Recommendations for Developing TMDL Effectiveness Monitoring Plans

Background

Section 303(d) of the federal Clean Water Act (CWA) requires states to identify all waters that do not meet water quality standards. Water bodies not meeting water quality standards are considered impaired and are placed on a state's 303(d) list, which is submitted to the U.S. Environmental Protection Agency (EPA) for review and approval. Total Maximum Daily Loads (TMDLs) must be developed for all water bodies on the approved 303(d) list. Development of a TMDL is only the beginning of the process that will ultimately lead to water body restoration. TMDL *implementation* is a critical step in the water body recovery process, and requires a concerted effort from all stakeholders - local, state, and federal. TMDL implementation can include a vast array of policies and on-the-ground management actions intended to restore previously impaired water bodies to conditions allowing them to meet designated uses. Even with the enormous efforts and resources invested by citizens and government at all levels, very few impaired waters have been fully restored. This is due to a number of factors, including:

1. An historical focus on the "worst first" approach to TMDL development;
2. The significant lag time between removal or reduction of a pollutant source and the corresponding response in the water body; and
3. The lack of regulatory authority over most nonpoint sources (NPS) of pollution.

With more than 44,000 TMDLs developed nationwide, and limited successes reported, the importance of demonstrating incremental improvements in water quality cannot be overstated. TMDL effectiveness monitoring helps to measure progress towards attainment of water quality standards and informs future management actions. Documenting improvements in water quality is important because demonstrated success, or effectiveness, is often tied to funding; further, a lack of demonstrated success can undermine the scientific credibility and hard work of the many stakeholders involved in TMDL development and implementation. In a review of stream restoration project monitoring activities in the Pacific Northwest, 64% of project managers indicated that, in retrospect, they would have used a more methodical monitoring design that allowed for a scientific evaluation of project effectiveness, citing future funding as a benefit to such an approach[1]. Washington Department of Ecology describes the benefits of effectiveness monitoring as follows:

1. Provides a measure of progress toward implementation of recommendations (i.e., how much watershed restoration has been achieved, how much more effort is required);
2. Supports the decision-making process on more efficient allocation of funding and optimization in planning (i.e., identifying recommendations or restoration activities that worked, which restoration activity achieved the most success for the resources); and

[1] Jeanne M. Rumps, Stephen L. Katz, Katie Barnas, Mark D. Morehead, Robin Jenkinson, Stephen R. Clayton, and Peter Goodwin. (2007). Stream Restoration in the Pacific Northwest: Analysis of Interviews with Project Managers. Restoration Ecology Vol. 15, No. 3, pp. 506–515.

3. Provides technical feedback that is useful for refinements of initial modeling analysis (e.g., for a TMDL) or refinements in the planning of best management practices (BMPs), permits, and other pollutant reduction and watershed restoration strategies.

Most states rely on their current ambient monitoring network to evaluate TMDL effectiveness. However, statewide ambient water quality monitoring networks are generally not set up to provide a significant amount of data on targeted water bodies; rather, ambient monitoring networks provide basic data for many different uses, including supporting the development of water quality criteria, reporting on the condition of the state's waters, and identifying impaired waters. As a result, ambient monitoring networks are typically designed to characterize water quality conditions at a broad scale. This approach is not consistent with the data needs of a TMDL effectiveness monitoring program, which requires targeted, more rigorous sample collection of specific parameters and information on designated use support on a smaller (i.e., subwatershed) scale.

EPA's current strategic plan calls for the improvement of water quality in impaired watersheds using the watershed approach (Measure SP-12) and the full or partial restoration of water bodies that are primarily NPS-impaired (Measure WQ-10). The strategic plan refers to these targets as National Water Program Guidance Measures. Measure SP-12 is a "demonstration" measure, used to document water quality successes that result from application of the watershed approach; the intent is not to inventory all instances where success has occurred or is underway in a state. Measure WQ-10 is the main long-term environmental results measure for the NPS program and requires that a designated use be restored or that one or more pollutants causing impairment of a designated use meet applicable criteria.

For a watershed to be counted under SP-12, states can use one of three options for demonstrating water quality improvement. Under Option 1, states must demonstrate improvement as the removal of one or more of the impairment causes identified in 2002 for at least 40 percent of the impaired water bodies or impaired miles/acres. Under Option 2a, states must use valid scientific information and statistical procedures to demonstrate that significant improvement has occurred with a 90 percent or greater level of confidence, where improvement is defined as a significant watershed-wide improvement in one or more water quality parameters associated with the impairments. Under Option 2b, states can use a multiple lines of evidence approach to demonstrate watershed improvement. A "multiple lines of evidence approach" means that the cumulative weight of several lines of evidence is used to assess whether a watershed-wide improvement has occurred.

