ENVIRONMENTAL LAW INSTITUTE®

AN INDEPENDENT, NON-PARTISAN ENVIRONMENTAL EDUCATION AND POLICY RESEARCH CENTER.

Evaluating the Water Quality Effects of TMDL Implementation

How States Have Done It and the Lessons Learned

This project is made possible through a cooperative agreement with the United States Environmental Protection Agency

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The Environmental Law Institute (ELI) gratefully acknowledges the Watershed Branch of the U.S. Environmental Protection Agency (EPA), Office of Wetlands, Oceans & Watersheds, for its support of this project, undertaken pursuant to Cooperative Agreement No. X7 - 83592101 - 0. ELI also would like to thank the members of its Technical Assistance Planning Group for this project, which included state, state association, and EPA representatives from monitoring, assessment and listing, TMDL, and nonpoint source programs: Pam Anderson, Jeff Berckes, Rosaura Conde, Cyd Curtis, Julie Espy, Julian Gonzalez, Jim Havard, Jasper Hobbs, Richard Mitchell, Eric Monschein, Pete Ode, Bill Richardson, Diane Switzer, and Alan Wittmuss. In addition, ELI extends its appreciation to the many respondents to the formal inquiries that were part of the research for this project as well as the many reviewers, including Emily Cira, Greg DeAngelo, Ben Rau, Marylou Renshaw, Robert Voss, Jeff White, and Meg Wiitala. ELI staff contributing to this project were Adam Schempp, John Hare-Grogg, and Jessye Waxman.

Except where expressly noted, the views expressed in this document should not be attributed to EPA, or to other federal or state agencies, nor should any official endorsement or approval be inferred. The information contained in this document was gathered from many and varied sources, and ELI alone is responsible for errors or inaccuracies.

I. Overview

Evaluating the water quality effects of Total Maximum Daily Load (TMDL) implementation is challenging but important. It is needed to reveal whether a TMDL and implementation actions are working or should be revised. It also provides information about restoration and protection, ideally making future efforts more effective and efficient, and it helps to demonstrate to the public the impact of the Clean Water Act (CWA) 303(d) Program and other CWA programs.

This document highlights the diversity of approaches to evaluating the water quality effects of TMDL implementation, explains some of those methods, and conveys lessons learned. It also details terminology challenges and identifies relevant resource materials. This document is intended to facilitate communication among water quality programs, especially CWA 303(d) programs, and help generate new ideas in this underexplored area. With the ten-year horizon of the 2013 CWA 303(d) Program Vision nearing its end and the next iteration of the Vision soon to begin, this is an opportune time to consider more thoroughly how the effects of TMDL implementation are and will be evaluated, particularly as it relates to the 2013 Vision Assessment Goal (likely to be the Data and Analysis Goal in the next Vision).

What follows is an organized summary of findings from primary and secondary research conducted by the Environmental Law Institute (ELI) on approaches to, experiences with, and resources that help in evaluating the effectiveness of TMDL implementation. The information about state practices comes primarily from answers to a questionnaire distributed in the fall of 2017 and completed by staff from 39 states and the District of Columbia (collectively referred to as "states" hereafter). Follow-up conversations and independent literature reviews by ELI staff supplement the information from the questionnaire, especially with regard to the section on terminology, specific examples from practice, and the Collection of Monitoring Resources. The information provided here is not intended to be comprehensive.

The next section of this document (Section II) contains a brief review of monitoring types relevant to TMDLs, highlighting the evolution and differences in the meaning of various terms and emphasizing the importance of being clear when communicating about monitoring. The lack of a common vocabulary for monitoring types highlights the value of explicitly defining monitoring terms when they are used, ideally providing a clear statement of the purpose of the monitoring. With this understanding, ELI's questionnaire, referenced above, did not put forward a common vocabulary from which to elicit responses and instead provided a statement of purpose: the intent of the questionnaire, and ultimately the compendium, is to collect "materials to aid state CWA 303(d) programs in supporting and developing data-gathering strategies to evaluate the water quality effects of TMDL-relevant restoration efforts." Likewise, this document does not set a common vocabulary to describe different types of monitoring, as it might mischaracterize the meaning behind state responses.

Section III of this document categorizes approaches to evaluating the water quality effects of TMDL implementation, as identified in the questionnaire, with examples for each of them. Some of these efforts have been site-specific while others have sought to be more geographically comprehensive; some have involved tailored monitoring and tools, and others have relied on existing processes and information. This diversity reflects differences in states' environmental

features, land uses, monitoring resources, and other characteristics. There is no one right way, and many states use or have used multiple approaches.

Section IV provides a wide variety of lessons learned. In general, states affirm the importance of developing strategic monitoring plans at the beginning of a project, and of collaborating and communicating with other programs and agencies as well as the public. Faced with the long time spans and uncertainties involved in demonstrating water quality improvements, several states have found it helpful to set interim measures of success, and to treat observed improvements with cautious optimism rather than conclusive satisfaction.

As demonstrated in Section V, states are finding creative ways to get the monitoring data necessary for effectiveness analyses. Some states' CWA 303(d) programs collaborate with other state water quality programs, other agencies, local governments, educational institutions, and nonprofit conservation groups to collect and share information. Other states are focusing on efforts to manage data more effectively, by establishing platforms to share information between programs and departments, and by using available data to improve the collection of additional data.

Section VI provides a brief description of and link to an annotated bibliography of resource materials relevant to analyzing the effectiveness of TMDLs. The bibliography identifies the general stages of the process addressed by each resource: planning, implementation, and analysis. It specifies the level of detail provided by each resource: basic, mid-level, detailed, and highly detailed. It also specifies where in each resource information on types of monitoring and specific sub-topics can be found and offers some general notes on each resource.

From all of this information, a few key takeaways appear. First, there are many approaches to evaluating the water quality effects of TMDL implementation; their usefulness can depend on the circumstances; and they can complement one another. Second, the more tailored the monitoring effort is to the purposes for which the data will be used, the more useful the data will be. Third, identifying data needs and working with potential sources of those data as early as possible increases the chances of getting more and better data for an effects evaluation. Fourth, estimating when TMDL implementation actions are likely to result in changes to water quality and other ecological indicators can be difficult, but it can help make for a more efficient process and set more realistic expectations. This timing will vary depending on the pollutant at issue, what actions were taken, and when the actions were taken, among other factors, so the estimate may be rough, but it still can provide value for designing and potentially modifying plans. Finally, patience and creativity, particularly in light of limited resources, are important.

