test can be used to determine whether a set of observations could be derived from a probability distribution (for example, normal distribution). The K-sample Anderson-Darling test allows testing of whether several sets of observations can be modeled as derived from a single population.
- The Kendall rank correlation coefficient (also known as Kendall's tau) can be used to determine the similarity of ordering of data based on rank variables.
- The Kolmogorov-Smirnov test can be used to compare two samples or to compare a sample to a reference probability distribution, based on distance between an empirical distribution function and a cumulative distribution function.
- The Mann-Whitney U test (or Wilcox test) can be used to test whether two samples derive from the same population; as such, it is a nonparametric alternative to a Student's t-test.

Table 1. Analytical approaches and associated case studies and references for stressor identification.

Analytical approach	Case studies and references
Tests of significant difference	Clements (1994), Townsend and others (2008).
Multiple regression	Clements and others (2000), Norton and others (2000), Townsend and others (2008), Black and others (2011), Vander Laan (2012).
Quantile regression	Cade and Noon (2003), Cade and others (2005, 2008), Cade (2011), Schmidt and others (2012), Knight and others (2014).
Ordination methods	Norton and others (2000), Morris and others (2006), Matthaei and others (2010), U.S. Environmental Protection Agency (2010a), Allan and others (2013), Smucker and others (2013), Palmer (2016).
Discriminant analysis	Clements (1994), Norton and others (2000).
Classification and regression trees	De'ath and Fabricius (2000), Prasad and others (2006), Usio and others (2006), Steen and others (2008).
Random forests	Breiman (2001), Prasad and others (2006), Cutler and others (2007), Maloney and others (2009, 2013), Carlisle and others (2010), Black and others (2011), Vander Laan and others (2013), Liaw and Wiener (2014), Gregorutti and others (2017).
Boosted regression trees	De'ath (2007), Elith and others (2008), Smucker and others (2013).

Uses and Advantages

Where parametric assumptions can be met, analysis of variance (ANOVA) or t-tests can be used to test for significant differences between or among stations or reaches (Clements, 1994; EPA, 2010a). ANOVA can also be used to support SI through controlled experiments (Townsend and others, 2008). Nonparametric methods have the advantage of not requiring normally distributed data; however, certain nonparametric tests may have their own assumptions or constraints.

Constraints and Limitations

Parametric approaches (such as t-tests or ANOVA) assume normality of data distribution, adequate sample size, and random sampling. In some practical cases of SI, these assumptions may not be valid.

Multiple Regression

Linear and nonlinear regression models can be useful in SI to model the relations between one or more candidate stressors and a biological response. If multiple candidate stressors are modeled (as is often the case), the regression analysis is referred to as multiple regression. Regression models can be evaluated based on goodness-of-fit statistics that describe how well the model fits the data, as well as by examining regression coefficients.

Uses and Advantages

Regression analysis allows the evaluation of a functional form between a candidate stressor and biological response. In addition, the difference between observed values and expected values based on regression equations (in other words, the regression model residuals) can be useful in investigating other variables that were not included in the original regression model. Where identification of stressor-response relations is made more difficult because of confounding correlations based on natural variation along gradients, regression residuals may be more appropriate than raw biological data for use in subsequent analysis. For example, because some explanatory and response variables in streams naturally correlate with stream size (such as drainage area), Norton and others (2000) first regressed these variables on drainage area and then used drainage-area residuals in subsequent analysis. Similarly, Vander Laan (2012) defined potential stress as a site-specific difference between an observed level of a physical or chemical variable and its expected value based on natural conditions.

A common issue in stream ecology is that two or more variables produce interactive effects on biological responses. To address this common issue, interaction terms can be fitted into regression models. For example, Townsend and others (2008) fitted interaction terms (such as sediment × nutrients) to model interactive effects on biological response. Another common issue in regression models for stream ecology is the challenge of selecting the appropriate explanatory variables, such that important predictors are not left out but also that the model is not overfitted. Stepwise multiple regression can be used to select explanatory variables, either with forward selection (starting with no variables and systematically adding variables based on a statistical criterion for entry into the model) or backward selection (starting with all candidate variables and removing variables systematically based on a criterion for removal). For example, using forward selection, Clements and others (2000) constructed regression models for 16 biological response variables and found that heavy-metal concentration was the most important predictor, with other physical and chemical characteristics explaining smaller amounts of variation in community structure.

Constraints and Limitations

Linear regression, although simple in application and interpretation, requires an assumption of a constant relation between explanatory and response variables across all values of both variables. In stream ecology, nonlinear relations are common. The appropriateness of a linear regression model can be assessed by a quantile-quantile plot of the model residuals (EPA, 2010a). Other assumptions that should be checked prior to acceptance of linear regression results include assumptions of constant sampling variance and of independent samples (for example, lack of spatial autocorrelation).

Quantile Regression

Quantile regression estimates the conditional quantiles of a biological response variable in a linear model. Quantile regression can be useful in cases where there is a weak or no predictive relation between the mean value of the biological response variable and the explanatory variables, but relations do exist with other parts of the biological response variable distribution (Cade and Noon, 2003; Armstrong and others, 2011) (fig. 4).

Uses and Advantages

Quantile regression is particularly applicable in cases of multiple measured and unmeasured limiting factors (in other words, physical or chemical variables that constrain a biological response such as species richness). In stream ecology, including SI processes, multiple unmeasured limiting factors are commonly present. In such cases, examination of quantile regression parameters for multiple quantile levels can be useful in discerning “ecological limits” set by the explanatory variable in question. Quantile process plots show the slope and intercept parameter estimates for all quantile levels and can include confidence intervals (for example, 95-percent confidence) that can be compared to zero to test whether (at the given level of confidence) the parameter estimate is different from zero.

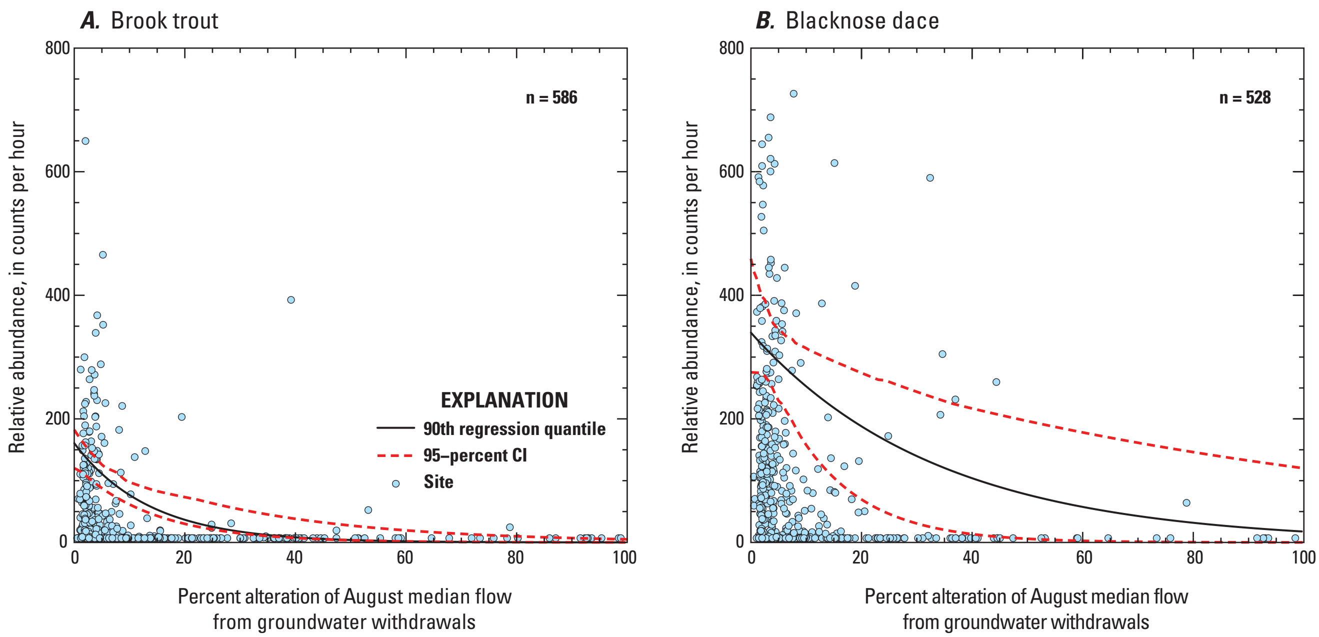


Figure 4. Quantile regression relations between relative abundance of *A*, brook trout and *B*, blacknose dace and percent alteration of August median flow from groundwater withdrawals (from Armstrong and others, 2011). [CI, confidence interval; n, number of sites]

Constraints and Limitations

If the number of model parameters is large relative to the sample size, parameter estimate precision may be poor (Cade and others, 2008).

Ordination Methods

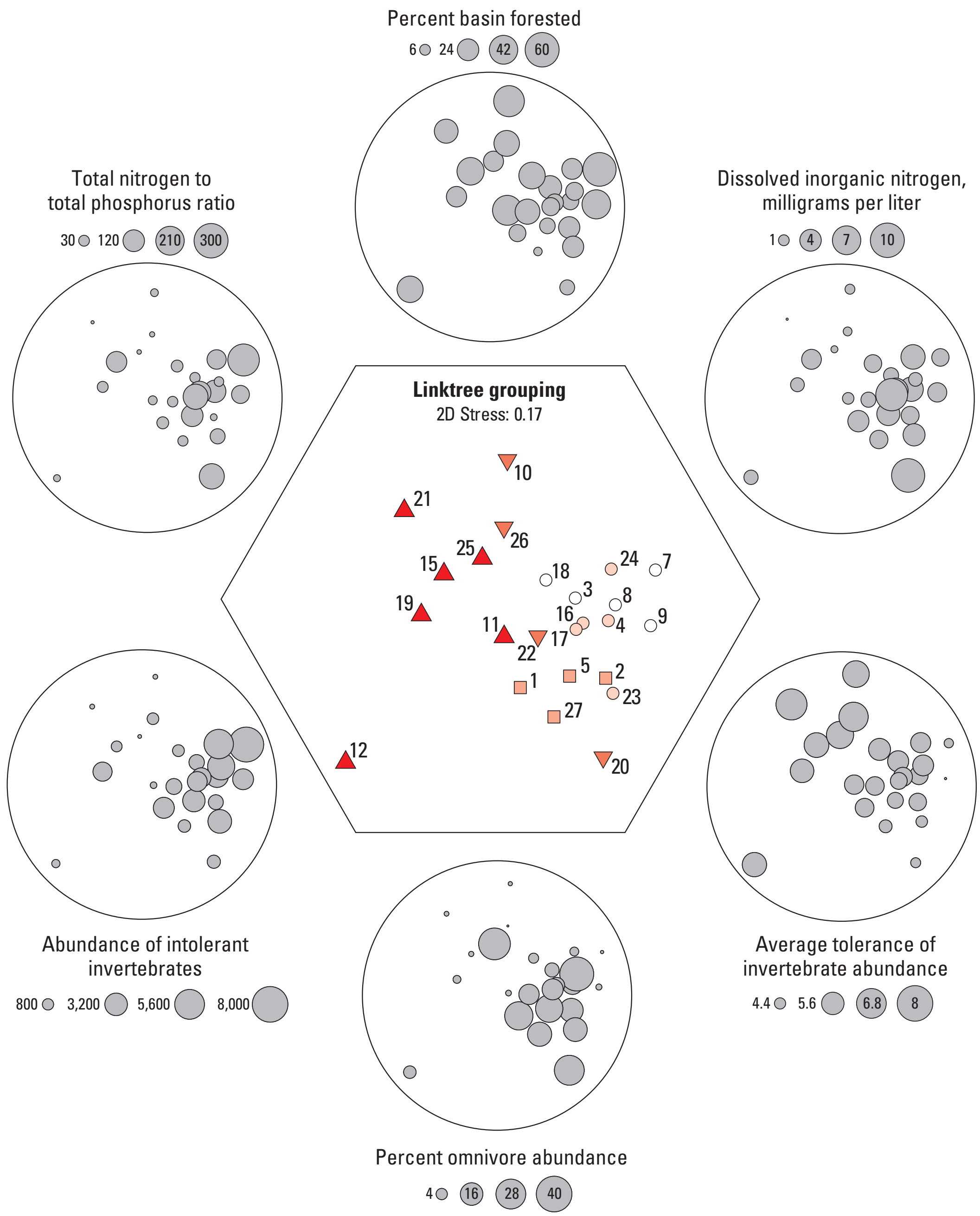
Ordination methods are a family of statistical procedures that share a common goal: to synthesize highly multidimensional data into a smaller number of dimensions. This allows data to be plotted in two-dimensional ordination space (fig. 5), such that similar observations are plotted closer together and dissimilar observations are plotted farther apart. Commonly used ordination methods include principal component analysis, cluster analysis, principal factor analysis, detrended correspondence analysis, and nonmetric multidimensional scaling.

Uses and Advantages

Ordination methods are useful in finding patterns in datasets that are otherwise too complicated (in other words, too highly dimensional) to interpret. Environmental variables may be highly correlated or “redundant” with one another. For example, soil pH, calcium, magnesium, and cation exchange capacity are usually very tightly correlated. If so, any one of these variables could potentially be used as a proxy for all the others. Alternatively, an ordination method such as principal components analysis could be used to distill and summarize the dataset, such that a principal component might represent the combined effect of soil pH, calcium, magnesium, and cation exchange capacity, rather than representing each variable individually.

Constraints and Limitations

Ordination methods are generally considered exploratory tools and are less well suited to statistical hypothesis testing, meaning that they do not support rejection of a null hypothesis that no significant relations exist in a dataset. When applying ordination approaches to community composition data along a gradient (for example, species counts along a DO gradient), the beta diversity (how different the community composition is among samples along the gradient) should be considered in selecting the appropriate method. For example, principal components analysis functions best with low beta diversity, whereas detrended correspondence analysis and canonical correlation analysis perform better with high beta diversity (Palmer, 2016).