A watershed approach is focused on hydrologically defined areas, involves key stakeholders, uses an adaptive management process to address priority water resource goals, and uses an integrated set of tools and programs to meet those goals. To meet the requirements of SP-12, use of the watershed approach must be demonstrated by:

1. Identifying the HUC-12 in which the approach was applied;
2. Identifying key stakeholders involved and their roles in implementing the watershed approach to achieve the reported water quality improvements;
3. Describing the watershed plan that was developed and how it was implemented to achieve the reported water quality improvements; and
4. Documenting the restoration actions or BMPs that resulted in water quality improvement.

Developing an Effectiveness Monitoring Plan

Effectiveness monitoring should not be an afterthought, but a critical component of the TMDL development process. It should form the foundation of an adaptive management approach, providing critical information on what is working where and why. Ideally, an effectiveness monitoring plan is developed during, or immediately following, TMDL development. Although the specifics of an effectiveness monitoring plan will vary from one watershed to another, the general steps involved are common to all (Figure 1).

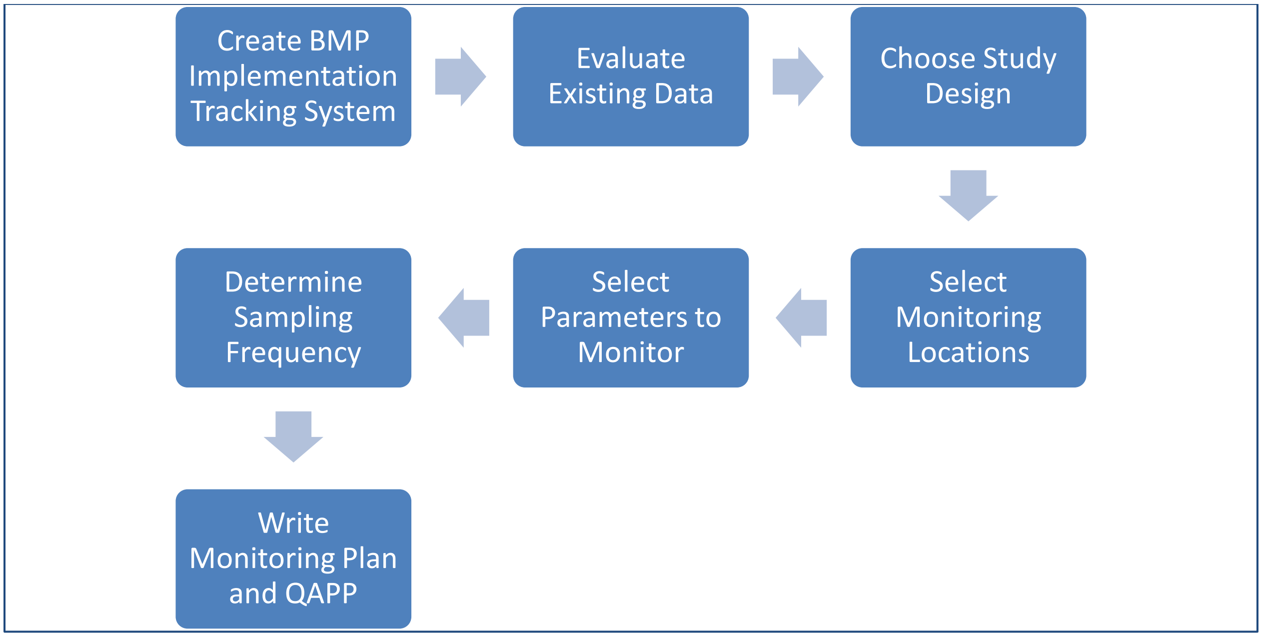

Figure 1 General steps in the development of a TMDL effectiveness monitoring plan.

Create BMP Implementation Tracking System

In order to meet the requirements of SP-12, the TMDL implementation measures that resulted in water quality improvement must be documented. However, the importance of BMP implementation tracking goes beyond receiving credit for measure SP-12. Effectiveness monitoring is a critical component of an adaptive management approach to water quality improvement. While one can certainly determine trends in water quality over time without knowing anything about management actions in the watershed, those trends cannot be attributed to specific TMDL implementation activities without certain critical pieces of information. The minimum pieces of information that should be tracked for all BMPs include:

- Type of BMP;
- Latitude and longitude of BMP location;
- Date of implementation; and
- Maintenance schedule and activities.