II. Terminology

The CWA and supporting regulations do not include terms and definitions for the various types of monitoring that may occur to support all CWA programs, including the CWA 303(d) Program. Hence, for this compendium, key CWA 303(d) Program documents and related literature were reviewed to identify the types of monitoring that can support TMDL evaluation and adaptation. This review revealed that a variety of terms, and definitions of those terms, have been employed by various governmental and non-governmental entities to describe types of monitoring. The review also revealed that this variety can cause communication challenges. Below is a brief history of nationally relevant monitoring types that have been described with a direct or indirect connection to TMDLs. The history is not based on an exhaustive review of available program documents and literature, and thus is not necessarily complete. It simply is intended to illustrate the diversity of terms and the ways in which those terms have been used. This section concludes with recommendations for minimizing communication challenges that may occur due to differences in terminology within and among entities collaborating on TMDL-related monitoring projects.

A Brief History of Nationally Relevant Monitoring Types Related to TMDLs

Different types of monitoring that have an indirect or direct relationship to the CWA 303(d) Program have been described several times at the national level over the past four decades. In 1980, the U.S. Forest Service issued a technical guide to their hydrologists for preparing water quality monitoring plans. This early guidance, written by Potyondy, discusses two types of monitoring: (1) baseline and (2) project.[1] A document by Solomon concerning federal land nonpoint source management strategies was published almost a decade later. It discusses four types of monitoring: (1) implementation, (2) effectiveness, (3) validation, and (4) trend.[2] Table 1 depicts the types of monitoring identified in these and other documents discussed in this section.

Table 1. Examples of Terms Used to Describe Types of Monitoring

	Federal Resource				
Monitoring Term	**Potyondy 1980**	**Solomon 1989**	**MacDonald et al. 1991**	**EPA 1999**	**EPA 2013**
Baseline	✓		✓	✓	✓
Trend		✓	✓	✓	
Project	✓		✓	✓	
Effectiveness		✓	✓	✓	✓
Implementation		✓	✓	✓	
Compliance			✓		
Validation		✓	✓	✓	
Planning					✓

[1] J.P. POTYONDY, TECHNICAL GUIDE FOR PREPARING WATER QUALITY MONITORING PLANS 5-6 (1980).
[2] R. Solomon, *Implementing Nonpoint Source Control: Should BMPs Equal Standards?, in* PROCEEDINGS OF THE SYMPOSIUM: THE FORESTED WETLANDS OF THE SOUTHERN UNITED STATES, ORLANDO, FL JULY 12-14, 1988 155-162 (D.D. Hook and R. Lea ed., 1989).

In the early 1990s, MacDonald, et al. described in a guide seven different types of monitoring to support the design of water quality monitoring projects for forest management and streams in the Pacific Northwest and Alaska. The document, which was funded under an EPA grant, focused on a particular land use type; however, the authors took a broad perspective, and much of the information is widely applicable. The authors acknowledge that other federal and state agencies had described types of monitoring carried out by their respective agencies, including Potyondy, 1980, and Solomon, 1989, but they observe that the definitions are not consistent and often result in "semantic confusion."[3] To help address the confusion, the authors offer the following seven types of monitoring, including definitions and observations on how the types of monitoring are not mutually exclusive, depending on the scope of the project and the entities involved.

- **Baseline:** measurements used "to characterize existing water quality conditions and to establish a data base for planning or future comparisons." The authors indicate that some individuals refer to this type of monitoring as inventory monitoring or assessment monitoring while others use "baseline" monitoring to refer to long-term trend monitoring.
- **Trend:** measurements "made at regular, well-spaced time intervals in order to determine the long-term trend in a particular parameter." The authors indicate that measurements from trend monitoring are not taken specifically to evaluate management practices (as in effectiveness monitoring), management activities (as in project monitoring), water quality models (as in validation monitoring), or water quality standards (as in compliance monitoring); however, they acknowledge that trend data may be used for one or more of these other purposes.
- **Project:** measurements used to assess "the impact of a particular activity or project." The authors offer several examples of such activities and projects (e.g., timber sales, construction of a ski run, a fish habitat improvement project) and indicate that measurements would be taken in the stream channel to make upstream and downstream comparisons or before and after comparisons. The authors acknowledge that, because such comparisons may indicate the overall effectiveness of best management practices (BMPs) and other mitigation measures associated with a project, some entities consider "project" monitoring to be a subset of effectiveness monitoring.
- **Effectiveness:** measurements used "to evaluate whether the specified activities had the desired effect." The authors acknowledge that, "confusion arises over whether effectiveness monitoring should be limited to evaluating individual BMPs, or whether it can also be used to evaluate the total effect of an entire set of practices." The authors indicate that monitoring of individual BMPs (i.e., the "narrow" definition of effectiveness monitoring), such as spacing of water bars on skid trails, often occurs at, or immediately adjacent to, the BMP site, and therefore occurs outside of the stream channel. The authors offer that such measurements may be qualitative in order to determine whether proper implementation of BMPs protects water quality (e.g., observing road drainage problems during storm events). In contrast, monitoring of the overall effectiveness of BMPs (i.e., the "broad" definition of effectiveness monitoring) relies on measurements in the steam channel. The authors acknowledge that the distinction between the broader definition of "effectiveness" monitoring "becomes blurred" with other terms such as project and compliance monitoring.
- **Implementation:** measurements used "to assess whether activities were carried out as planned." The authors indicate that the most common type of implementation monitoring is

[3] LEE MACDONALD ET AL., MONITORING GUIDELINES TO EVALUATE EFFECTS OF FORESTRY ACTIVITIES ON STREAMS IN THE PACIFIC NORTHWEST AND ALASKA 6 (1991).

to determine whether BMPs were implemented as intended and does not involve water quality measurements. The authors also indicate that implementation monitoring is one of the few monitoring terms that "has a relatively widespread and consistent definition."