EXPLANATION FOR LINKTREE GROUPING

- ▲ **A** Wooded vegetation in 25-meter-reach buffer < 80%
- ▼ **B** Crops in 50-meter-segment buffer > 15%
- ■ **C** Dissolved organic carbon > 4 mg/L
- ● **D** Dissolved inorganic nitrogen > 4.2 mg/L
- ○ **E** Dissolved inorganic nitrogen < 3.2 mg/L

Figure 5. Plot series illustrating nonmetric multidimensional scaling (MDS) ordinations of Delmarva Peninsula algal community composition from Bray-Curtis similarity matrices (from Calhoun and others, 2008). Top graphs are environmental variables, and bottom graphs are autecological algal indicators. [mg/L, milligram per liter; >, greater than; %, percent; <, less than]

Discriminant Analysis

Discriminant analysis (DA) describes statistical methods for predictive models to classify observations into a set of classes based on a set of continuous response variables. The dataset used for DA includes a set of observations with known classification and values for the continuous variables. For example, a set of N sites that are already classified as "good," "medium," and "poor" for water quality, and which have data on Y continuous variables (such as nitrate, temperature, DO, and turbidity), could be a dataset used for DA. The discriminant model would use the Y continuous variables to assign each water-quality site as being "good," "medium," or "poor." This output is used to generate a misclassification summary that provides an indicator of the goodness-of-fit of the model. The model can then be used predictively to classify future sites as "good," "medium," or "poor" based on their values for the Y continuous variables.

Uses and Advantages

DA can be used for predictive classification. In other words, DA can be used to assign "new" observations to predetermined categories based on continuous variables. DA can also detect which variables are most useful for differentiating between or among categories.

Constraints and Limitations

Linear DA assumes that the independent variables are normally distributed and that the variance/covariance matrices of variables are homogenous across classes. In addition, significance tests for linear DA may not be valid if the means for variables across classes are correlated with the variances.

Classification and Regression Trees

Classification and regression tree (CART) analysis is a machine-learning technique that involves recursive splitting of a dataset into progressively smaller and more homogenous groups by using combinations of explanatory variables. The output is a tree diagram with branches determined by the splitting rules derived from the input data (fig. 6).

Uses and Advantages

CART and ensemble tree-based methods can model nonlinear relations and interactions between explanatory variables. These approaches do not assume normality of data distributions and are able to handle missing values both for response and explanatory variables. In addition, CART and ensemble tree-based methods can be performed on a variety of data types, including continuous, binomial, and categorical data. Unlike ensemble tree-based methods, CART models produce a single decision tree, which makes results relatively easy to interpret. As such, CART models may be more "user friendly" to nonstatisticians than ensemble methods. For example, Usio and others (2006) used CART to identify thresholds of pH and water temperature (6.5 and 14.3 degrees Celsius [°C], respectively) associated with stream habitat for an invasive crayfish species in Japan. Similarly, Steen and others (2008) identified a water-temperature threshold of 18.66 °C associated with brown bullhead (*Ameiurus nebulosus*) habitat in Michigan streams. In each case, the decision trees provide clearly organized information on the relative importance of predictor variables in explaining patterns of the dependent variable, along with thresholds on predictor variables and an overall metric of the prediction accuracy of the CART model. CART and similar tree-based methods could be useful in the SI process for demonstrating evidence for or against several characteristics of causation, including co-occurrence and sufficiency. For example, if a CART analysis demonstrates a threshold response for a particular candidate stressor, and that stressor co-occurs at the study site at levels well exceeding that threshold, then the CART analysis could be used to help illustrate sufficiency of the candidate cause.

Constraints and Limitations

Depending on the size of the dataset, and the number and nature of the explanatory variables, individual trees from CART analysis may be unstable. For example, if a small subset of the original dataset is randomly selected and withheld from analysis, the resulting tree may have a different first split (indicating a different explanatory variable as being of primary importance) than if that subset was included. To deal with this issue of instability, ensemble tree-based methods were developed as described in the report sections Random Forests and Boosted Regression Trees.

Random Forests

Random forests (RF) is an ensemble tree-based method, meaning that it fits many (classification or regression) decision trees to an input dataset, unlike CART which produces only a single tree for each analysis run. The use of many "trees" produced the "forests" part of the method name. RF uses bootstrap aggregation (bagging); in other words, uniform samples are taken from the input dataset and used to create trees. Each decision tree is grown using a randomized subset of explanatory variables.

Useful outputs of RF include variable-importance plots and partial-dependence plots. Variable importance represents the magnitude of contribution of each of the explanatory variables to the recursive split algorithm (in other words, an estimate of how much worse the prediction would be for the dataset if that explanatory variable were removed from analysis). Partial dependence represents the relation of an individual explanatory variable with the response variable if all other explanatory variables were held constant.

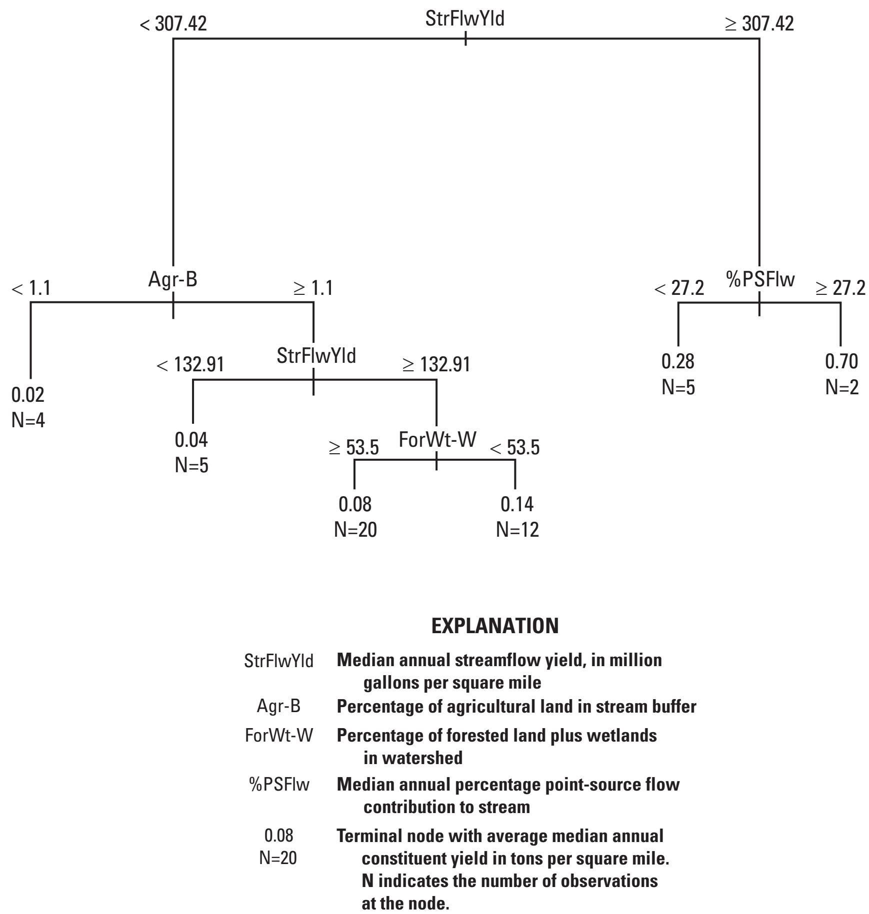


Figure 6. Example of a regression tree identifying predictor variables that best explain observed variations in median annual yields of total phosphorus (from Harden and others, 2013).

Uses and Advantages

RF can be used for classification and for regression. When used for classification, RF classification accuracy is generally high. Like other tree-based methods, RF can model nonlinear relations and interactions among explanatory variables and is not adversely affected by missing values. Because it is an ensemble method, meaning that it averages results across many trees, results are generally more stable than from single-tree approaches such as CART. A common application of RF methods is to produce species-distribution models that predict occurrence (presence and absence) for species of interest based on a suite of habitat metrics. For example, Maloney and others (2013) used conditional RF models to predict the probability of presence for 30 common fish species in streams of the Chesapeake Bay watershed. Such predictive models can then be used to generate predictive maps for types of information such as species occurrence or stream biological condition for locations where these variables have not been measured directly. For example, using a conditional random forest model, Maloney and others (2009) predicted that approximately 34 percent of small nontidal stream reaches in the Maryland portion of the Chesapeake Bay watershed were in “fair” condition, with 30, 23, and 13 percent in “good,” “poor,” and “very poor” biological condition, respectively. Similar to CART and other tree-based approaches, RF could be useful in the SI process for demonstrating evidence for or against co-occurrence and (or) sufficiency of candidate causes. For example, an RF model could be used to predict that a site with a given set of physical parameters, such as water-quality and habitat metrics, would be biologically impaired. If a study site has those physical parameters in the ranges specified in the RF model and also demonstrates biological impairment, this could constitute evidence in favor of the candidate cause(s) included in the RF model.

Constraints and Limitations

When used for regression, RF cannot predict beyond the range in the training dataset. Additionally, RF may result in overfitting, especially when applied to datasets with a low signal-to-noise ratio. Strong correlation among explanatory variables may produce bias in the variable-importance plots produced in RF, such that variable importance is inflated for highly correlated variables. Bias in variable importance due to correlated explanatory variables may be addressed by using recursive feature elimination. RF and other ensemble tree methods have been referred to as “black box” approaches because interpretation of their results is less straightforward than for single decision trees, such as CART.

Boosted Regression Trees

Like RF, boosted regression trees (BRT) is an ensemble tree-based method which fits many decision trees to an input dataset, using recursive partitioning based on explanatory variables. In BRT, trees are created to handle residuals leftover from previously constructed trees; in other words, each tree attempts to explain remaining variation in data that was not explained by previous trees. As each new tree is built, the data are re-weighted to emphasize cases that were poorly predicted by previous trees.

BRT produces variable-importance plots and partial-dependence plots that are similar to those from RF.

Uses and Advantages

Like other tree-based methods, BRT can model nonlinear relations and interactions among explanatory variables and is not adversely affected by missing values. BRTs are generally robust to extreme outliers and to the inclusion of irrelevant explanatory variables. Because it is an ensemble method, meaning that it averages results across many trees, results are generally more stable than those from single-tree approaches, such as CART. De’ath (2007) asserted that BRTs “consistently outperform [RF], particularly for regression.” Similar to both CART and RF, BRTs could be useful for evaluating the co-occurrence and sufficiency characteristics of causation.

Constraints and Limitations

Like other ensemble tree approaches (for example, RF), BRT has been referred to as a “black box” approach because interpretation of results is less straightforward than for single decision trees, such as CART.

Stressor Identification of Biologically Impaired Streams in Mississippi

Biological Impairment in Mississippi: Data Collection and Assessment

MDEQ annually collects benthic macroinvertebrate samples from streams statewide as part of their ambient surface-water monitoring program. Biological samples are processed by using taxonomic identification and counting of individual macroinvertebrates to produce biological metrics, which are then used to assess the health of wadable streams and to identify impaired streams based on differences in taxa richness and percent composition, feeding and habitat richness, and composition and tolerance traits. This statewide monitoring also involves assessment of habitat condition, sediment particle size, water-quality characteristics, and remotely sensed land-use and land-cover percentages. Calculation of an M-BISQ value for each sampling site involves comparing selected biological metrics for the site to their respective impairment-threshold values. The biological metrics used, and the stream reaches chosen to represent the impairment threshold, vary according to bioregions of the State. Bioregions represent geographically distinct natural assemblages of benthic macroinvertebrates (Stribling and others, 2016) and were determined by modifying level IV ecoregions (Omernik, 1987) with biological data to produce bioregions with relative biological homogeneity (fig. 7). For each bioregion, the impairment threshold is calculated as the 25th percentile of the M-BISQ scores for a set of streams considered to be least disturbed by human activities. If the M-BISQ score for a sampling site falls below the impairment threshold for its bioregion, the site and its associated stream reach are listed as biologically impaired. Biological sample collection, processing, and subsequent data analysis procedures are explained in detail in the Quality Assurance Project Plan for the §303(d) List Assessment and Calibration of the Index of Biological Integrity for Wadable Streams in Mississippi (Stribling and others, 2016).

The M-BISQ assessment process identifies biologically impaired streams but does not seek to explain the causes of impairment. For many stream reaches, the causes of impairment are unknown, thus requiring an SI process to clarify the specific potential causes of impairment. Stream reaches that are listed as impaired based on an M-BISQ score are generally similar in the type, amount, and nature of data available for SI. Therefore, a standardized process for SI in these impaired reaches of Mississippi was developed by adapting the EPA framework for stressor identification (EPA, 2000; Cormier and others, 2003). The goals of the Mississippi SI framework are to (1) provide a common guide for users to maintain consistency and transparency in assessment and documentation for impaired streams, and (2) serve as a foundation for sustainable and adaptive management of aquatic resources.

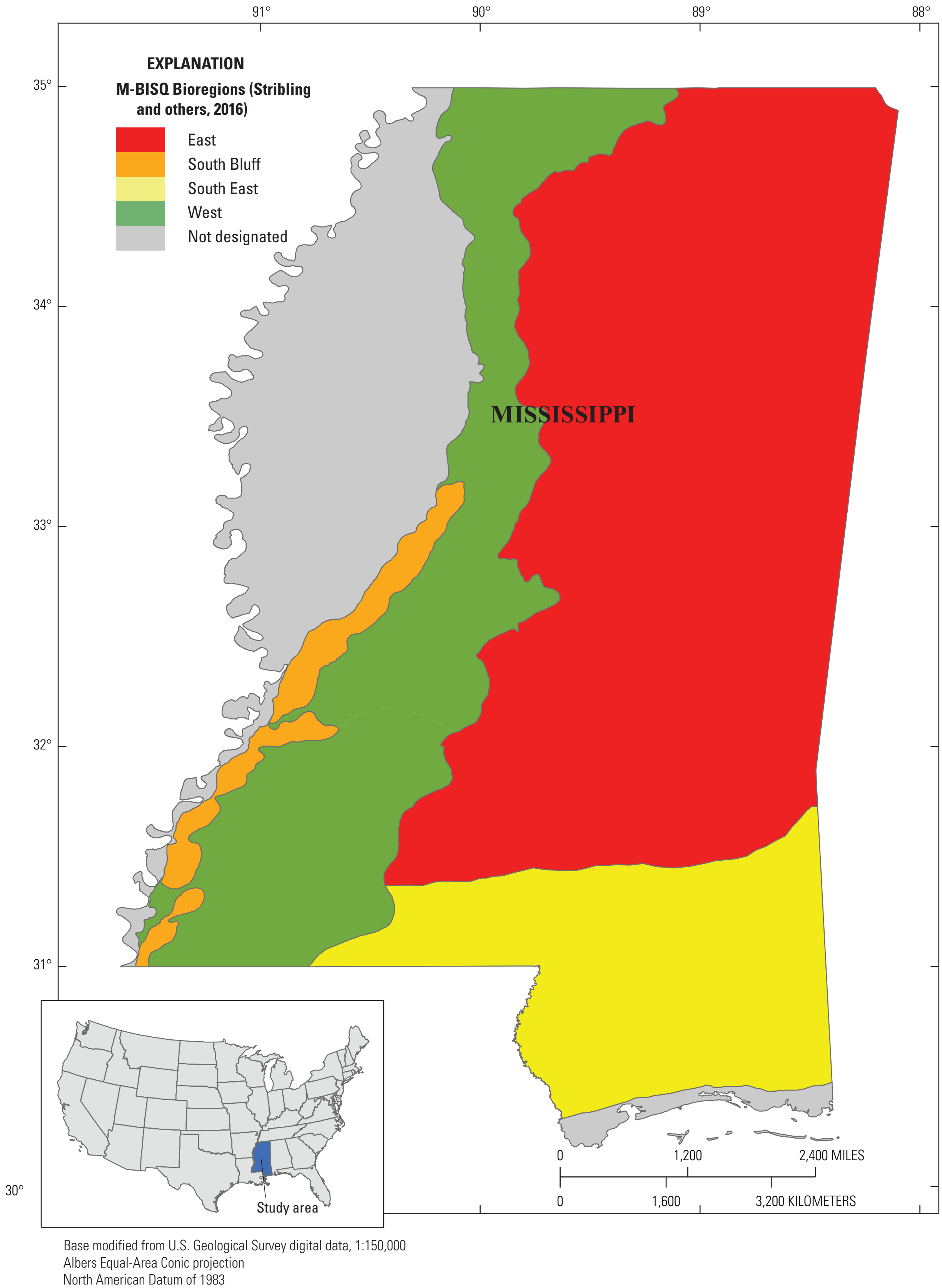


Figure 7. Mississippi Benthic Index of Stream Quality (M-BISQ) bioregions used to stratify analysis of ecological response to stressors in biologically impaired streams in Mississippi.