These minimum pieces of information are necessary for demonstrating a linkage between water quality improvement and TMDL implementation. Ideally, BMP implementation would be tracked in a geospatial database that records all information about each BMP and its location. Some management measures, such as public education programs or lawn fertilizer bans, may not have specific latitudes and longitudes associated with them. However, the geographic areas in which these types of management measures are implemented can still be represented in a geospatial database.

Evaluations of individual BMP effectiveness can also be carried out to supplement this information. This can provide even stronger support that the management measures implemented are responsible for water quality improvement. However, BMP effectiveness is highly variable and the effectiveness of one BMP cannot be assumed to represent all BMPs of that type or in that location. Since every BMP in a watershed cannot be monitored, individual BMP effectiveness monitoring does not necessarily provide significant information relevant to watershed-wide improvements. General estimates of BMP effectiveness are already available from sources such as the International Stormwater BMP Database (www.bmpdatabase.org) and use of these estimates is a cost-effective method for choosing BMPs to implement and evaluating their likely effectiveness.

Evaluate Existing Data

Existing data are an important component of the development of any monitoring plan. Exploratory analysis of these data can help to inform the study design by revealing apparent trends, important covariates, and baseline conditions. Apparent trends can most simply be evaluated by plotting the parameter of interest in a scatterplot over time. This will often reveal apparent trends that can then be tested with more formal statistical tools. Linear regression and Kendall trend tests are two such tools. Covariates are other variables that are significantly correlated with the water quality parameter(s) of interest. Inclusion of covariates in the monitoring plan and subsequent analyses can increase the chances of detecting a statistically significant trend, where one exists. Data from two or more different sites can also be compared to one another to visualize differences between upstream and downstream locations. This is most simply achieved with boxplots. The existing data should also be evaluated for normality if possible. Normally distributed data are required for parametric statistical techniques, such as linear regression or t-tests (although these two techniques are relatively robust with large sample sizes). If the raw data do not appear to be normally distributed, transformations should be tested to see if they improve the normality. Common transformations for water quality data include the logarithmic transformation and the square root transformation. If the normality of the data cannot be improved, nonparametric statistical techniques (e.g., Kendall trend test, rank sum tests) may have to be used instead.[2]

It is not necessary to conduct a full statistical analysis of the existing data. Quite frequently, in fact, there will not be sufficient data available to conduct any meaningful statistical analysis. The point of the existing data evaluation is simply to provide insight into the nature of the data such that an appropriate study design can be chosen. Most importantly, existing data allow for estimation of the error variance in the water quality parameter(s). This is critical for determining the necessary sample size. The higher the variance, the more samples that will need to be collected to demonstrate trends or changes with statistical significance. The lower the variance, the fewer samples that will be required.

[2] For guidance on appropriate statistical methods, see: Helsel, D.R. and R. M. Hirsch, 2002. Statistical Methods in Water Resources. Techniques of Water Resources Investigations, Book 4, Chapter A3. U.S. Geological Survey. 522 pages.

Existing data are typically available for the development of TMDL effectiveness monitoring plans, as TMDLs themselves are not frequently developed without monitoring data to confirm the impairment or characterize current conditions. In the event that existing data are not available, however, one season of sampling is typically sufficient to estimate the error variance and identify some covariates.

Choose Study Design

Choosing a study design requires a thorough understanding of any existing data, the TMDL implementation plan, and requirements for delisting the water body (e.g., a 30-day geometric mean based on a minimum of 5 samples). It is important to understand how the data to be collected under the effectiveness monitoring program will be used in future analyses, as this will impact the monitoring design. For example, will pre-implementation data be compared to post-implementation data (before, after, control, impact design), or will trends be evaluated continuously over the entire time frame (trend design)? Will one watershed be compared to another, possibly a reference watershed (paired watershed or upstream/downstream design)? The answers to these questions should inform your selection of monitoring locations and sampling frequency.