- **Compliance:** measurements used to "determine whether specific water quality criteria are being met." The authors acknowledge that, because compliance and trend data can help indicate the effectiveness of BMPs in protecting water quality, some entities consider them to be another type of effectiveness monitoring.
- **Validation:** measurements used to support "quantitative evaluation of a proposed water quality model to predict a particular water quality parameter." The authors acknowledge that other entities have considered "validation" monitoring to include validation of water quality standards, and indicate that the data set used for validation should be different from the data set used to construct and calibrate the model.[4]

The authors draw connections between these seven types of monitoring and the CWA, including TMDLs; however, a discussion of how the seven types of monitoring relate to TMDL evaluation and adaptation is not provided.

In the late 1990s and early 2000s, EPA issued "Protocol" documents for developing nutrient, sediment, and pathogen TMDLs that included recommendations for developing TMDL follow-up monitoring and evaluation plans. For example, the recommendations in the Nutrient Protocol document include descriptions of six different types of monitoring, modified from the terminology provided by MacDonald, et al. The types of monitoring and associated descriptions are presented below.

- **Baseline:** "describes existing conditions and provides a basis for future comparisons. This type of monitoring is not always necessary for the monitoring plan."
- **Trend:** "assesses the effectiveness of management actions and the changes in conditions over time relative to the baseline and identified target values. Trend monitoring is the primary type of follow-up monitoring, assuming the other elements of the TMDL are appropriately developed. It would address the changing conditions in the waterbody that results from TMDL-specific activities and other land management activities over time. This is the most critical component of the monitoring program, because it also serves to document progress toward achieving the desired water quality conditions."
- **Project or Effectiveness:** "Specific projects undertaken in the context of the TMDL or separate from the TMDL but potentially affecting water quality conditions for the watershed area under consideration should be monitored both to determine their immediate effects and the effects on the water quality downstream of the project."
- **Implementation:** "would ensure that identified management actions (such as specific BMPs or resource restoration or enhancement projects) are undertaken."
- **Validation:** "used to re-evaluate the selection of indicators, numeric targets, and source analysis methods."[5]

The Nutrient Protocol document notes that baseline monitoring may not always be necessary for TMDL follow-up monitoring and evaluation plans because some baseline data already exist (e.g., data and information that resulted in the waterbody being included on a state's CWA 303(d) list) and were considered during TMDL development. In contrast to MacDonald, et. al.,

[4] *Id.* at 6-7.
[5] U.S. ENVTL. PROTECTION AGENCY, PROTOCOL FOR DEVELOPING NUTRIENT TMDLS 78 (1999).

the Protocol document does not include the term "compliance monitoring," nor does it provide an explanation for not including the term.

In 2013, EPA announced a new collaborative framework for implementing the CWA 303(d) Program with states, territories, and tribes – *A Long-Term Vision for Assessment, Restoration and Protection under the Clean Water Act Section 303(d) Program*. This CWA 303(d) Program Vision contains six goals, including the Assessment Goal, which has the following milestones focused on baseline and effectiveness monitoring:

- "States develop plans to complete 'baseline' monitoring to gather needed data to assess pre-implementation conditions in priority areas."
- "States develop plans to complete 'effectiveness' monitoring to gather needed data to assess post-implementation conditions in priority areas."[6]

Although the milestones present baseline and effectiveness monitoring in quotations, the Vision document and supporting information do not provide definitions for these terms or reference documents that define the terms. At the National CWA 303(d) Training Workshop in 2015, working draft definitions were offered for these terms as well as an additional term (i.e., not identified in the Assessment Goal) to capture monitoring that may occur in-between baseline and effectiveness monitoring.

- **Baseline:** "performed to allow initial or ongoing assessment of ambient site-specific conditions."
- **Effectiveness:** "performed to assess ambient site-specific conditions post implementation activities."
- **Planning:** "performed, if needed, to support development of planning documents."[7]

The working draft definitions were based on EPA-state dialogue during development of the Assessment Goal in the Vision and were not intended to tie back to any historic definitions, such as those discussed above.

Recommendations

This brief history illustrates that, since the CWA was enacted, varied terms have been used to reference different types of monitoring, and a common vocabulary does not exist. As such, there is a very real possibility for miscommunication, not only within the CWA 303(d) Program but also with other programs, agencies, and organizations involved in TMDL evaluation and adaptation. To help minimize potential communication challenges, the following recommendations from MacDonald, et. al. are still relevant today. Those authors recommend that each monitoring plan explicitly define the monitoring terminology being used. Those authors also offer that, in "most cases[,] a clear statement of the purpose of the monitoring will be the best method of defining the type of monitoring, and it then is simply a matter of attaching a mutually agreeable label to that particular type of monitoring."[8]

---

[6] U.S. ENVTL. PROTECTION AGENCY, MEMORANDUM: A NEW LONG-TERM VISION FOR ASSESSMENT, RESTORATION, AND PROTECTION UNDER THE CLEAN WATER ACT SECTION 303(d) PROGRAM 7 (December 5, 2013).

[7] ENVTL. LAW INSTITUTE, 2015 NATIONAL TRAINING WORKSHOP ON CWA 303(D) LISTING & TMDLS, FINAL PROJECT REPORT & TRAINING WORKSHOP PROCEEDINGS 14 (2015).

[8] LEE MACDONALD ET AL., MONITORING GUIDELINES TO EVALUATE EFFECTS OF FORESTRY ACTIVITIES ON STREAMS IN THE PACIFIC NORTHWEST AND ALASKA 6 (1991).