Components of the Stressor Identification Framework for Biologically Impaired Streams in Mississippi

The Mississippi stressor identification framework follows six general steps:

1. Define the impairment,
2. List the candidate causes of impairment and develop a conceptual model,
3. Compile all data relevant to the impairment and conceptual model,
4. Evaluate the data,
5. Identify the probable causes of impairment by using a WOE approach, and
6. Generate a report of results.

Step 1. Define the Impairment

The first step is to define the geographic and temporal scope of the biological impairment for the impaired stream reach. Basic information includes the reach name, a general location description, bioregion (including geographic details of any site spanning multiple bioregions), date the original biological sample was collected that resulted in the listing for biological impairment, dates of any subsequent biological sampling, drainage area, dominant current land uses and land cover, and historical land uses and landscape changes. Defining impairment for a given stream reach requires documentation of the biological metrics from the sample that was used to list the site as impaired, such as low numbers of sensitive taxa, low overall richness of taxa, or low diversity of functional feeding groups. To document the impairment, each biological metric value for the site should be listed along with the corresponding impairment-threshold value for the bioregion in which the impaired site is located. Other relevant biological information can be compiled in this step, such as M-BISQ scores and metrics from samples obtained at the same location or within the same watershed on different dates, or aquatic community data for taxonomic groups other than benthic macroinvertebrates. Additional documentation to define the impairment may come from field notes, information obtained during reconnaissance of the drainage area and site locations, §303(d) listing information, and remote-sensing data. In rare cases, stream reaches may have additional information that predates the use of M-BISQ to confirm impairment, such as TMDLs developed and actions taken to remediate the impairment. In addition, data from other local, State, Federal, nonprofit, and private sector entities may be considered for use.

Step 2. List the Candidate Causes of Impairment and Develop a Conceptual Model

This step includes the development of a list of candidate causes of stress (stressors) and a conceptual model to explain the potential linkages between stressors and the nature of biological impairment documented in step 1. The conceptual model can also suggest linkages among stressors and potential sources of stress, which can include a variety of human activities and artifacts within the watershed. A template list of common potential sources of stress (table 2) may be used as a guide and may be modified or expanded based on a site's individual history or landscape context. Some potential sources of stress may be identified or eliminated based on land-use and land-cover characteristics of the watershed and field reconnaissance information. Geospatial analysis of mapped locations of point sources (for example, wastewater effluent discharge sites), septic tanks, and miscellaneous sources of contaminants may also help identify or eliminate potential sources of stress.

Once a list of potential sources of stress within the watershed has been compiled, associated proximate and intermediate causes can be identified. Proximate causes are defined as the immediate and effective causes of the biological impairment. MDEQ considers (Shawn Clark, MDEQ, personal communication, September 4, 2017) five main categories of proximate causes of impairment:

1. Decrease in suitable habitat,
2. Alteration to thermal regime,
3. Decrease in DO due to nutrient enrichment,
4. Decrease in DO due to organic enrichment, and
5. Increased toxicity, including ionic strength and other contaminants.

Intermediate causes are part of the causal pathway but do not directly cause the impairment. For example, sources of stress within a watershed, including water withdrawals, urban areas, agriculture, and channel alterations, could produce intermediate causes such as flow-regime alterations, changes in channel morphology, decreases in riparian canopy cover and instream habitat, and increases in solar input, suspended sediment, and erosion (fig. 8). These intermediate causes do not produce the biological impairment in and of themselves, but rather contribute to the proximate causes of impairment which could include decreased suitable habitat, altered thermal regime, and increased toxicity (fig. 9).

Table 2. Common potential sources of stress in biologically impaired streams.

Source of stress	Examples of more specific sources	References for example studies
Resource extraction	Sand and gravel mine Oil field Coal mine Water withdrawal	Clements (1994), Clements and others (2000), Carlisle and others (2011), Merriam and others (2011), Bernhardt and others (2012), Lavoie and others (2018).
Silviculture	Clear cutting Specific trees on the farm	Allan and others (2012).
Urbanization	Urban construction Dumping Fertilizer application Impervious surfaces	Onorato and others (1998), Wang and others (2000), Walsh and others (2005), DeGasperi and others (2009), Miguntanna and others (2010), Merriam and others (2011), Porter and others (2015), Maryland Department of the Environment (2016), Lavoie and others (2018).
Wastewater discharge	Specific size or permitted discharge pollutants	Blazer and others (2012), Lavoie and others (2018).
Unsewered residential	Septic tanks Landfill	Lapointe and others (2012).
Agriculture	Specific crops Specific animal operations	Cormier and others (2000), Norton and others (2000), Matthaei and others (2010), Black and others (2011), Blazer and others (2012), Lange and others (2014), Maryland Department of the Environment (2016), Pearson and others (2016), Munn and others (2018).
Channel alteration	Channelization Dredging Ditch construction Road crossings Bank erosion	Cormier and others (2000), Carlisle and others (2009), Doll (2011), Allan and others (2012), Voss and others (2012), Porter and others (2015).
Impoundments	Dams and other barriers	Ohio Environmental Protection Agency (2010), Carlisle and others (2011).
Roads	Unpaved roads Paved roads, highways	None.

A conceptual model can then be developed to show linkages between potential sources of stress, intermediate causes/pathways, proximate stressors, and biological responses (fig. 10). The conceptual model represents the plausible relations among possible sources of stress and the biological response variables, including the intermediate pathways. A conceptual model can consist of a single diagram or multiple diagrams that each represent a proximate cause of impairment. The process of conceptual model development may help eliminate some candidate stressors because of the lack of a source and (or) pathway. For example, a candidate stressor might be eliminated if a causal agent or source of stress were believed to be either mechanistically implausible or absent from the watershed in question.

Lists of causes and templates in figures 8 through 10 can help guide assessments but should be tailored to site-specific watershed histories and characteristics as needed. In some cases, sources of impairment may be present in a watershed that are not included in the preliminary list. These may be revealed by field investigations, geospatial analysis of the watershed, or other sources of information. Such additional candidate sources of impairment should be evaluated in the relevant scientific literature to determine the causal pathway linking intermediate causes to proximate causes to the response indicator. Although it is practical to consider intermediate causes that are commonly measured and available, caution should be taken by also considering variables that may not have been measured. For example, pesticides, invasive species, pathogens, endocrine disruptors, pharmaceuticals, and polycyclic aromatic hydrocarbons are not routinely measured, but may be important in the actual causal-chain scenario. It is important to identify all the potential sources of stress and the potential causes of stress that are associated, because even though the data may not be available, these sources and causes may need to be explored. Future data collection, analyses, and refinement of linkages may allow for more specific, direct, and comprehensive proximate causes to be developed over time.

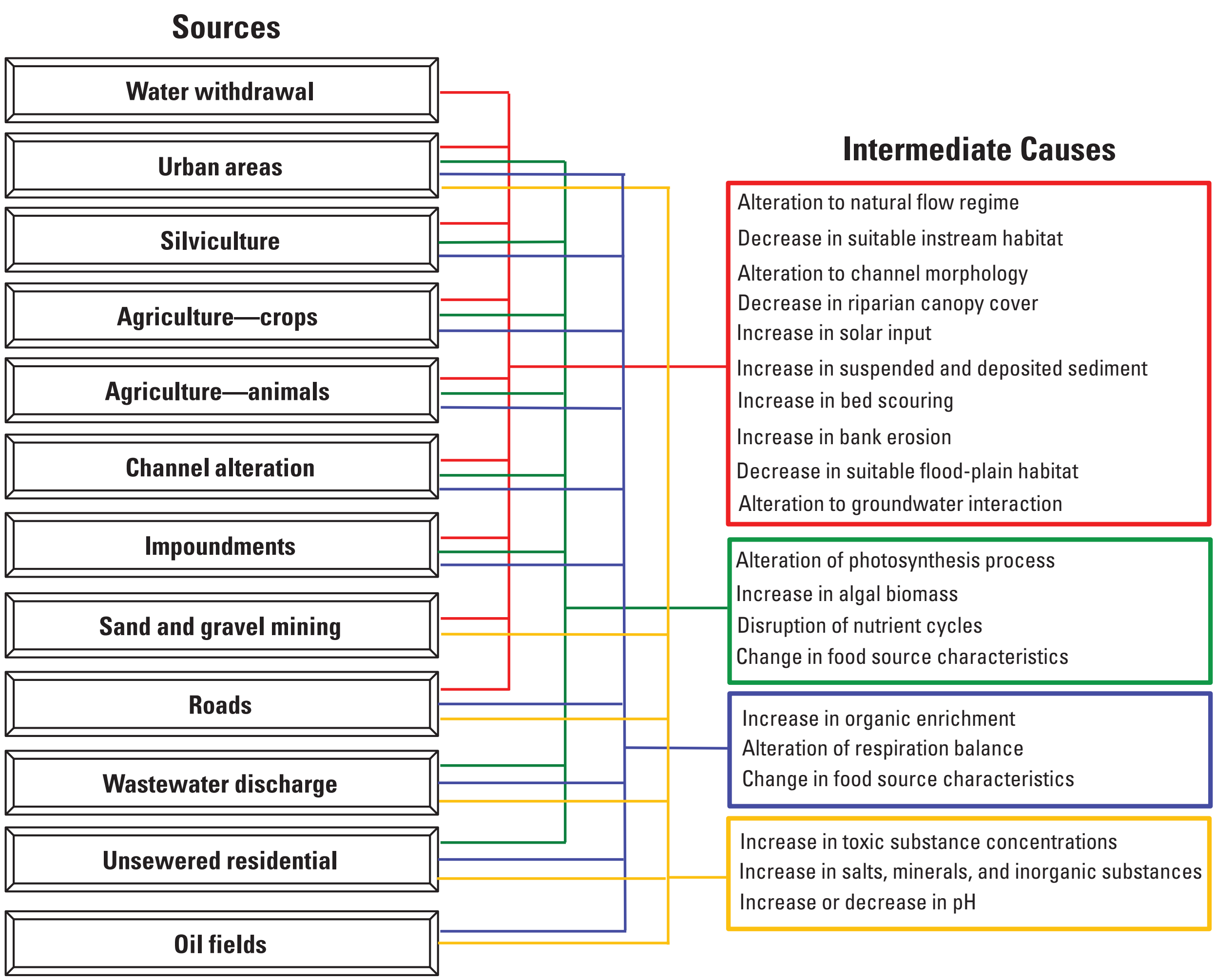


Figure 8. Examples of sources of stress and linkages to intermediate causes of impairment.

Step 3. Compile All Data Relevant to the Impairment and Conceptual Model

By using the conceptual model as a guide, data compilation should focus mainly on data that may characterize proximate and intermediate causes linked to the potential sources of stress identified within the watershed.

Sources of Data

A reconnaissance survey of the stream reach and its watershed is typically conducted as part of this step to begin assembling data relevant to causes of impairment. This survey will provide a rapid geomorphic assessment of the stream reach, field measurements of water-quality and physical characteristics of the stream, and observations from upland locations throughout the watershed. Examples of the field forms that can be used for such surveys are given in appendix 1.

Following reconnaissance, all available water-quality, biological, physical, and hydrologic data available for the stream reach, watershed, and bioregion need to be gathered and organized. Data sources include information collected during preliminary site visits, MDEQ databases, the EPA Legacy Storage and Retrieval (STORET) database, USGS databases, land-cover geospatial datasets, and other government agency databases and scientific literature. Supporting information about the data is also compiled at this stage, including data sources, data collection dates and locations, data collection methods, and quality assurance protocols.

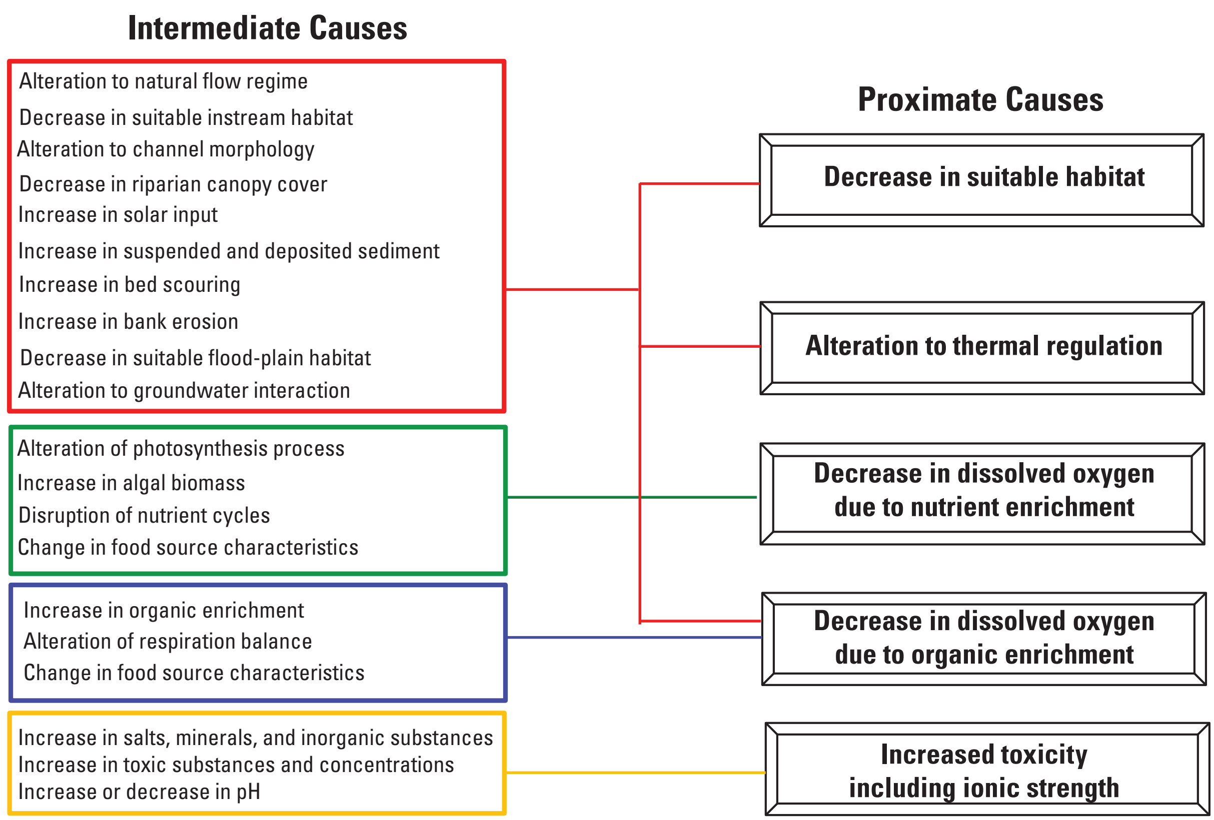


Figure 9. Examples of intermediate causes and linkages to proximate causes of impairment.

M-BISQ Data and Organization

Common data collected by MDEQ during biological stream monitoring that may be compiled in this step include stressor indicators of habitat, substrate particle size, physical and chemical water-quality indicators (table 3); response indicators of benthic macroinvertebrate community metrics; and taxonomic information. In addition to data from the impaired site, statewide data from a master database of stations monitored during 2001–14 may be used to compare the impaired site to other local sites that are not biologically impaired, termed "site-specific comparators" (SSCs), and a range of conditions from least disturbed (LD) and nonimpaired (NI) stream reaches. The statewide master database includes all physical, chemical, biological, and watershed data collected during 2001–14 and is used to define LD and NI conditions and to select SSCs. LD and NI conditions may be defined as the interquartile (25th to 75th percentile) ranges of LD and NI streams within the same bioregion. Appendix 2 contains details on the selection of SSCs.

Step 4. Evaluate the Data

The goal of this step is to perform quantitative and qualitative analyses to determine strengths of association between candidate stressors and observed biological impairments. Evaluation of data and information is performed to support a WOE approach that considers all relevant evidence from diverse sources. These sources include those described in step 3, in addition to interpretation of observations, summary statistics, and statistical and mathematical modeling, and weighing of evidence based on categorization of the types of data. CADDIS includes a scoring system that can be used to summarize the degree to which each type of available information strengthens or weakens the case for a candidate cause. The specifics of the scoring system for the Mississippi framework, a modification of the CADDIS scoring system, are detailed in step 5. In step 4, it is important to document (1) the data analysis methods used, (2) the scoring results, and (3) a summary of the ways in which each type of available evidence either strengthens or weakens the case for a candidate cause.

The objective of step 4 is to evaluate whether fundamental characteristics of a causal relation are present by using various types of evidence (table 4) (EPA, 2000; Suter and others, 2002; Cormier and others, 2003). Confidence in conclusions generally increases as more types of evidence are evaluated.

In most cases, data available for evaluation are those collected as part of the MDEQ statewide biological monitoring program. Therefore, the types of evidence typically available in practice relate to three primary types of causal inference: co-occurrence, antecedence, and sufficiency.

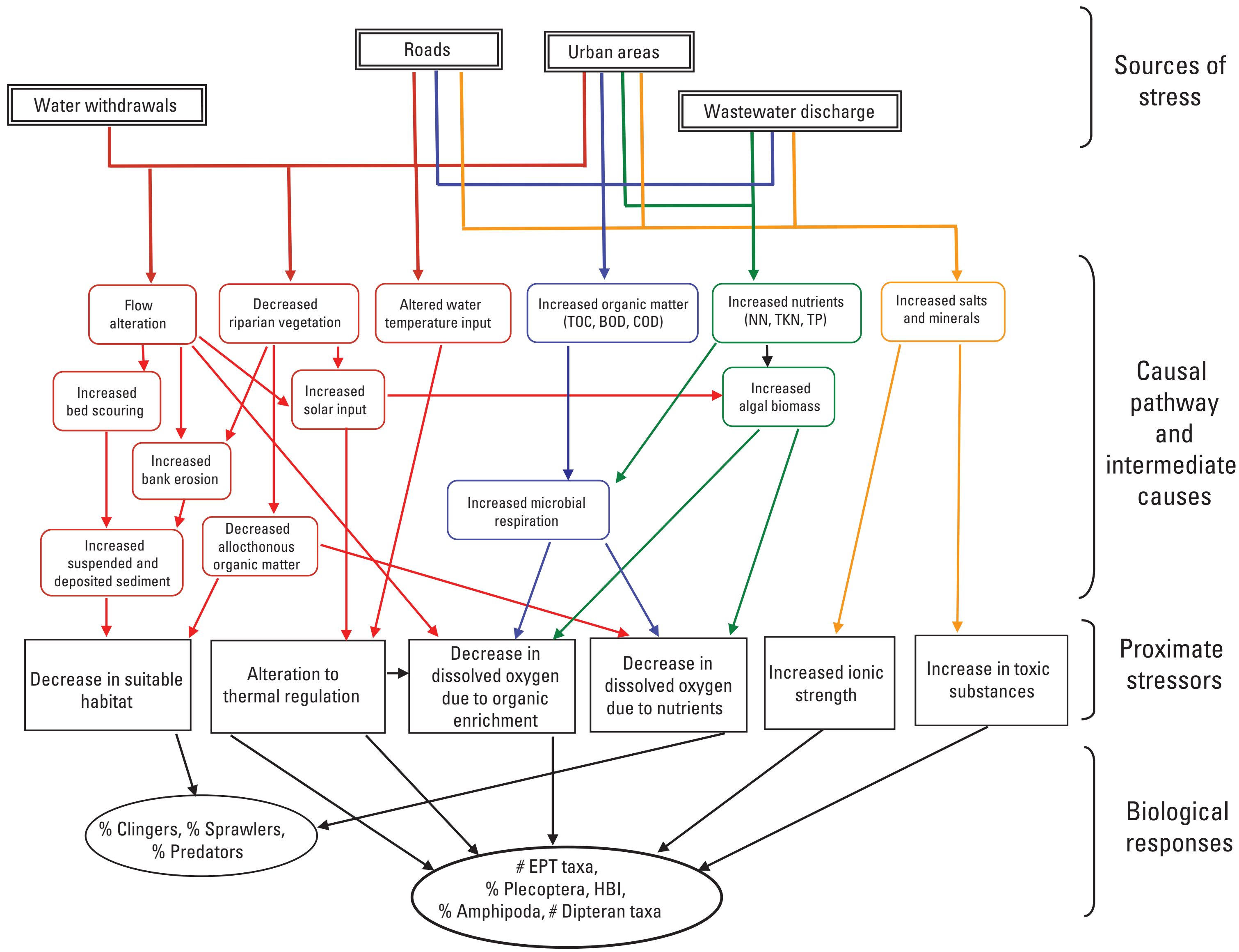


Figure 10. An example of a conceptual model linking potential sources of stress within a watershed to intermediate causes and proximate stressors which produce observed biological responses. [TOC, total organic carbon; BOD, biological oxygen demand; COD, chemical oxygen demand; NN, Nitrate + Nitrite; TKN, total Kjeldahl nitrogen; TP, total phosphorus; %, percent; #, number; EPT, Ephemeroptera, Plecoptera, and Trichoptera; HBI, Hilsenhof's Biotic Index]

Table 3. Examples of common physical and chemical water-quality indicators of proximate causes and the typical availability of the indicator.

[RBP, rapid bioassessment protocol; Q, streamflow; 7Q10, lowest 7-day average flow that occurs (on average) once every 10 years; N, nitrogen]

Proximate cause	Indicator	Availability
Decrease in suitable habitat	RBP habitat assessment	usually
	Quantified habitat measures	rarely (site specific)
	Turbidity	rarely (site specific)
	Suspended-sediment concentration	rarely (site specific)
	Particle-size distribution	usually
	Continuous Q	rarely (site specific)
	Regional ratings	7Q10
Alteration to thermal regime	Temperature	usually (not seasonal or diel)
Decrease in dissolved oxygen due to nutrient enrichment	Dissolved oxygen	usually (not diel)
	Total nitrogen	usually
	Ammonia	usually
	Inorganic N	usually
	Total phosphorus	usually
	Dissolved inorganic phosphorus	rarely (site specific)
Decrease in dissolved oxygen due to organic enrichment	Dissolved oxygen	usually (not diel)
	Total organic carbon	usually
	Chemical oxygen demand	usually
	Biological oxygen demand	rarely (site specific)
Increased toxicity including ionic strength	Specific conductance	usually
	Total chlorides	usually
	pH	usually
	Alkalinity	usually
	Herbicides	rarely (site specific)
	Insecticides	rarely (site specific)
	Metals	rarely (site specific)

Table 4. Examples of types of evidence based on data from within a case.

Causal inference	Type of evidence	General concept
Co-occurrence	Spatial/temporal co-occurrence	The biological effect is observed at the same time and place as the cause, or the stressor is greater compared to other places where the effect is not observed.
Time order	Temporal sequence	The cause precedes the biological effect in the same location.
Antecedence	Causal pathway	A pathway exists from the potential source of stress to the affected organism, their food source, or some entity that can affect them.
Interaction	Evidence of exposure	Measurements of the biota exhibit signs that the causal agent has occurred and the biota was exposed.
Alteration interaction	Laboratory tests of site media	Laboratory tests using media (such as water) from the site in a controlled environment to mimic effects observed in the field.
Sufficiency	Stressor-response relations from the site	The candidate cause from the impaired site is of sufficient magnitude to cause the observed biological effects based on an association of nearby sites (typically a strong r-value on a regression plot).

Co-occurrence and Antecedence

When biological impairment is identified by an M-BISQ score below the impairment threshold, evidence for causal inference may take the form of co-occurrence or antecedence. Co-occurrence is a type of evidence wherein the biological effect is observed at the same time and place as the cause and is not observed when and where the cause is absent. Antecedence is a form of evidence that suggests that the cause and effect are likely to have co-occurred, such as when evidence of the stressor (or a surrogate for the stressor) exists, but not necessarily at the same time as the observed biological effect.

The original process of documenting the observed biological effect—leading to an M-BISQ score below the impairment threshold—may have produced helpful information such as individual community metrics and taxonomic data associated with a candidate. Co-occurrence considers indicators of proximate causes collected at the same time and place as the biological sample. A cause is considered present based on MDEQ water-quality standards, LD/NI conditions, and SSC conditions. Figure 11 is an example of a comparison report used to document these results. Antecedence may consider information gathered during a reconnaissance visit, from point-source discharge records, or from other data from the site or watershed not necessarily collected at the same time as the biological sample.

Sufficiency

Evidence for a candidate cause may also derive from comparison of measures of exposure to a stressor at a study site with lab results or field studies from other sites that demonstrate biological effects of the exposure. If lab tests, field tests, or observational studies from other sites indicate exposure-response relations between a candidate stressor and a biological outcome, an important consideration is whether the stressor is present at the study site in sufficient quantity and frequency for the anticipated result to occur. An additional consideration is that lab conditions and organisms may not accurately represent field conditions. If no lab or field data exist for a candidate cause, analogous or surrogate data may be used, although use of a surrogate will increase the uncertainty of the association.

Typically, when biological impairment is identified by an M-BISQ score below the impairment threshold, analysis of data from other places takes the form of scatterplots of regional data showing relations between specific stressors (causes) and biological community metrics (effects).

Scatterplots depict biological community metrics (typically on the vertical y-axis) plotted against abiotic variables (candidate causal physical/chemical indicators, typically on the horizontal x-axis), with points representing pairs of observed variables from throughout a given region (fig. 12). A regression line—such as from quantile regression—may also be plotted to show the overall regional relation between the candidate stressor and the biological metric. The study site is plotted as a single point on the scatterplot, showing how the study site relates to other sites in the region. Summary statistics from reference sites throughout the region can be added as well, for evaluating how the study site compares to reference sites in terms of the candidate stressor and biological metric. Following similar thresholds used by Knight and others (2014), the interquartile ranges (values between the 25th and 75th percentiles) of the candidate stressor and biological metric for reference sites in a study region can be used to establish reference profiles for the region; in other words, ranges of values for the candidate stressor and biological response are considered generally similar to values observed at reference sites. We then propose that each scatterplot can be scored based on the location of the study site on the scatterplot, with the following scoring options: *strong evidence (++), some evidence (+), incompatible (–), and uncertain (0)*.