Heeding the recommendation above, this document does not attempt to establish a common vocabulary to describe different types of monitoring that may be used to support TMDL evaluation and adaptation. To do so may have the unintended consequence of mischaracterizing the meaning behind the states' responses to the questionnaire conducted. The questionnaire sent to states did not put forward a common vocabulary from which to elicit responses and instead provided a "statement of purpose" (in the manner of MacDonald, et al., 1990) for the type(s) of monitoring being considered for the questionnaire. Specifically, the questionnaire indicated that the intent of the questionnaire, and ultimately the compendium, is to collect "materials to aid state CWA 303(d) programs in supporting and developing data-gathering strategies to evaluate the water quality effects of TMDL-relevant restoration efforts." The questionnaire also noted that, "If one views the TMDL process as being (1) the identification of impairment, (2) development of a TMDL, (3) implementation of the TMDL, and (4) evaluation and adaptation of that implementation, this project is focusing on how the last of those steps is accomplished." Furthermore, regarding TMDLs, the questionnaire asked, "In what ways is your state evaluating the effects of TMDL-influenced restoration efforts on water quality? (How are you determining whether progress is being made, whether and how restoration actions should be changed, whether and how TMDLs may need to be revised, and/or whether the objective/standard has been met?)." Throughout the remainder of this document, please consider that the terminology used by the respondents may not always be consistent with that used by your program, agency, or organization.

III. Types of Approaches

States have evaluated the effects of TMDL implementation in a wide variety of ways. The approach chosen and when and where monitoring is conducted commonly depend on the resources available and the circumstances of the specific restoration effort, but some states have built one or more approaches to evaluating TMDL implementation into their water quality assessment and restoration processes. Based on responses to ELI's 2017 questionnaire, below are brief descriptions of some of the approaches that states have taken, along with commentary and examples.

Using the Existing Water Quality Monitoring and Assessment Processes

Some of the most common means of evaluating the effects of TMDL implementation are through a state's existing water quality monitoring processes:

- Many states that implement a rotating basin monitoring strategy, as referenced in their questionnaire responses, use this approach to regularly monitor waterbodies or watersheds as a means of evaluating the effects of TMDL implementation activities. Coordination between the TMDL program and the monitoring program can improve the value of the data generated and its application to evaluating TMDL implementation effectiveness. Even with that coordination, this approach may provide less of the targeted data than other approaches. For example, a rotating basin strategy results in monitoring conducted at intervals dictated by the rotation schedule, as opposed to immediately before, during, and after the implementation actions or at other moments, when effects may be more observable.

- Several states, with and without a rotating basin monitoring strategy, referenced reliance on fixed monitoring stations for data relevant to evaluating the effects of TMDL implementation. The fixed stations can provide more regular data than the rotating basin monitoring strategy, but they may cover fewer TMDL waters. Furthermore, the locations of the fixed stations may not be ideal for an analysis of the impacts of the TMDL at issue, unless their siting is selected in part because of the TMDL.

- Some states noted in their questionnaire responses that the state's monitoring program's strategy includes effectiveness sampling after TMDL implementation has begun. Communication between the TMDL and monitoring programs is important in these instances as well, to identify shared priorities and when water quality improvements likely are to occur.

- Some states noted that they require permittees to sample the receiving water, as well as the effluent, which provides additional data that can be helpful in evaluating the effectiveness of TMDL implementation efforts. One of the practical challenges to this approach can be the accessibility of the data. Ideally, it should be stored in the same location as other water quality data.

Regardless of the monitoring strategy, coordination between TMDL and assessment staff is also likely necessary to evaluate the effects of TMDL implementation. Water quality assessment can include evaluation of the effect of corrective actions, such as TMDL implementation, but the

evaluation commonly is binary—a water meets or does not meet a standard. Additional analysis of the data beyond basic assessment purposes may be needed to identify trends and improvement short of meeting a standard. Knowing whether the effort is working, even if it is not yet fully successful, is important when evaluating the effects of TMDL implementation.

ELI Note: The existing water quality monitoring and assessment processes can be a good foundation or supplement to other approaches to evaluating the effects of TMDL implementation, providing breadth of coverage to supplement depth of coverage for priority sites through other approaches.

Missouri's Monitoring and Assessment Processes

The Missouri Department of Natural Resources (MoDNR) relies in part on its water quality assessments for biennial integrated reporting to evaluate progress toward meeting TMDL goals. Among other objectives, the annual monitoring schedule prioritizes monitoring in areas where restoration efforts, for point and/or nonpoint sources, are known to have occurred. The anticipated lag time between those restoration efforts and water quality improvements is carefully considered when prioritizing waters for effectiveness monitoring, so there is a higher likelihood of detecting water quality changes. MoDNR also analyzes the pre-implementation data and its variability to see how much new data would be needed to show a statistical difference.

For nonpoint source restoration efforts, the calculation as to when and what monitoring is necessary to observe water quality improvements can be challenging and depends on the pollutant, the scope of the problem, the severity of the impairment, and the scope of the restoration effort. Where major restoration efforts have occurred, MoDNR tends to do more intensive sampling in a shorter period. Where voluntary efforts will occur over several years, it may do routine sampling over a long period or wait until a critical mass of BMPs have occurred before sampling again. From this monitoring, MoDNR can identify improvements and the restoration of uses, as well as the degradation of restoration efforts.

From a NDPES perspective, MoDNR schedules intensive monitoring at several locations upstream and downstream of a facility over a 48-hour period during mid- to late-summer low flow conditions to determine if the receiving stream is meeting criteria or established 4b or TMDL targets. If not, the multiple locations sampled help determine the extent of impairment and whether the water quality restoration targets are still appropriate for attaining standards, as well as potentially to model the stream, calculate new permit limits, and evaluate the need to revise the existing TMDL. This facility will then look at plant optimization, construction of plant upgrades, regionalization, or moving to a no discharge/land application operation. MoDNR will work with the facility via permitting or the SRF program to evaluate and assist the facility with upgrades. Once the facility optimizations or upgrades are complete, MoDNR will again schedule intensive monitoring to assess the status of the water. Depending upon the original impairment listing, MoDNR also may schedule invertebrate monitoring to determine use attainment.

Leveraging Information from Effectiveness Evaluations of CWA 319-Funded Projects

Many states referenced in their questionnaire responses the value of the monitoring and analysis associated with CWA Section 319 water quality restoration projects when evaluating the effects of TMDL implementation. EPA's CWA 319 grant guidelines emphasize the importance of watershed-based plans developed and implemented with CWA 319 funds to include nine minimum elements, the last two of which are: (1) identify indicators to measure progress; and (2) develop a monitoring component. When the execution of a watershed-based plan implements a TMDL, in whole or in part, the indicators identified, monitoring conducted, and conclusions reached have the added benefit of identifying some of the effects of the TMDL implementation.

ELI Note: Not all TMDLs have CWA 319-funded projects associated with them. Even where there is overlap, the monitoring data might not be exactly what the CWA 303(d) Program is seeking, and the interim and ultimate objectives for the waterbody might differ among the two programs. Yet, coordination between TMDL and Nonpoint Source Program staff could help address some of these issues.

TMDL Implementation Analysis via Maryland's CWA 319 Program

Much of the evaluation of TMDL implementation effects in Maryland stem from CWA 319-related activities. Ten of the state's watershed-based plans being implemented as of early 2018 incorporate TMDL load allocations as watershed plan goals. In these instances, progress tracking includes BMP implementation and load reduction estimates. Progress is regularly reported to EPA through the state's CWA 319 Nonpoint Source Program Annual Report. Execution of watershed-based plans in Maryland has led to several waterbodies meeting water quality standards and revealed these accomplishments. For example, the Casselman River watershed was listed for low pH in 1996, and a TMDL concerning that impairment was first approved in 2008. Acid mine drainage remediation as part of watershed plan implementation has resulted in meeting water quality standards for pH, and thus the goals of the TMDL, for several stream segments in the watershed.

Developing Location-Specific Plans in the TMDL Implementation Plan

Some states noted in their questionnaire responses that they detail their approach to evaluating the effects of TMDL implementation in the TMDL or TMDL implementation plan itself. While many states reference "effectiveness monitoring" in at least some of their TMDLs or associated implementation plans, the breadth and depth of specifics regarding monitoring and the assessment of monitoring data – how, when, where, and by whom – vary significantly across states and across TMDLs within a state. [See examples of how states have covered this material here.]

ELI Note: This approach has the advantage of being tailored to the particular circumstances of each TMDL, with TMDL staff identifying the data they would need, and ideally when they would need it, to determine whether TMDL implementation is effective. In addition, if monitoring and evaluation plans are a regular part of TMDL development, the approach is rather comprehensive in its coverage of waters with approved TMDLs. Yet, coordination with monitoring and assessment staff, and probably other entities for implementation and monitoring help, will be necessary for the plan to be well-designed and fully executed. Developing and executing these plans can be resource intensive.

Florida's BMAPs

In consultation with local stakeholders, the Florida Department of Environmental Protection (FDEP) creates TMDL implementation plans known as Basin Management Action Plans (BMAPs), establishing a framework for water quality improvement projects and management strategies to address key TMDLs. As of 2021, over 30 BMAPs had been adopted or were under development. The BMAPs cover large, multi-county areas and typically address multiple TMDLs. A representative example is the Lower St. Johns River BMAP, which was developed in 2008 for dissolved oxygen for the river's marine reach and for chlorophyll *a* for its freshwater reach. The BMAP includes a highly detailed, basin-specific effectiveness monitoring plan, as well as adaptive management measures. It describes how water quality data will be collected and used to assess whether there is reasonable progress in BMAP implementation and water quality improvement. State law requires FDEP to review and potentially update the BMAPs every five years. The BMAP monitoring plan builds upon existing ambient monitoring stations and lists additional points where monitoring should be instituted. The BMAP describes how the number and locations of sampling sites were determined using statistical power analysis. It also explicitly refers to data management procedures, including the acceptance and vetting of external partners' data. In addition to FDEP, watershed management districts, utility authorities, and other entities are listed as responsible for monitoring at specific sites. The Lower St. Johns River Main Stem BMAP can be found at: https://floridadep.gov/sites/default/files/adopted-lsjr-bmap.pdf.

Establishing a Collaborative Monitoring Program

A few states highlighted in their questionnaire responses state programs specifically created to plan and execute monitoring of the effectiveness of specified efforts, often done collaboratively with partners. While the programs have different areas of focus, from nonpoint source-caused impairments to the work of various partners, they tend to identify progress in water quality restoration beyond just those efforts implementing a TMDL.

ELI Note: A collaborative monitoring program focused on effectiveness can help elevate the significance of this type of monitoring among the many water quality monitoring needs, engrain

it in monitoring plans, and establish a coordinated process for prioritizing such monitoring needs of multiple governmental and non-governmental entities.

Indiana's Performance Measures Monitoring Program

The Indiana Department of Environmental Management (IDEM) has developed a collaborative effectiveness monitoring approach, concentrating monitoring resources at locations where there is the greatest chance of seeing water quality improvements in the state. The effectiveness monitoring is not tied to specific TMDLs; rather, IDEM develops an annual statewide *Performance Measures Monitoring Plan*. Through this plan, IDEM identifies waterbodies where restoration activities have occurred (including work done through CWA 319 grants, USDA Farm Bill programs, and state water quality programs like Clean Water Indiana and the Lake and River Enhancement program) and where results are most likely to be seen. IDEM identifies these waterbodies in consultation with the Indiana Conservation Partnership, which includes soil and water conservation districts, the USDA Natural Resources Conservation Service, the Purdue Cooperative Extension Service, and other agencies and organizations. IDEM samples many sites a year during the typical field season, April to October, and for efficiency, prioritizes locations within the probabilistic river basin for the year. The collaborative approach helps IDEM meet EPA performance measures, but also helps it determine if restoration efforts are indeed successful.

Indiana's surface water quality monitoring work plans, including the performance measures monitoring work plan, can be accessed on the IDEM website at https://www.in.gov/idem/cleanwater/surface-water-monitoring/.

Adaptive Management through "Effectiveness Monitoring" Reports

One state referenced in its questionnaire response its "effectiveness monitoring" reports. The state's guidance document for TMDL "effectiveness monitoring" suggests that these reports include explanations of the history of water quality in the waterbody at issue, the restoration activities undertaken, the current water quality conditions, the results of data analyses, conclusions, and recommendations for subsequent action. The guidance also clarifies that "effectiveness monitoring" reporting is not necessarily a single, final report once all implementation has occurred and many of the results manifested; it can inform water quality managers and the public throughout the course of implementation and outcome realization. Such an iterative process provides transparency and aids adaptive management.