Figure 12 illustrates the scoring process. In this example, there is a negative relation at an ecoregional scale between specific conductance and the number of Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa observed. The upper (90th quantile) limit of this relation is illustrated by quantile regression. For any specific conductance value (X value), the line depicting the 90th quantile shows the Y value for which 90 percent of sites in the region have equal or lower Y values. For example, at a specific conductance value of 100 micromhos per centimeter, approximately 90 percent of sites have 12 or fewer EPT taxa. In the example scatterplot (fig. 12), the 25th percentile of the biological metric across reference sites is eight EPT taxa, meaning that 75 percent of reference sites have at least eight EPT taxa. Thus, any site that plots above that horizontal line (EPT taxa=8; in other words, in the upper left or upper right portions of the plot) could be considered to have the biological metric largely within the reference range observed for the region. The example scatterplot also includes a vertical line that represents the 75th percentile of specific conductance (the candidate stressor in this case). Any site that plots to the left of this vertical line has specific conductance less than this value and thus within the general range of reference sites of the region. A study site that plots in this left region of the plot therefore is unlikely to have specific conductance as a strong candidate stressor, because specific conductance values are actually observed to be generally within the reference range for the region. An additional vertical line is plotted based on the intersection point between the 90th quantile regression (diagonal) line and the 25th percentile of the biological metric (horizonal line). In figure 12, this vertical line represents a specific conductance value at which 90 percent of sites in the region have eight or fewer EPT taxa, in other words, fewer EPT taxa than the lower bound of the reference range for EPT taxa. Any study site that plots in the lower right portion of the scatterplot (in figure 12, having fewer EPT taxa than the reference range and having specific conductance greater than the value at which 90 percent of sites have fewer EPT taxa than the reference range) would be considered to have relatively strong evidence in favor of the candidate cause.

A

Parameter	Stressed site	Difference from LD condition		Least disturbed condition		Percentile used for LD
		East Bioregion LD	East nonimpaired	East LD	East nonimpaired	
Chemical parameters						
Specific conductance (µmho/cm)	0.00	lower	lower	61.40	50.00	75th
Dissolved oxygen (% saturation)	0.00	lower	lower	93.20	93.30	25th
Ammonia (mg/L)	0.00	lower	lower	0.10	0.10	75th
Nitrate + nitrite (mg/L)	0.00	lower	lower	0.19	0.23	75th
Total Kjeldahl nitrogen (mg/L)	0.00	lower	lower	0.44	0.43	75th
Total nitrogen (mg/L)	0.00	lower	lower	0.56	0.62	75th
pH	0.00	lower	lower	6.80	6.83	75th
Total phosphorus (mg/L)	0.00	lower	lower	0.04	0.04	75th
Temperature (°C)	0.00	lower	lower	8.10	8.47	50th
Total organic carbon (mg/L)	0.00	lower	lower	4.00	5.00	75th
Chemical oxygen demand (mg/L)	0.00	lower	lower	13.00	12.00	75th
Total chlorides (mg/L)	0.00	lower	lower	4.50	4.50	75th
Alkalinity (mg/L)	0.00	lower	lower	11.73	11.80	75th
Turbidity (NTU)	0.00	lower	lower	18.75	18.00	75th
Physical parameters						
Basin area (mi²)	0.00	lower	lower	11,046.16	14,422.42	50th
Total habitat score	0.00	lower	lower	122.00	119.00	25th
Instream cover habitat score	0.00	lower	lower	36.00	34.00	25th
Channel habitat score	0.00	lower	lower	52.75	49.00	25th
Bank habitat score	0.00	lower	lower	32.00	30.00	25th
Hydrohab	0.00	lower	lower	39.75	35.00	25th
% Silt/clay	0.00	lower	lower	18.00	13.00	50th
% Sand	0.00	lower	lower	66.00	72.00	50th
% Gravel	0.00	comparable	comparable	0.00	0.00	50th
% Cobble	0.00	comparable	comparable	0.00	0.00	50th
% Boulder	0.00	comparable	comparable	0.00	0.00	50th
% Bedrock	0.00	comparable	comparable	0.00	0.00	50th
% Hardpan clay	0.00	comparable	comparable	0.00	0.00	50th
Whole watershed natural LU/LC	0.00	lower	lower	83.39	76.00	25th
Whole watershed disturbed LU/LC	0.00	lower	lower	16.10	24.00	75th
Whole watershed, buffered natural LU/LC	0.00	lower	lower	83.37	79.13	25th
Whole watershed, buffered disturbed LU/LC	0.00	lower	lower	16.63	20.87	75th
1-km radius, 50,000-km buffered natural LU/LC	0.00	lower	lower	88.74	79.78	25th
1-km radius, 50,000-km buffered disturbed LU/LC	0.00	lower	lower	11.26	20.22	75th

Figure 11. Example comparison report to document and evaluate co-occurrence of stressors and responses from data collected at the time of the original sampling, *A*, comparison to East bioregion least disturbed and nonimpaired conditions, and *B*, comparison to site-specific comparators. [LD, least disturbed; %, percent; LU/LC, land use/land cover; km, kilometer; µmho/cm, micromhos per centimeter; mg/L, milligrams per liter; °C, degrees Celsius]

B

Parameter	Stressed site	Site-specific comparators				Percentile used for LD
		Creek A	Creek B	Creek C	Creek D	
Chemical parameters						
Specific conductance (µmho/cm)	0.00	0.00	0.00	0.00	0.00	75th
Dissolved oxygen (% saturation)	0.00	0.00	0.00	0.00	0.00	25th
Ammonia (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Nitrate + nitrite (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Total Kjeldahl nitrogen (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Total nitrogen (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
pH	0.00	0.00	0.00	0.00	0.00	75th
Total phosphorus (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Temperature (°C)	0.00	0.00	0.00	0.00	0.00	50th
Total organic carbon (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Chemical oxygen demand (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Total chlorides (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Alkalinity (mg/L)	0.00	0.00	0.00	0.00	0.00	75th
Turbidity (NTU)	0.00	0.00	0.00	0.00	0.00	75th
Physical parameters						
Basin area (mi^2)	0.00	0.00	0.00	0.00	0.00	50th
Total habitat score	0.00	0.00	0.00	0.00	0.00	25th
Instream cover habitat score	0.00	0.00	0.00	0.00	0.00	25th
Channel habitat score	0.00	0.00	0.00	0.00	0.00	25th
Bank habitat score	0.00	0.00	0.00	0.00	0.00	25th
Hydrohab	0.00	0.00	0.00	0.00	0.00	25th
% Silt/clay	0.00	0.00	0.00	0.00	0.00	50th
% Sand	0.00	0.00	0.00	0.00	0.00	50th
% Gravel	0.00	0.00	0.00	0.00	0.00	50th
% Cobble	0.00	0.00	0.00	0.00	0.00	50th
% Boulder	0.00	0.00	0.00	0.00	0.00	50th
% Bedrock	0.00	0.00	0.00	0.00	0.00	50th
% Hardpan clay	0.00	0.00	0.00	0.00	0.00	50th
Whole watershed natural LU/LC	0.00	0.00	0.00	0.00	0.00	25th
Whole watershed disturbed LU/LC	0.00	0.00	0.00	0.00	0.00	75th
Whole watershed, buffered natural LU/LC	0.00	0.00	0.00	0.00	0.00	25th
Whole watershed, buffered disturbed LU/LC	0.00	0.00	0.00	0.00	0.00	75th
1-km radius, 50,000-km buffered natural LU/LC	0.00	0.00	0.00	0.00	0.00	25th
1-km radius, 50,000-km buffered disturbed LU/LC	0.00	0.00	0.00	0.00	0.00	75th

Figure 11. Example comparison report to document and evaluate co-occurrence of stressors and responses from data collected at the time of the original sampling, *A*, comparison to East bioregion least disturbed and nonimpaired conditions, and *B*, comparison to site-specific comparators. [LD, least disturbed; %, percent; LU/LC, land use/land cover; km, kilometer; µmho/cm, micromhos per centimeter; mg/L, milligrams per liter; °C, degrees Celsius]—Continued

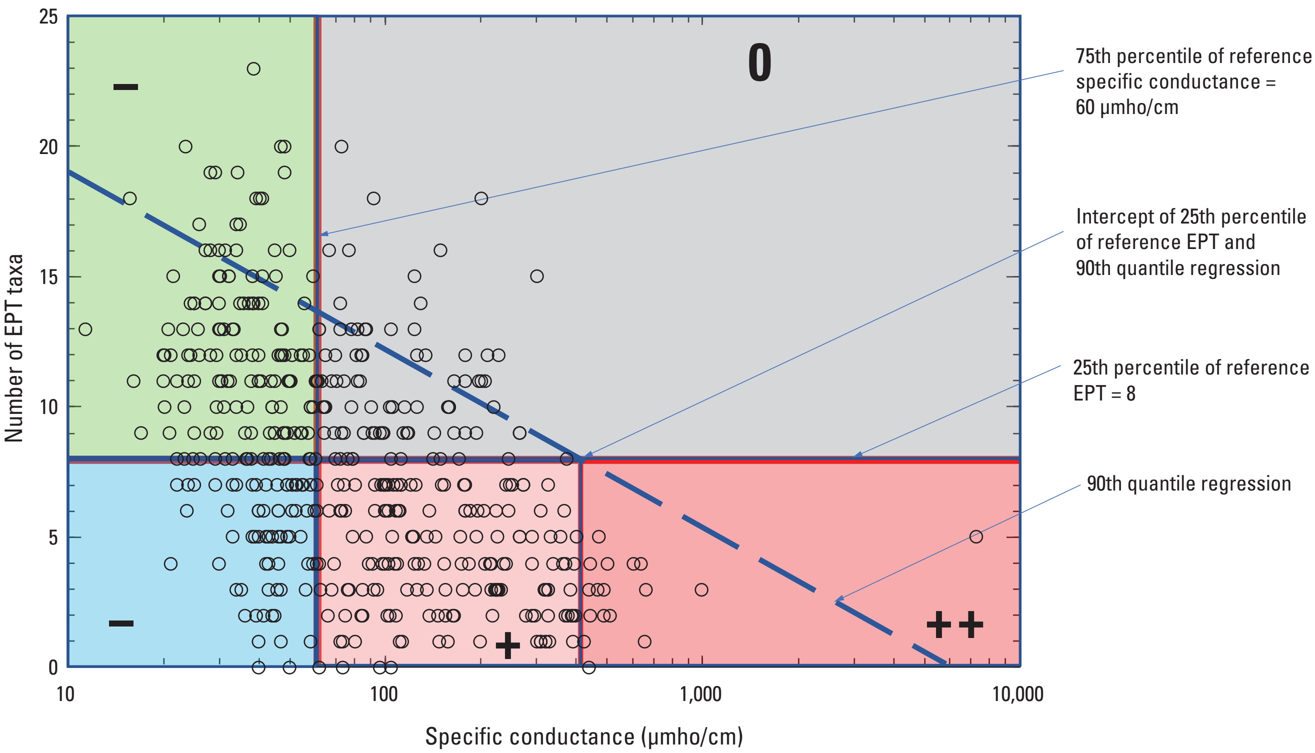


Figure 12. Example of a scatterplot divided into quadrants by using the 75th percentile of reference values of a physical variable (the candidate stressor being evaluated, in this example specific conductance) and the 25th percentile of the biological response (in this example, number of Ephemeroptera, Plecoptera and Trichoptera [EPT] taxa). Color-coded quadrants are interpreted below with corresponding colored text. The blue dashed line represents the 90th percentile quantile regression line. The intercept of the 25th percentile of the reference response variable and the 90th percentile quantile regression distinguishes between some evidence (+) and strong evidence (++). [µmho/cm, micromhos per centimeter]

Hypothesis for Evaluation: High Specific Conductance at Site A Is a Possible Cause of Biological Impairment

Upper left quadrant—If site A falls in this quadrant of the regional distribution, the scatterplot contradicts the hypothesis and is scored as ***incompatible (–)***. In this example, specific conductance is lower than the 75th percentile of specific conductance values from reference sites in the East Bioregion (fig. 7), indicating specific conductance that is generally in accordance with the reference range of specific conductance, and the number of EPT taxa is greater than the reference 25th percentile (generally in accordance with the range of numbers of EPT taxa observed at reference sites), suggesting that specific conductance is not causing biological impairment at site A.

Upper right quadrant—If site A falls in this quadrant, the scatterplot does not support or contradict the hypothesis and is scored as ***uncertain (0)***. In the example, specific conductance is greater than the 75th percentile of specific conductance values from reference sites in the East Bioregion, suggesting that it may be a stressor, but the number of EPT taxa is greater than the reference 25th percentile for the region, which does not support a conclusion of specific conductance causing biological impairment.

Lower left quadrant—If site A falls in this quadrant, the scatterplot contradicts the hypothesis, and is scored as ***incompatible (–).*** In this example, specific conductance is lower than the 75th percentile of specific conductance values from reference sites in the East Bioregion (suggesting that specific conductance is generally in accordance with reference conditions), but the number of EPT taxa is less than the reference 25th percentile, suggesting biological impairment. If site A falls in this quadrant, evidence from the scatterplot suggests that some physical or chemical variable(s) *other than the one plotted* are likely causing the biological impairment. In other words, for this example, biological impairment is illustrated by low numbers of EPT taxa, but this biological impairment is apparently being caused by factors other than specific conductance.

Lower right quadrant—This quadrant is divided into two parts by using the intercept of the 25th percentile of the response reference condition and the 90th quantile regression. If site A falls in this quadrant, the scatterplot supports the hypothesis, and is scored as ***some evidence (+)*** if the specific conductance value from a site is less than the intercept value of the reference percentile and the quantile regression, ***or strong evidence (++)*** if the specific conductance value is greater than that value.

Each of the three types of causal inference—co-occurrence, antecedence, and sufficiency—are scored enabling an overall summary score to be assigned for each causal inference:

1. Co-occurrence—To evaluate co-occurrence, comparison reports may be constructed to document and evaluate physical, chemical, and biological data collected at the time of the original sampling. Figure 11 is an example of a comparison report. Based on an overall evaluation of these comparisons, this line of evidence will result in one of six scores (table 5). Stressor and response variables are compared to SSC values and to bioregional background conditions. There may be different causes for different effects. For example, a low number of Ephemeroptera taxa may be the result of a different cause than a low percentage of dominant filter feeders. It is useful to deconstruct the index or even a metric to find the most relevant effect signal.

2. Antecedence—Five components need to be scored for this line of evidence: (A) causal pathway for indirect stressors relative to comparison sites (fig. 11*B*); (B) causal pathway for indirect stressors relative to bioregion background (fig. 11*A*); (C) reconnaissance data from the stream reach; (D) watershed data (from remote sensing and [or] reconnaissance information); and (E) data from any point sources (such as from compliance reports) that may be influencing the site, and data from other locations in the watershed that drain to the site, obtained from other water-quality databases (non-MDEQ data) and not collected during the original sampling. For any component, if insufficient data are available for evaluation, then the score can be assigned "ND" for no data (table 5). Finally, a single summary score is given for the antecedence causal inference, based on holistic consideration of the five individual scores (or the ones that are available if some are "ND").