ELI Note: While not a stand-alone approach, reports like these can provide a valuable subsequent step to plans for evaluating the effects of TMDL implementation, and better ensure the execution of those plans.

Washington's Effectiveness Monitoring Reports

The Washington Department of Ecology (Ecology) has developed effectiveness monitoring reports following TMDL implementation for several waterbodies in the state. These publications, since 2002, are cataloged at: https://fortress.wa.gov/ecy/publications/UIPages/PublicationList.aspx?IndexTypeName=Topic&NameValue=Effectiveness+Monitoring+for+Water+Quality+Improvement+Projects+(TMDLs)&DocumentTypeName=Publication. Ecology's guidance document for TMDL effectiveness monitoring, available here: https://fortress.wa.gov/ecy/publications/documents/1303024.pdf, includes a section focused on reporting.

An example of such a report is the *Henderson Inlet Fecal Coliform Total Maximum Daily Load Water Quality Effectiveness Monitoring Report*. It details the water quality classifications, regulatory criteria, and land use characteristics of the relevant area; the general goal and specific objectives of the data analysis; what data were used; and the results of the analysis. The report focuses on assessing progress, not just whether water quality standards were met. It includes TMDL implementation actions undertaken, trends in water quality, and links between implementation and water quality changes, ultimately identifying several key conclusions and recommending specific actions going forward. The report can be found at: https://fortress.wa.gov/ecy/publications/documents/1703001.pdf.

IV. Lessons from Experience

States have drawn a variety of lessons from their experiences evaluating the effects of TMDL-relevant water quality restoration efforts. These lessons are largely harmonious but address many different challenges. In response to the fall 2017 questionnaire by ELI, several states noted that maximizing the value of monitoring requires identifying the purposes of monitoring, and planning accordingly, before beginning project implementation or data collection. Also, several states indicated that effective communication and collaboration with data collection partners, and with the general public, are important characteristics of an effective program. States cautioned against underestimating funding constraints, data management needs, and time periods needed to observe water quality improvements, and provided some insights to help overcome those challenges. Below is an organized list of many of the lessons learned that were noted in the responses. The terms used in these statements are those provided by the questionnaire respondents; their meanings may not match interpretations noted in Section II above.

Identify objectives, needs, and procedures at the outset.

- Data needs must be determined early in the process of watershed planning, and a monitoring scheme developed prior to project implementation. (NE)
- Clearly define the water quality and data quality objectives prior to data collection for successful results. (MS)
- A long-term monitoring/evaluation plan is necessary with any project aimed at the restoration or improvement of water quality. This requires dedicated individuals and funding. (MO)
- Initially define the scope of the assessment relative to flow, as the results of restoration efforts may be noticeable at normal/base flows but not at high flows. (KS)
- Listing methodologies should be considered, as well as method detection limits relative to applicable water quality standards. (DC)
- Consider the amount of data needed to show improvement. If delisting a water is the ultimate goal, the data requirements to show attainment must be incorporated into the sample plan. (AZ)
- Put time into study questions, study design, and tool and method development in order to collect the right data. (OR)
- Delay robust monitoring until significant progress implementing restoration activities has been made. (KS)

Monitor strategically.

- The monitoring plan should be designed to provide pre-project baseline conditions and comparable post-implementation data. Monitoring sites should be strategically positioned to isolate the area(s) of treatment. If a rotation scheme is planned for long-range watershed treatment, monitoring sites should be positioned to isolate future areas of treatment as well. (NE)
- Collect detailed landscape condition and practice information in order to better understand or predict water quality response from specific actions. (OR)
- If ideal sampling frequencies are too costly, target sampling activities to capture samples during the desired flow conditions and season. (KS)

- Collect and evaluate data from a variety of hydrological years to account for the dry year/wet year effects, since improvements may be overstated during dry years. (KS)

Monitoring the effectiveness of implementation efforts can be resource intensive.

- Be prepared to devote a significant amount of time and resources to the monitoring effort. I estimate project monitoring consumes approximately 20 percent of our TMDL staff's time but is well worth the investment weighed against the importance of the data for documenting effectiveness and soliciting adoption of restoration goals by partners. (UT)
- Gathering effectiveness data requires extensive effort and planning, and managers need to consider program capacity to do the work. (IN)
- Depending on the size and scope of the remedial efforts, effectiveness monitoring may be as resource intensive as development sampling for the TMDL or alternative plan. (AZ)
- Funding and staffing for post-implementation monitoring is generally a much lower priority then getting funding for implementation projects. (ID)
- Typically, there is not enough data, which is a function of staffing, resources, and competing priorities. (KS)
- There is not nearly enough money to do the amount of monitoring we need to do to tell an accurate story about the state of our waters and the effects of implementation. (WA)
- Restrictions in staffing and funding make it essential that time and resources are used efficiently. (NM)

Monitoring by partners can be critical to determining the effectiveness of implementation efforts.

- Collaboration has been essential to our program for post-TMDL monitoring, whether with a university, federal agency, permitted entity, or municipality. (MA)
- We have to work collaboratively with other groups to collect adequate data. (CO)
- Collaboration with consultants or qualified non-profits can be very helpful. (ME)

Teaching partners how to collect samples and the importance of those samples to the overall process and objective can be worth the effort.

- We have found it useful to train local project sponsors to collect the water quality samples. Generally, NDEQ staff collects samples the first year and then train staff of the project sponsor to collect samples in subsequent years. They are trained to collect and record field data and to collect and ship water samples to NDEQ for analysis. This added efficiency allows sampling more often than agency staff can accommodate. A side benefit is that the experience encourages local sponsors to develop their own monitoring programs. (NE)
- The better understanding that sample collectors have of what it is they are to collect, the more motivated they are to gather useful information. Some coordinators have been shown how to collect samples and told when to collect them, but they end up losing interest and quit or find excuses for why they were not collected. If the coordinator has a vested interest in what their project is doing and know what the goal is, they will make greater efforts in what they do leading to a successful project. (SD)

Communication with stakeholders and the public is important.

- Water quality restoration is ultimately achieved when the stakeholders/public are given the opportunity, through public outreach, to weigh in on the project. Transparency and action lead to buy in and a vested interest in the outcome of the plan. Sustainability is achieved when stakeholders take the initiative to continue implementation and protection of a watershed long after the plan has been developed. (TX)
- Engage with local governments and stakeholders early in the process. (LA)
- It is important to have communication with stakeholders, to know the status, timing, and location of TMDL implementation efforts. (AL)

Collaboration among programs and agencies is important.

- Collaboration is needed to ensure data collection meets standards required for assessment. (NM)
- It is necessary to have good cooperation and coordination among TMDL, monitoring, and nonpoint source program staff. It is also necessary that staff from these programs are looking for potential "success stories", because sometimes restoration actions may occur but may not be directly linked to the TMDL. (MI)
- Communication between TMDL, NPS, and monitoring staff is essential. We have an individual from the monitoring program, the TMDL program, and NPS program that consult together on choosing waterbodies to sample, what parameters and methods to sample, and assessment of the data, as well as communicating successes to US EPA. (IN)
- Different programs or entities often have different reasons and goals for implementing projects; attainment of surface water quality standards may not be an explicit goal of the USFS, for example. In those instances, we may have to settle for improved water quality but not attainment. (AZ)

Data management is important.

- Data management is at least of equal importance to data collection. (NM)
- Colorado has a data sharing network (CDSN) to provide an accessible and affordable way to manage, analyze, and share data. The state initially helped fund the project through federal grants, and the CDSN continues now with fees and donations by users. (CO)
- Changes in monitoring strategies, labs, or QA protocols can affect evaluations. It is best to try and maintain some minimum level of standardization or data flow to ensure consistent evaluations. (FL)

Water quality improvement, let alone restoration, takes time and can be hard to prove.

- To evaluate any positive results of implementation takes time. No best management practice will produce results overnight. (MO)
- It takes time to see results, especially if the primary measure of water quality is bio-assessment. (ID)
- We have found that it has taken well over ten years before the impact of BMP implementation can be seen. (GA)
- Be prepared for disappointment – even if several watersheds per year are sampled, and even when lag time for BMP effectiveness has been taken into consideration – water quality restoration simply takes a long time to be accomplished. (IN)

- The lag time associated with water quality improvements is often difficult for those implementing (or funding) BMPs to understand. (KS)
- The levels of noise for most parameters of interest in a water quality dataset makes declaring improvement in a waterbody challenging. It is better to be cautiously optimistic about the trend being seen than to declare a successful restoration project. (KS)

Water quality improvement milestones help with long lag times.

- Interim measures of success are necessary. (NC)
- Most of our impairment problems will take actions by many over a long time, and some of us are thinking that we need indicators during the transition period that will have relevance with landowners and watershed group personnel as well as being cost-effective. We are exploring riparian health indicators at a watershed scale to help track progress in many locations (as well as nutrient data in places); this is consistent with many of the indicator measures used to help develop allocations in TMDLs. (MT)

V. Approaches for Getting Data to Evaluate TMDL Effectiveness

States reported in ELI's 2017 questionnaire a variety of routine and less common means of obtaining water quality data for evaluating TMDL effectiveness.

Routine Approaches

Many states evaluate the effectiveness of restoration efforts broadly through the information provided by routine monitoring of the monitoring program, via long-term monitoring stations, basin-rotation monitoring, etc. Commonly, these data inform water quality assessments, and "successful" restoration or protection is indicated through the details of biennial Integrated Reports. In addition, activities of state nonpoint source programs, notably via 9-element watershed plans and annual program reports, include plans for monitoring effectiveness, and resulting data can be relevant to TMDLs.

Less Common Approaches

In addition to these routine approaches, what follows is an organized list of some of the less common means by which states procure the data necessary to evaluate the water quality effects of TMDL-relevant restoration efforts, as noted in responses to the questionnaire. The terms used in these statements are those provided by questionnaire respondents; their meanings may not match interpretations noted in Section II above.

Effectiveness Program

- Indiana: The Performance Monitoring program was specifically developed to look at previous impairments where restoration work (or other water pollution control measure, like a new permitted facility) is known to have occurred, and re-sample the CWA 303(d)-listed stream segments. This captures not only "TMDL-influenced restoration," but also work done through CWA 319 grants and partner programs (like USDA Farm Bill programs and state water quality programs like the Clean Water Indiana and the Lake and River Enhancement Program). The state samples multiple sites a year during the typical field season (April-October), and for efficiency, prioritizes locations within the probabilistic river basin for the year.
- New Mexico: The Surface Water Quality Bureau Watershed Protection Section has implemented an effectiveness monitoring program that collects baseline water quality information for comparison to post-treatment conditions. This program is used specifically for nonpoint source impairments and improvements.

Formal Collaborations

- Alaska: DEC participates in Fish Habitat Partnerships throughout the state, with said groups often looking at the same high priority waters as the state.
- Indiana: Among other means, IDEM attempts to evaluate restoration efforts by analyzing calculated load reductions from BMPs based on 12-digit HUC watersheds. As part of the Indiana Conservation Partnership, the agency is able to get partner restoration practices and load reductions aggregated on the 12-digit HUC. This includes practices from the Indiana State Department of Agriculture, the Indiana Department of Natural Resources, IDEM's NPS Program, and USDA Natural Resources Conservation Service and Farm

Service Agency Farm Bill programs. Load reductions are calculated using the EPA Region 5 Model.

- Oregon: The Conservation Effectiveness Partnership (CEP) is a collaboration between multiple state agencies and the Natural Resource Conservation Service (NRCS). The purpose of the collaboration is to evaluate how management or restoration actions have contributed to water quality improvement. A number of studies that include both monitoring and/or modeling have been initiated in select areas. As part of this collaboration, the state has a data sharing MOA with NRCS that allows NRCS to share practice data with the state as long as it is used only to evaluate conservation effectiveness.