3. Sufficiency—This line of evidence is evaluated by using scatterplots of biological and stressor data from Mississippi stratified by bioregions (fig. 7). A single stressor variable may have multiple scatterplots. If the preponderance of plots provides strong evidence for or against the cause, the summary score would be ++ or –, respectively. If some plots provide evidence for the cause while others are uncertain, then a summary score of + may be appropriate, and conversely a summary of – may be appropriate where some plots provide evidence against the cause while others are uncertain. If all plots are uncertain or a roughly equivalent number of plots appear to provide evidence supporting and contradicting the cause, then a 0 score may be assigned (table 5). Sufficiency may also be evaluated by comparing site data to biological and stressor thresholds published by other States, in the literature, or from laboratory experiments.

Table 5. Scoring system for co-occurrence, antecedence, and sufficiency from within a case.

[The number of plusses (+) and minuses (−) increases with the degree to which the evidence either supports or weakens the argument for a candidate cause, and a "0" is used when it is uncertain. Alternatively, a score for ND means "no data."]

Finding	Interpretation	Score
Strong evidence	All of the data support the case for the candidate cause	++
Some evidence	Some of the data support the case for the candidate cause; no data weaken the case for the candidate cause	+
Uncertain	The data neither support nor weaken the case for the candidate cause; an equal amount of data supports and weakens the case, evidence is ambiguous	0
Less likely	Some of the data weakens the case for the candidate cause; no data supports the case	–
Not supported	All of the data weakens the case for the candidate cause	– –
No data	Data do not exist to evaluate this case for the candidate cause	ND

Step 5. Identify the Probable Causes of Impairment by Using a Weight-of-Evidence Approach

The goal of this step is to integrate the assessments of candidate stressors undertaken in previous steps to draw conclusions about probable causes of biological impairment at a study site. In this step, evidence and scores developed in step 4 are compared across all the candidate causes. Evidence for each candidate cause is weighed based upon the quality and quantity of data that were used for evaluation in an effort to inform overall confidence in conclusions. In the case of contrasting evidence, conclusions may be uncertain or may be tentative and accompanied by caveats regarding the level of confidence in the conclusion. Regardless of the degree of confidence, all available evidence should be evaluated before drawing a conclusion. At a minimum, this step can be used to eliminate certain candidate causes from further consideration and to identify data needs for future data collection efforts.

Consistency of Evidence

In the WOE worksheet (fig. 13), scores from co-occurrence, antecedence, and sufficiency are used to make an evaluation for the "consistency of evidence" conclusion. The scores for a candidate cause are not added, but rather are used to inform an overall subjective evaluation of the quality, quantity, consistency, and credibility of the total body of evidence available. The conclusion of consistency of evidence will fall into one of these six categories:

1. "All support,"
2. "Some support" (some are uncertain, but none are incompatible),
3. "Uncertain" (all are inconclusive),
4. "None support,"
5. "Inconsistent" (some support, some do not support), or
6. "No data" (there is no information to evaluate the cause).

Evaluation for this step produces a summary scoring table for each candidate cause and a consistency-of-evidence assessment across multiple lines of evidence. In some cases, the analysis may point clearly to a probable cause or causes. In most cases, it is possible to reduce the number of candidate causes by eliminating some causes from consideration (for example, those evaluated as "none support" based on the consistency-of-evidence assessment). Based on the consistency-of-evidence assessment, each candidate cause can then be assigned to one of five categories of probability of cause resulting in effect:

1. Probable primary cause of biological impairment,
2. Probable secondary cause of biological impairment,
3. Less probable cause of biological impairment,
4. Unlikely cause of biological impairment, or
5. Insufficient evidence to determine causality of biological impairment.

Figure 13 is an example of the WOE worksheet scored for each causal inference, consistency of evidence, and an overall assignment of probability of cause for decrease in suitable habitat as a proximate cause.

Step 6. Generate a Report of the Results

The results of the previous steps are compiled into a report in the last step to present to stakeholders and decision makers. The report of results will generally include an identification of probable causes of impairment, accompanied by a detailed description of the conceptual, analytic, and logical bases for the conclusions drawn. In addition, an evaluation of uncertainty relating to the SI process should be documented. Results in the summary report may include the following:

1. The reason for the SI,
2. A list of the candidate causes and the information used to support their selection,
3. The sources of data used in the analysis,
4. Conceptual models of causal pathways,
5. Tables of evidence derived from the data,
6. Key evidence that strengthens the probable cause(s) and weakens the others,
7. Confidence in conclusions (data quality, data quantity, consistency, and credibility), and
8. Next steps and recommendations (for example, to address a lack of data through future data-collection efforts).

Decrease in suitable habitat			
Causal inference	**Type of evidence**	**Evidence**	**Score**
Co-occurrence	Stressor level relative to site-specific comparators (see fig. 11*B*)	Document the habitat scores here and specific taxa that support this causal inference	
	Stressor level relative to bioregional background (see fig. 11*A*)	Document the habitat scores here and specific taxa that support this causal inference	
Summary for evidence of co-occurrence			Summary score for above two
Antecedence	Causal pathway for indirect stressors relative to comparison sites (see fig. 11*B*)		
	Causal pathway for indirect stressors relative to bioregion background (see fig. 11*A*)		
	Reconnaissance of station		
	Reconnaissance of watershed		
	Point sources, enSPIRE (MDEQ's water-quality database), other data sources		
Summary for evidence of antecedence			Summary score for above five
Sufficiency	Biological data and stressor data compared to position on bioregion scatterplot and threshold (fig. 12)		
	Biological data and stressor data compared to thresholds published by other States or in literature		
	Biological data and stressor data compared to thresholds from laboratory, mesocosm, experiments		
Summary for evidence of sufficiency			Summary score for above three
Consistency of evidence			Scored based on the consistency of the above three summary scores
Probability of cause resulting in effect			One of the five categories of probability

Figure 13. Example weight-of-evidence worksheet for decrease in suitable habitat.

Options for Future Enhancement of This Framework

The framework presented here for impaired streams based on biological community assessment in Mississippi allows future enhancements to improve the quality and efficiency and reduce the cost and uncertainty of the SI process. Potential options for improvement fall into two categories: development of advanced analytical approaches and monitoring for additional stressor and response indicators.

Development of Advanced Analytical Approaches

CADDIS outlines several advanced analytical approaches for causal analysis, which could be incorporated as added WOE tools. The presence of a large statewide database of benthic macroinvertebrate community data coupled with physical and chemical environmental data enables a number of analytical approaches, which are discussed in the following sections, that could strengthen and streamline the SI process.

Predicting Environmental Conditions from Biological Observations

This approach is based on the theory that different taxa generally require different environmental conditions to persist. The environmental requirements of different taxa can be represented with taxon-environment relation. A taxon-environment relation quantifies the relation between the probability of observing a particular taxon and the value of one or more environmental variables. Existing taxon-environment relations can be used (EPA, 2010a) or can be estimated by using regional datasets. A variety of statistical methods can be used to infer environmental conditions at a site by using study-site biological samples and taxon-environment relations, including weighted averaging, maximum likelihood, and categorical tolerance (Yuan, 2007). Given the extent of Mississippi's statewide database, with hundreds of field sites sampled across a wide range of disturbances, opportunities exist to develop taxon-stressor relations for many different combinations of taxonomic groups and stressors. Such an approach could be especially beneficial for nonchemical stressors where dose-response studies are difficult to conduct in the laboratory.

Propensity Score Analysis

Analysis of the role of a given stressor may be inaccurate if confounding variables are not considered. One method for developing a deconfounded analysis of observational data involves the calculation of a propensity score. First, environmental data, such as land cover, habitat metrics, topographic metrics, and soil or geologic variables, are used to model a variable representing a candidate stressor (for example, total nitrogen; Yuan, 2010). Modeled values for the candidate stressor (for example, predicted total nitrogen) become propensity scores that are then used to stratify the dataset into bins based on propensity score. Finally, relations between propensity scores and biological metrics of interest are modeled. This approach has the advantage of minimizing the confounding effects of environmental variables other than the candidate stressor of interest because they are accounted for in calculating the propensity scores, which reduces bias and allows variations in biological metrics to be attributed more directly to the candidate stressors of interest (Pearson and others, 2016). Propensity scores are only useful with very large and complete datasets.

Evidence of Alteration

Evidence of alteration is becoming more common with larger datasets. Similar to symptoms and signs of disease used for human and veterinary diagnostics, differential occurrence and proportion of taxa can be indicative of a particular cause (Coffey and others, 2014). Evidence of alteration can be derived from regional models of different communities associated with different stressors. The genera at the test site are used to predict the probable cause. This could involve use of the probability of a stressor from a nonmetric multidimensional scaling model of the taxon from impaired sites as evidence of alteration (Zheng and others, 2015). Models of specific alteration using taxon data can be very compelling. If there are not enough impaired sites of a particular type, data from Mississippi may be combined with other available data.

Concept of Stressor Syndromes

Groups of stressors are not randomly assigned across the landscape, but instead typically result from human-development activities and land-use and resource-management practices. As a result, stream sites with common land-use patterns in their watersheds may share similar physical and water-quality conditions. For example, “urban stream syndrome” is an increasingly used term to describe a set of commonly observed parameters across streams with high levels of urban development in their watersheds, including increasing “flashiness” of streamflow, warmer water temperatures, nutrient enrichment, decreased channel complexity, and loss of sensitive species (Walsh and others, 2005).

In Mississippi, relatively few watersheds are heavily urban, but many are dominated by agricultural land use, especially row-crop agriculture. Thus, exploratory work into a “row-crop syndrome” framework could be useful in Mississippi, to the degree that large numbers of sites in watersheds dominated by row-crop agriculture appear to share common physical parameters reflecting water quality and habitat. For example, Munn and others (2018) examined a suite of related stressors common to agricultural streams in the midwestern United States, including nutrients, herbicides, fungicides, sediment loads, and altered streamflow. Such an approach could be paired with propensity score analysis to explore row-crop agriculture (for example, as a proportion of land use within each watershed) as a driver of multiple candidate stressors such as nutrients (total nitrogen, total phosphorous, etc.), DO, and other water-quality indicators. Propensity score analysis could allow other potentially confounding factors such as watershed geology, topography, and soil type (Carlisle and others, 2009) to be accounted for explicitly to minimize bias in evaluating the effects of row-crop agriculture on candidate stressors. Multivariate machine-learning approaches such as boosted regression trees or random forest models (Munn and others, 2018; Waite and others, 2019) could also be employed to model the interactive effects of many physical parameters simultaneously and to enable quantification of the relative importance of each parameter to various biological metrics related to stream impairment.

Monitoring for Additional Stressor and Response Indicators

To support future SI efforts, MDEQ may wish to consider investment in developing new datasets and tools to better evaluate potential intermediate and proximate causes not currently represented in MDEQ's routine monitoring program. Nonroutine data could include indicators of hydrologic and geomorphic alteration from reference conditions, instream temperature monitoring data (historically, most temperature measurements occur in winter as a point sample and hence do not represent temperature cycles diurnally or across seasons), DO (similarly, most sampling historically occurs in winter as a point sample), alteration of food sources for aquatic organisms, and prevalence of toxic substances such as pesticides, herbicides, fungicides, and other organic contaminants. Furthermore, to complement MDEQ's extensive database of macroinvertebrate data, monitoring protocols and datasets could be developed to evaluate other trophic-level communities and taxonomic groups (such as fish and algae) which may also be influenced by similar sets of stressors.

Summary

In order to calculate total maximum daily loadings for Mississippi streams listed as biologically impaired on the §303(d) list, the actual stressor(s) causing the impairment must be known. The U.S. Geological Survey, in cooperation with the Mississippi Department of Environmental Quality, developed a framework to identify stressors in biologically impaired streams and rivers of Mississippi. This framework was adapted from the existing U.S. Environmental Protection Agency framework for stressor identification and involves six general steps: (1) define the impairment, (2) list the candidate causes of impairment and develop a conceptual model, (3) compile all data relevant to the impairment and conceptual model, (4) evaluate the data, (5) identify the probable causes of impairment by using a weight-of-evidence approach, and (6) generate a report of results. The intent is for the Mississippi stressor identification framework to reduce subjectivity, provide consistency, and allow for adaptation as the science evolves.

References Cited

Allan, J.D., McIntyre, P.B., Smith, S.D.P., Halpern, B.S., Boyer, G.L., Buchsbaum, A., Burton, G.A., Campbell, L.M., Chadderton, W.L., Ciborowski, J.J.H., Doran, P.J., Eder, T., Infante, D.M., Johnson, L.B., Joseph, C.A., Marino, A.L., Prusevich, A., Read, J.G., Rose, J.B., Rutherford, E.S., Sowa, S.P., and Steinman, A.D., 2013, Joint analysis of stressors and ecosystem services to enhance restoration effectiveness: Proceedings of the National Academy of Sciences of the United States of America, v. 110, no. 1, p. 372–377. [Also available at https://doi.org/10.1073/pnas.1213841110.]

Allan, J.D., Yuan, L., Black, P., Stockton, T., Davies, P., Magierowski, R., and Read, S., 2012, Investigating the relationships between environmental stressors and stream condition using Bayesian belief networks: Freshwater Biology, v. 57, p. 58–73. [Also available at https://doi.org/10.1111/j.1365-2427.2011.02683.x.]

Armstrong, D.S., Richards, T.A., and Levin, S.B., 2011, Factors influencing riverine fish assemblages in Massachusetts: U.S. Geological Survey Scientific Investigations Report 2011–5193, 58 p. [Also available at https://doi.org/10.3133/sir20115193.]

Bernhardt, E.S., Lutz, B.D., King, R.S., Fay, J.P., Carter, C.E., Helton, A.M., Campagna, D., and Amos, J., 2012, How many mountains can we mine? Assessing the regional degradation of central Appalachian rivers by surface coal mining: Environmental Science & Technology, v. 46, no. 15, p. 8115–8122. [Also available at https://doi.org/10.1021/es301144q.]

Black, R.W., Moran, P.W., and Frankforter, J.D., 2011, Response of algal metrics to nutrients and physical factors and identification of nutrient thresholds in agricultural streams: Environmental Monitoring and Assessment, v. 175, p. 397–417. [Also available at https://doi.org/10.1007/s10661-010-1539-8.]

Blazer, V.S., Iwanowicz, L.R., Henderson, H., Mazik, P.M., Jenkins, J.A., Alvarez, D.A., and Young, J.A., 2012, Reproductive endocrine disruption in smallmouth bass (Micropterus dolomieu) in the Potomac River basin—Spatial and temporal comparisons of biological effects: Environmental Monitoring and Assessment, v. 184, no. 7, p. 4309–4334. [Also available at https://doi.org/10.1007/s10661-011-2266-5.]