Informal Collaborations

- Florida: DEP coordinates with local entities that collect data, occasionally providing some financial support or laboratory services.
- Massachusetts: DEP is identifying locations and monitoring needs and working with partners to develop effectiveness monitoring plans.
- Minnesota: Water quality monitoring by local units of government is sometimes completed in a fashion that allows an evaluation of water quality change.
- Missouri: In order to share data collection efforts, DNR collaborates with partners and other agencies where areas of interest overlap.
- Montana: Overall, DEQ has good working relationships with many watershed groups and thus maintain an awareness of TMDL and other nonpoint source implementation actions in the state's priority watersheds. DEQ routinely meets with the U.S. Forest Service, which has documented multiple successful BMP and watershed restoration projects, and DEQ is in communication with many BLM offices that routinely share BMP implementation plans and activities.
- Ohio: Ohio EPA has supported and/or partnered with local entities, such as Northeast Ohio Regional Sewer District, Cincinnati MSD, Clermont County, and Miami Conservancy District, to do their own stream monitoring and share the results with the agency. Some of the data is Level 3 credible, so it is useful for the Integrated Report.
- Utah: The Division of Water Quality is currently expanding the monitoring of nonpoint source projects through work plans developed cooperatively with partner agencies, including local conservation districts, the State Department of Agriculture and Food, NRCS, Division of Wildlife Resources, and volunteers.

Citizen Science

- Arizona: ADEQ has been expanding its citizen science program to include effectiveness monitoring.
- Kansas: KDHE has evaluated additional data collected independently by the WRAPS groups.
- Missouri: DNR recruits and trains volunteers to collect data through its volunteer water quality monitoring program.
- Montana: DEQ supports some monitoring networks on larger rivers and some volunteer monitoring.

Data Management

- Alaska: The state CWA 303(d) program is starting to have conversations with the state permitting program about putting ambient water quality data in the same place (STORET).
- Ohio: A few years ago, Ohio EPA worked with state and federal partners to try to increase the amount of data being provided to the Water Quality Portal, for easier analysis and compilation. The effort met with limited success for a variety of reasons, but they are still trying to accomplish it, since they think it would help make their evaluations more robust.
- Washington: Washington is developing a TMDL implementation database to specifically track implementation of the nonpoint source practices intended to comply with TMDL load allocations. They already have a system to track NPDES permits. The objective of the database is to track implementation so they know where more work is needed or can tell when enough implementation has occurred that it is logical to do some water monitoring, to see if the waterbody has improved.

Monitoring Program Procedures

- District of Columbia: The CWA 303(d) program provides ongoing input to monitoring strategy revisions.
- Georgia: Targeted sampling stations often are located on CWA 303(d) listed segments where TMDLs and TMDL implementation plans have been prepared, to determine if improvements in water quality have occurred.
- Michigan: Water quality monitoring staff identify impaired waters and approved TMDLs in planning watershed surveys, and sometimes will conduct follow-up monitoring to see if the impairment is still present and/or whether there has been an improvement in water quality. Generally, this follow-up monitoring will only occur if actions have been taken and there is reason to believe that water quality may have changed.
- Missouri: As part of the monitoring and assessment process, the scheduling of annual monitoring efforts prioritizes monitoring in areas where restoration efforts are known to have occurred (e.g., CWA 319 projects and land reclamation projects).
- Utah: CWA 303(d) program staff regularly participate in the intensive monitoring site selection process and coordinate partner agencies' cooperative monitoring efforts to ensure necessary data are collected to evaluate water quality effects of restoration efforts.

Requiring or Encouraging the Monitoring of Effectiveness by Point Sources

- Kansas: KDHE conditions permits on providing appropriate monitoring to evaluate TMDL implementation efforts pertaining to point sources.
- Minnesota: Permittees often are required to sample the receiving water.
- Missouri: Collaborative agreements with permittees have monitoring components to evaluate progress in meeting TMDL goals.

Requesting Monitoring

- Connecticut: For some TMDLs, once implementation begins, the CWA 303(d) program requests monitoring data from the state monitoring program.
- Idaho: During the process of TMDL development, DEQ meets with a watershed advisory group and other federal and state agencies, requesting their existing data and encouraging them to collect additional data.
- Maryland: MDE is encouraging Phase I MS4 counties to collect and report habitat assessment data for randomly selected stations in non-tidal sediment TMDL watersheds. In waters with bacteria TMDLs, MDE is encouraging MS4 jurisdictions to set up long-term trend monitoring stations at the same station locations as those used for TMDL development. MDE also is encouraging MS4s to create one or two long-term monitoring stations in highly urbanized watersheds where winter salt application is likely. MDE will analyze these data for water quality trends, ideally indicating improving conditions resulting from stormwater BMPs, focused efforts to reduce human sources of bacteria, and salt management plans.

Requiring Data of Grant and Loan Recipients

- Washington: The Department of Ecology requires grant and loan recipients to provide the data collected following an agency-approved QAPP.

Requesting Data

- Connecticut: DEEP asks people to provide information on their implementation activities.
- Mississippi: DEQ solicits all available information collected by its partners through Mississippi's Basin Management Approach.
- Nebraska: Through the watershed planning process, DEQ requests copies of any studies or projects its partners have conducted that include water quality data or other relevant information that may help determine current watershed and water quality conditions and the potential for improving both. This is done through general requests for information and one-on-one contacts.
- New Mexico: Every two years, the Monitoring, Assessment, and Standards Section sends out a request for data to be included in the updated assessment. Data submitted through the call is verified through a QA process before being accepted. The Watershed Protection Section Effectiveness Coordinator works with cooperators, contractors, and agency partners to provide training and support for outside data collection.
- West Virginia: DEP requests submission of third-party water quality data as part of its solicitation of data for CWA 303(d) list development every two years.

VI. Collection of Monitoring Resources

An annotated bibliography of monitoring resources that may help inform evaluation of water quality restoration plan effectiveness is available here. In addition to the title of the resource, the list details the stage(s) of the monitoring process addressed by the resource, the level of detail provided by the resource, the type(s) of monitoring referenced in the resource (using the definitions in MacDonald, et al., 1991), the topic(s) that the resource addresses, and general notes about what is covered in each resource. The list is not intended to be comprehensive.