Breiman, L., 2001, Random forests: Machine Learning, v. 45, no. 1, p. 5–32. [Also available at https://doi.org/10.1023/A:1010933404324.]

Cade, B.S., 2011, Estimating equivalence with quantile regression: Ecological Applications, v. 21, no. 1, p. 281–289. [Also available at https://doi.org/10.1890/09-2060.1.]

Cade, B.S., and Noon, B.R., 2003, A gentle introduction to quantile regression for ecologists: Frontiers in Ecology and the Environment, v. 1, no. 8, p. 412–420. [Also available at https://doi.org/10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2.]

Cade, B.S., Noon, B.R., and Flather, C.H., 2005, Quantile regression reveals hidden bias and uncertainty in habitat models: Ecology, v. 86, no. 3, p. 786–800. [Also available at https://doi.org/10.1890/04-0785.]

Cade, B.S., Terrell, J.W., and Porath, M.T., 2008, Estimating fish body condition with quantile regression: North American Journal of Fisheries Management, v. 28, no. 2, p. 349–359. [Also available at https://doi.org/10.1577/M07-048.1.]

Calhoun, D.L., Gregory, M.B., and Weyers, H.S., 2008, Algal and invertebrate community composition along agricultural gradients—A comparative study from two regions of the eastern United States: U.S. Geological Survey Scientific Investigations Report 2008–5046. [Also available at https://pubs.usgs.gov/sir/2008/5046/.]

Carlisle, D.M., Falcone, J., and Meador, M.R., 2009, Predicting the biological condition of streams—Use of geospatial indicators of natural and anthropogenic characteristics of watersheds: Environmental Monitoring and Assessment, v. 151, p. 143–160. [Also available at https://doi.org/10.1007/s10661-008-0256-z.]

Carlisle, D.M., Falcone, J., Wolock, D.M., Meador, M.R., and Norris, R.H., 2010, Predicting the natural flow regime—Models for assessing hydrologic alteration in streams: River Research and Applications, v. 2, p. 118–136. [Also available at https://doi.org/10.1002/rra.1247.]

Carlisle, D.M., Wolock, D.M., and Meador, M.R., 2011, Alteration of streamflow magnitudes and potential ecological consequences—A multiregional assessment: Frontiers in Ecology and the Environment, v. 9, no. 5, p. 264–270. [Also available at https://doi.org/10.1890/100053.]

Clements, W.H., 1994, Benthic invertebrate community responses to heavy metals in the Upper Arkansas River Basin, Colorado: Journal of the North American Benthological Society, v. 13, no. 1, p. 30–44. [Also available at https://doi.org/10.2307/1467263.]

Clements, W.H., Carlisle, D.M., Lazorchak, J.M., and Johnson, P.C., 2000, Heavy metals structure benthic communities in Colorado mountain streams: Ecological Applications, v. 10, no. 2, p. 626–638. [Also available at https://doi.org/10.1890/1051-0761(2000)010[0626:HMSBCI]2.0.CO;2.]

Coffey, D.B., Cormier, S., and Harwood, J., 2014, Using field-based species sensitivity distributions to infer multiple causes: Human and Ecological Risk Assessment, v. 20, no. 2, p. 402–432. [Also available at https://doi.org/10.1080/10807039.2013.767071.]

Cormier, S.M., Norton, S.B., and Suter, G.W., 2003, The U.S. Environmental Protection Agency's stressor identification guidance—A process for determining the probable causes of biological impairments: Human and Ecological Risk Assessment, v. 9, no. 6, p. 1431–1443. [Also available at https://doi.org/10.1080/10807030390250930.]

Cormier, S.M., Smith, M., Norton, S., and Neiheisel, T., 2000, Assessing ecological risk in watersheds—A case study of problem formulation in the Big Darby Creek watershed, Ohio, USA: Environmental Toxicology and Chemistry, v. 19, no. 4, p. 1082–1096. [Also available at https://doi.org/10.1002/etc.5620190439.]

Crane, M., Burton, G.A., Culp, J.M., Greenberg, M.S., Munkittrick, K.R., Ribeiro, R., Salazar, M.H., and St-Jean, S.D., 2007, Review of aquatic in situ approaches for stressor and effect diagnosis: Integrated Environmental Assessment and Management, v. 3, no. 2, p. 234–245. [Also available at https://doi.org/10.1897/IEAM_2006-027.1.]

Cutler, D.R., Edwards, T.C., Jr., Beard, K.H., Cutler, A., Hess, K., Gibson, J., and Lawler, J.J., 2007, Random forests for classification in ecology: Ecology, v. 88, no. 11, p. 2783–2792. [Also available at https://doi.org/10.1890/07-0539.1.]

De'ath, G., 2007, Boosted regression trees for ecological modeling and prediction: Ecology, v. 88, no. 1, p. 243–251. [Also available at https://doi.org/10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2.]

De'ath, G., and Fabricius, K.E., 2000, Classification and regression trees—A powerful yet simple technique for ecological data analysis: Ecology, v. 81, no. 11, p. 3178–3192. [Also available at https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2.]

DeGasperi, C.L., Berge, H.B., Whiting, K.R., Burkey, J.J., Cassin, J.L., and Fuerstenberg, R.R., 2009, Linking hydrologic alteration to biological impairment in urbanizing streams of the Puget Lowland, Washington, USA: Journal of the American Water Resources Association, v. 45, no. 2, p. 512–533. [Also available at https://doi.org/10.1111/j.1752-1688.2009.00306.x.]

Dodds, W., and Welch, E., 2000, Establishing nutrient criteria in streams: Journal of the North American Benthological Society, v. 19, no. 1, p. 186–196. [Also available at https://doi.org/10.2307/1468291.]

Doll, J.C., 2011, Predicting biological impairment from habitat assessments: Environmental Monitoring and Assessment, v. 182, p. 259–277. [Also available at https://doi.org/10.1007/s10661-011-1874-4.]

Downes, B.J., 2010, Back to the future—Little-used tools and principles of scientific inference can help disentangle effects of multiple stressors on freshwater ecosystems: Freshwater Biology, v. 55, p. 60–79. [Also available at https://doi.org/10.1111/j.1365-2427.2009.02377.x.]

Elith, J., Leathwick, J.R., and Hastie, T., 2008, A working guide to boosted regression trees: Journal of Animal Ecology, v. 77, no. 4, p. 802–813. [Also available at https://doi.org/10.1111/j.1365-2656.2008.01390.x.]

Gregorutti, B., Michel, B., and Saint-Pierre, P., 2017, Correlation and variable importance in random forests: Statistics and Computing, v. 27, p. 659–678. [Also available at https://doi.org/10.1007/s11222-016-9646-1.]

Griggs, A., and Buchanan, C., 2012, Stressor identification methods used for impaired waters in the Chesapeake Bay watershed States, DRAFT Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup: Interstate Commission on the Potomac River Basin Report 12-05.

Harden, S.L., Cuffney, T.F., Terziotti, S., and Kolb, K.R., 2013, Relation of watershed setting and stream nutrient yields at selected sites in central and eastern North Carolina, 1997–2008: U.S. Geological Survey Scientific Investigations Report 2013–5007. [Also available at https://doi.org/10.3133/sir20135007.]

Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., and Schlosser, I.J., 1986, Assessing biological integrity in running waters—A method and its rationale: Illinois Natural History Survey, Special Publication 5, 28 p. [Also available at https://semspub.epa.gov/work/01/554353.pdf.]

Klein, R.D., 1979, Urbanization and stream quality impairment: Water Resources Bulletin, v. 15, no. 4, p. 948–963. [Also available at https://doi.org/10.1111/j.1752-1688.1979.tb01074.x.]

Knight, R.R., Murphy, J.C., Wolfe, W.J., Saylor, C.F., and Wales, A.K., 2014, Ecological limit functions relating fish community response to hydrologic departures of the ecological flow regime in the Tennessee River basin, United States: Ecohydrology, v. 7, p. 1262–1280. [Also available at https://doi.org/10.1002/eco.1460.]

Lange, K., Townsend, C.R., and Matthaei, C.D., 2014, Can biological traits of stream invertebrates help disentangle the effects of multiple stressors in an agricultural catchment?: Freshwater Biology, v. 59, no. 12, p. 2431–2446. [Also available at https://doi.org/10.1111/fwb.12437.]

Lapointe, B.E., Herren, L.W., and Bedford, B.J., 2012, Effects of hurricanes, land use, and water management on nutrient and microbial pollution—St. Lucie Estuary, southeast Florida: Journal of Coastal Research, v. 28, p. 1345–1361. [Also available at https://doi.org/10.2112/JCOASTRES-D-12-00070.1.]

Lavoie, I., Morin, S., Laderriere, V., and Fortin, C., 2018, Freshwater diatoms as indicators of combined long-term mining and urban stressors in Junction creek (Ontario, Canada): Environments, v. 5, no. 2. [Also available at https://doi.org/10.3390/environments5020030.]

Liaw, W., and Wiener, M., 2014, RandomForest: Breiman and Cutler's random forests for classifications and regression, accessed May 2020 at http://cran.r-project.org/web/packages/randomForest/.

Maloney, K.O., Weller, D.E., Michaelson, D.E., and Ciccotto, P.J., 2013, Species distribution models of freshwater stream fishes in Maryland and their implications for management: Environmental Modeling and Assessment, v. 18, no. 1, p. 1–12. [Also available at https://doi.org/10.1007/s10666-012-9325-3.]

Maloney, K.O., Weller, D.E., Russell, M.J., and Hothorn, T., 2009, Classifying the biological condition of small streams—An example using benthic macroinvertebrates: Freshwater Science, v. 28, p. 869–884. [Also available at https://doi.org/10.1899/08-142.1.]

Maryland Department of the Environment, 2016, Watershed report for biological impairment of the non-tidal Potomac River Frederick County watershed, Frederick and Washington Counties Maryland biological stressor identification analysis results and interpretation: Baltimore, Maryland, Maryland Department of the Environment.

Matthaei, C.D., Piggott, J.J., and Townsend, C.R., 2010, Multiple stressors in agricultural streams—Interactions among sediment addition, nutrient enrichment and water abstraction: Journal of Applied Ecology, v. 47, no. 3, p. 639–649. [Also available at https://doi.org/10.1111/j.1365-2664.2010.01809.x.]

Merriam, E.R., Petty, J.T., Merovich, G.T., Jr., Fulton, J.B., and Strager, M.P., 2011, Additive effects of mining and residential development on stream conditions in a central Appalachian watershed: Journal of the North American Benthological Society, v. 30, no. 2, p. 399–418. [Also available at https://doi.org/10.1899/10-079.1.]

Miguntanna, N.S., Egodawatta, P., Kokot, S., and Goonetilleke, A., 2010, Determination of a set of surrogate parameters to assess urban stormwater quality: The Science of the Total Environment, v. 408, no. 24, p. 6251–6259. [Also available at https://doi.org/10.1016/j.scitotenv.2010.09.015.]

Minnesota Pollution Control Agency, 2013, St. Louis River watershed stressor identification report: Saint Paul, Minn., Minnesota Pollution Control Agency.

Morris, C.C., Simon, T.P., and Newhouse, S.A., 2006, A local-scale in situ approach for stressor identification of biologically impaired aquatic systems: Archives of Environmental Contamination and Toxicology, v. 50, no. 3, p. 325–334. [Also available at https://doi.org/10.1007/s00244-005-1066-8.]

Munn, M.D., Waite, I.R., and Konrad, C.P., 2018, Assessing the influence of multiple stressors on stream diatom metrics in the upper Midwest, USA: Ecological Indicators, v. 85, p. 1239–1248. [Also available at https://doi.org/10.1016/j.ecolind.2017.09.005.]

Norton, S.B., Cormier, S.M., Smith, M., and Jones, R.C., 2000, Can biological assessments discriminate among types of stress? A case study from the eastern corn belt plains ecoregion: Environmental Toxicology and Chemistry, v. 19, no. 4, p. 1113–1119. [Also available at https://doi.org/10.1002/etc.5620190442.]

Ohio Environmental Protection Agency, 2010, Biological and habitat study of the Olentangy River, 2005 and 2009: Groveport, Ohio, Ohio Environmental Protection Agency.

Omernik, J.M., 1987, Ecoregions of the conterminous United States: Annals of the Association of American Geographers, v. 77, no. 1, p. 118–125. [Also available at https://doi.org/10.1111/j.1467-8306.1987.tb00149.x.]

Onorato, D., Marion, K.R., and Angus, R.A., 1998, Longitudinal variations in the ichthyofaunal assemblages of the upper Cahaba River—Possible effects of urbanization in a watershed: Journal of Freshwater Ecology, v. 13, no. 2, p. 139–154. [Also available at https://doi.org/10.1080/02705060.1998.9663602.]

Palmer, M., 2016, Ordination methods for ecologists: Stillwater, Okla., Oklahoma State University, accessed May 2020 at http://ordination.okstate.edu/.

Pearson, C.E., Ormerod, S.J., Symondson, W.O., and Vaughan, I.P., 2016, Resolving large-scale pressures on species and ecosystems—Propensity modelling identifies agricultural effects on streams: Journal of Applied Ecology, v. 53, no. 2, p. 408–417. [Also available at https://doi.org/10.1111/1365-2664.12586.]

Porter, C.M., Butler, D.R., and Janz, D.M., 2015, Central Oklahoma bioassessment study—Evaluation of stream health by using fish and macroinvertebrate communities as biological indicators: Proceedings of the Oklahoma Academy of Science, v. 80, p. 61–70.

Prasad, A.M., Iverson, L.R., and Liaw, A., 2006, Newer classification and regression tree techniques—Bagging and random forests for ecological prediction: Ecosystems (New York, N.Y.), v. 9, no. 2, p. 181–199. [Also available at https://doi.org/10.1007/s10021-005-0054-1.]

Schleiter, K.E., 2009, Proving causation in environmental litigation: AMA Journal of Ethics, v. 11, no. 6, p. 456–460. [Also available at https://doi.org/10.1001/virtualmentor.2009.11.6.hlaw1-0906.]

Schmidt, T.S., Clements, W.H., and Cade, B.S., 2012, Estimating risks to aquatic life using quantile regression: Freshwater Science, v. 31, no. 3, p. 709–723. [Also available at https://doi.org/10.1899/11-133.1.]

Smucker, N.J., Becker, M., Detenbeck, N.E., and Morrison, A.C., 2013, Using algal metrics and biomass to evaluate multiple ways of defining concentration-based nutrient criteria in streams and their ecological relevance: Ecological Indicators, v. 32, p. 51–61. [Also available at https://doi.org/10.1016/j.ecolind.2013.03.018.]

Steen, P.J., Zorn, T.G., Seelbach, P.W., and Schaeffer, J.S., 2008, Classification tree models for predicting distributions of Michigan stream fish from landscape variables: Transactions of the American Fisheries Society, v. 137, no. 4, p. 976–996. [Also available at https://doi.org/10.1577/T07-119.1.]

Stribling, J.B., Jessup, B.K., and Leppo, E.W., 2016, The Mississippi-Benthic Index of Stream Quality (M-BISQ)—Recalibration and testing: Prepared for Mississippi Department of Environmental Quality, Office of Pollution Control, Jackson, Mississippi. [Also available at https://www.mdeq.ms.gov/wp-content/uploads/2017/05/MBISQ2015-with-appendixes-FINAL-20160203c.pdf.]

Suter, G.W., II, Norton, S.B., and Cormier, S.M., 2002, A methodology for inferring the causes of observed impairments in aquatic ecosystems: Environmental Toxicology and Chemistry, v. 21, p. 1101–1111.

Townsend, C.R., Uhlmann, S.S., and Matthaei, C.D., 2008, Individual and combined responses of stream ecosystems to multiple stressors: Journal of Applied Ecology, v. 45, no. 6, p. 1810–1819. [Also available at https://doi.org/10.1111/j.1365-2664.2008.01548.x.]

U.S. Army Corps of Engineers, 2009, National inventory of dams, accessed May 2020 at http://nid.usace.army.mil/cm_apex/f?p=838:12.

U.S. Environmental Protection Agency [EPA], 1997, Guidelines for preparation of the comprehensive State water quality assessments (305(b) Reports) and electronic updates—Report contents: Washington, D.C., U.S. Environmental Protection Agency, Office of Water, EPA-841-B-97-0002A.

U.S. Environmental Protection Agency [EPA], 2000, Stressor identification guidance document: Washington, D.C., U.S. Environmental Protection Agency, EPA-822-B00-025, 228 p.

U.S. Environmental Protection Agency [EPA], 2002, Consolidated assessment and listing methodology toward a compendium of best practices: Washington, D.C., U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds.

U.S. Environmental Protection Agency [EPA], 2010a, Causal Analysis/Diagnosis Decision Information System (CADDIS), accessed May 2020 at http://www.epa.gov/caddis.

U.S. Environmental Protection Agency [EPA], 2010b, Inferring causes of biological impairment in the Clear Fork watershed, West Virginia (Final): Washington, D.C., U.S. Environmental Protection Agency, EPA/600/R-08/146.

U.S. Geological Survey, 2020, National Hydrography Dataset, accessed May 11, 2020, at https://www.usgs.gov/core-science-systems/ngp/national-hydrography/national-hydrography-dataset?qt-science_support_page_related_con=0#qt-science_support_page_related_con.

Usio, N., Nakajima, H., Kamiyama, R., Wakana, I., Hiruta, S., and Takamura, N., 2006, Predicting the distribution of invasive crayfish (Pacifastacus leniusculus) in a Kusiro Moor marsh (Japan) using classification and regression trees: Ecological Research, v. 21, no. 2, p. 271–277. [Also available at https://doi.org/10.1007/s11284-005-0120-3.]

Vander Laan, J.J., 2012, Environmental assessment of streams—Linking land use, instream stressors, and biological indices to infer likely causes of ecological impairment: Utah State University.

Vander Laan, J.J., Hawkins, C.P., Olson, J.R., and Hill, R.A., 2013, Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams: Freshwater Science, v. 32, no. 3, p. 801–820. [Also available at https://doi.org/10.1899/12-186.1.]

Virginia Department of Environmental Quality, 2014, Benthic TMDL development stressor analysis report, Happy Creek, Warren County: Virginia, Virginia Department of Environmental Quality.

Voss, K.A., Pohlman, A., Viswanathan, S., Gibson, D., and Purohit, J., 2012, A study of the effect of physical and chemical stressors on biological integrity within the San Diego hydrologic region: Environmental Monitoring and Assessment, v. 184, no. 3, p. 1603–1616. [Also available at https://doi.org/10.1007/s10661-011-2064-0.]

Wagenhoff, A., Townsend, C.R., Phillips, N., and Matthaei, C.D., 2011, Subsidy-stress and multiple-stressor effects along gradients of deposited fine sediment and dissolved nutrients in a regional set of streams and rivers: Freshwater Ecology, v. 56, no. 9, p. 1916–1936. [Also available at https://doi.org/10.1111/j.1365-2427.2011.02619.x.]

Waite, I.R., Munn, M.D., Moran, P.W., Konrad, C.P., Nowell, L.H., Meador, M.R., Van Metre, P.C., and Carlisle, D.M., 2019, Effects of urban multi-stressors on three stream biotic assemblages: The Science of the Total Environment, v. 660, p. 1472–1485. [Also available at https://doi.org/10.1016/j.scitotenv.2018.12.240.]

Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M., and Morgan, R.P., II, 2005, The urban stream syndrome—Current knowledge and the search for a cure: Journal of the North American Benthological Society, v. 24, no. 3, p. 706–723. [Also available at https://doi.org/10.1899/04-028.1.]

Wang, L., Lyons, J., Kanehi, P., Bannerman, R., and Emmons, E., 2000, Watershed urbanization and changes in fish communities in southeastern Wisconsin streams: Journal of the American Water Resources Association, v. 36, no. 5, p. 1173–1189. [Also available at https://doi.org/10.1111/j.1752-1688.2000.tb05719.x.]

Yuan, L.L., 2007, Maximum likelihood method for predicting environmental conditions from assemblage composition—The R package bio.infer: Journal of Statistical Software, v. 22, no. 3, p. 1–20. [Also available at https://doi.org/10.18637/jss.v022.i03.]

Yuan, L.L., 2010, Estimating the effects of excess nutrients on stream invertebrates from observational data: Ecological Applications, v. 20, no. 1, p. 110–125. [Also available at https://doi.org/10.1890/08-1750.1.]

Zheng, L., Gerritsen, J., and Cormier, S.M., 2015, The Clear Fork watershed case study—The value of State monitoring programs, chapter 4 of Norton, S.B., Cormier, S.M., and Suter, G.W., II, eds., Ecological causal assessment: Boca Raton, Fla. CRC Press, Taylor and Francis Group, p. 47–64.

Appendix 1. Field Forms Used During Reconnaissance in Stressor Identification

Reconnaissance Field Form for Physical, Chemical, and Other Data and Observations

CA Sample Site Data Form______________________Creek at ________________________(Road)

Time__________________Date__________________IBI #_____________ Basin ______________

Weather Conditions _____________________________ GPS N______◦_______’______.______”

Taken By___________________________ W______◦_______’______.______”

Temp (◦C)	Cond. (mg/l)	TDS (mg/l)	D.O. (mg/l)	%D.O.	pH	Sample Depth (feet)	Bottom Depth (feet)

Photo File #’s Upsteam____________Downstream__________Other_____________

Flow Status FLOWING PONDED DRY

Water Clarity/Turbidity CLEAR SLIGHT MODERATE SEVERE

Flow Alteration YES / NO Channel Alteration YES / NO

Morphology B C D E F G Channel Evolution I II III IV V VI

Entrenchment LOW MODERATE HIGH HEAD CUT/Nick Point?___

Bank Erodibility HIGH BNKS LEANING TREES LOW ROOT MASS OTHER________________

Canopy OPEN FILTERED OTHER___________________

Riparian ________Meters Age_______TREES SHRUBS GRASS LU_________________

Bed Material Consolidated YES / NO CLAY SILT SAND GRAVEL OTHER______________

Channel Structure RIFFLE POOL RUN POINT BAR

Siltation NONE SLIGHT MODERATE SEVERE

ALGAE FILM TRASH OTHER_______________________

Habitat CUT BANKS LEAF PACK WOODY BOTTOM

NOTES :____

Figure 1.1. Reconnaissance field form for physical, chemical, and other data and observations. [CA, causal analysis]

enSPIRE#	RAPID GEOMORPHIC ASSESSMENT (RGA) FORM CHANNEL STABILITY RANKING SCHEME		Project:
Station Name:			Lat:
Station Location:			Long:
Date and time:	Team	Basin	
County		Photo (US, DS, others:)	Pattern (**M**eander/**S**traight/**B**raided):

1. Primary bed material

Hardpan	Boulder/Cobble	Gravel	Sand	Silt/Clay
0	1	2	3	4

2. Bed/bank protection (rip rap, concrete, sheetpile, etc.)

No	Bed	(with) 1 bank protected	2 banks protected
0	1	2	3

3. Degree of incision (relative elevation of "normal" low water if flood plain/terrace is 100%)

0-10%	11-25%	26-50%	51-75%	76-100%
4	3	2	1	0

4. Degree of constriction (relative decrease in top-bank width from upstream to downstream)

0-10%	11-25%	26-50%	51-75%	76-100%
0	1	2	3	4

5. Streambank erosion (dominant process each bank)

	None	Fluvial	mass wasting (failures)
Left	0	1	2
Right	0	1	2

6. Streambank instability (percent of each bank failing)

	0-10%	11-25%	26-50%	51-75%	76-100%
Left	0	0.5	1	1.5	2
Right	0	0.5	1	1.5	2

7. Established riparian vegetative cover (woody or stabilizing perennial grasses each bank)

	0-10%	11-25%	26-50%	51-75%	76-100%
Left	2	1.5	1	0.5	0
Right	2	1.5	1	0.5	0

8. Occurrence of bank accretion (percent of each bank with fluvial deposition)

	0-10%	11-25%	26-50%	51-75%	76-100%
Left	2	1.5	1	0.5	0
Right	2	1.5	1	0.5	0

9. Stage of channel evolution (I and VI are generally stable < 11 total score, III and IV least stable generally > 15)

I	II	III	IV	V	VI
0	1	2	4	3	1.5

10. TOTAL____(35 Max)

Figure 1.2. Rapid geomorphic assessment field form.

Appendix 2. Tools for M-BISQ Data Compilation and Evaluation

Comparison reports (for physical and chemical data and for biological data) are documented for each bioregion by using spreadsheets. The purpose of the comparison reports is to organize parameter and indicator data from the impaired site to enable comparison against corresponding data from least disturbed (LD) and nonimpaired (NI) stream reaches and site-specific comparator (SSC) sites. Physical and chemical parameters included are those commonly collected as part of Mississippi Department of Environmental Quality's (MDEQ's) routine data collection and assessment program. Biological data to be included are typically the overall Mississippi Benthic Index of Stream Quality (M-BISQ) score, biological metrics used in the calculation of the M-BISQ score, and the metrics found to be of use in scatterplot evaluations for the bioregion.

Conditions for LD and NI are calculated by using the interquartile range of values for all streams listed as LD or assessed as NI from the master database and located in the bioregion of the study site. Offsite stream-reach SSCs are sites located nearby, within the same 8-digit hydrologic unit code (HUC) and the same bioregion, with relatively similar drainage area, and with an M-BISQ score above the bioregion impairment threshold (Stribling and others, 2016).

Selection of Site-Specific Comparators (SSCs)

Several guidelines can support selection of a useful set of SSCs. First, no more than four SSCs are typically selected. A comparator site does not necessarily have to be a "high-quality" reference site. If a comparator site is not a part of the same aquatic system, it is important to ensure that, aside from the influence of anthropogenic stressors, the comparator and test sites are as similar as possible in terms of natural environmental factors (for example, elevation, stream size, drainage area, climate, slope, and geology).

The process of SSC selection involves using geospatial analysis (for example ArcGIS Esri software) with the following data layers: 8-digit HUC, bioregion, points representing all potential SSC sites (with associated data in the table), county, and National Hydrography Dataset stream coverage maps (U.S. Geological Survey, 2020). Starting with NI sites that are nearest the impaired site, perform a quick scan of similarly sized drainage areas to select three or ideally four SSCs that are relatively in the same size stream class as the impaired site. If more than four potential SSCs are in close proximity and of similar stream size, then other abiotic data (for example, substrate particle size or habitat assessment score) may be used to select a final set of four SSCs that are fairly typical for the given bioregion. Sometimes the four potential SSCs may be located relatively distantly from the impaired site, in which case effort should be made to select SSCs in the same major river basin, the same physiographic province, and most importantly, the same bioregion. The process of selecting SSCs is inherently subjective, and in many cases, there is no one "optimal" set of SSCs.

Stressor-Response Relations from Regional Data

Data from MDEQ's master database of biological data are used to derive stressor-response relations in the form of scatterplots. For sites with multiple samples obtained over time, only the most recent sample should be used.

Spearman's rank correlation coefficients (ρ) are calculated for relations between stressors (physical and chemical indicators) and responses (biological community metrics) and are stratified by bioregion. The 12 stressor-response relations with the highest absolute values of correlation coefficients are visually examined as scatterplots. Scatterplots with clearly discernible positive or negative relations are retained. In cases where biological metrics are highly autocorrelated ($\rho > 0.75$ for relations between biological metrics), the metric with the strongest correlation to the stressor is retained. For the retained group of scatterplots, the quantile regression curve (90th or 10th percentile, for positive and negative relations, respectively) is added to better visualize the response. The scatterplots are divided into quadrants by using the 25th or 75th percentile of a physical variable (the candidate stressor being evaluated) from reference sites and by using the 25th or 75th percentile of a biological response from reference sites, with the quadrants color-coded. The lower right quadrant is further divided by the intercept of the quantile regression and the 25th or 75th percentile of a biological response from reference sites. The result is a set number of scatterplots for each of the four bioregions in Mississippi (fig. 7), grouped by major category of proximate cause.

For more information about this publication, contact

Director, Lower Mississippi-Gulf Water Science Center
U.S. Geological Survey
640 Grassmere Park, Suite 100
Nashville, TN 37211

For additional information visit
https://www.usgs.gov/centers/lmg-water/

Publishing support provided by
Lafayette Publishing Service Center

Hicks and Cartwright—**Stressor Identification Framework of Biological Impairment in Mississippi Streams to Support Watershed Restoration**—SIR 2020–5098

ISSN 2328-0328 (online)
https://doi.org/10.3133/sir20205098