Select Monitoring Locations

Selection of water quality monitoring locations typically follows either a targeted design or a probabilistic design. Targeted designs are most commonly used for evaluating TMDL effectiveness because the same monitoring site can be maintained over a long period of time and sampled with a consistent and representative frequency. This allows for statistically valid trend analyses to be conducted on the data. EPA's National Water Program Guidance Measure SP-12 requires that results be documented at the HUC-12, or similar, scale. If one of the objectives of the effectiveness monitoring program is to obtain data that will be used in demonstrating improvement under SP-12, monitoring locations should include sites at HUC-12 outlets. This allows for trends to be detected at the appropriate scale and the trends can be related to known BMPs that have been implemented in the contributing drainage area.

Local watershed stakeholders may have additional objectives for the monitoring plan. For example, a particular river reach may be targeted for riparian reforestation and the effectiveness of this individual project may be of interest. In this case, monitoring at the reach scale will also be appropriate. In another example, urban jurisdictions may be interested in the collective effectiveness of their water quality management actions. Key streams that intersect the jurisdiction's boundaries may be targeted for monitoring in this case. A single monitoring program can have multiple objectives and be carried out at multiple scales. It is important to distinguish between these scales however, as data analysis should be conducted on only one scale at a time (i.e., reach and watershed scale data should not typically be combined in the same analysis).

Select Parameters to Monitor

In addition to the primary water quality parameter(s) of concern, numerous other parameters may also be monitored as part of an effectiveness monitoring program. The important covariates identified in the evaluation of existing data should be included. These can be used in subsequent analyses to help demonstrate trends in the parameter(s) of concern. Flow should be measured where possible. Flow data allow for the calculation of loads and flow-weighted averages, which are useful in trend analyses and comparisons of before/after data. Existing flow gages may be operated by one or more agencies; existing gages should be identified and utilized as appropriate. Rainfall and air temperature are also monitored at many locations by multiple agencies and these data can be useful in statistical analyses to detect change over time. Changing land use patterns over time can often confound trend analyses if they are not accounted for. Any information on changes in land cover/land use patterns should be monitored if possible. In coastal areas (often including hundreds of miles inland), the National Oceanic and Atmospheric Administration's Coastal Change Analysis Program (www.csc.noaa.gov/digitalcoast/data/ccapregional/) monitors changes in land cover approximately every 5-years. These data can also be incorporated into subsequent analyses to help explain variability in the data. Biological and habitat data can be monitored to provide additional supporting information for an SP-12 determination. Option 2b uses a multiple lines of evidence approach where biological and habitat data can be incorporated. Macroinvertebrate, fish, and/or periphyton assemblage sampling, for example, allows for assessments of biotic integrity using existing protocols that can provide supporting information to an observed trend in a water quality parameter. Biological data are also a more direct measure of aquatic life designated use support.

Determine Sampling Frequency

It is important that the results of a TMDL effectiveness analysis are statistically significant. This is required by Option 2a of Measure SP-12 (if the results are not statistically significant, but there are additional lines of evidence, then Option 2b can be used). A statistically significant improving trend may also provide justification for delisting a previously impaired water body that is now meeting water quality standards. Ensuring statistical significance can be aided by a power analysis. To conduct a power analysis, one must be able to estimate two variables:

1. The error variance in the measured parameter, which can be estimated by calculating the standard deviation of the residuals of a model that includes known covariates (e.g., season).
2. The expected change in the response variable's concentration, which should be estimated based on the expected effects of BMP implementation. If the expected reduction is uncertain, a conservative estimate should be made.

The power analysis results in a recommended sample size for detecting the expected change in the parameter with the specified error variance. This sample size represents the total number of samples necessary to detect that change, but says nothing about when these samples should be collected. An estimate of the time frame that the expected change will take place allows the total sample size to be distributed over an appropriate number of years (keep in mind that there may be a significant lag time between BMP implementation and an observed water quality response). For example, a power analysis might result in a recommended sample size of 60 for an expected change that is estimated to take 5 years to occur. Year-round, monthly sampling would provide a sufficient number of samples to detect this change with statistical significance. Alternatively, if only the summer months are of concern or suitable for field sampling, samples could be collected biweekly from May to October in order to arrive at the same total sample size after 5 years. Calculated sample sizes will vary depending on the type of statistical test that will be used to analyze the data. This is why it is important to evaluate existing data and think about the statistical method that will be most appropriate for analyzing the data resulting from the effectiveness monitoring program.

Write Monitoring Plan and Quality Assurance Project Plan

The monitoring plan and quality assurance project plan (QAPP) document the decisions resulting from the elements discussed above. The monitoring objectives (evaluating effectiveness of TMDL implementation) and study design should be fully described in this document. Monitoring locations should be listed and mapped, water quality parameters defined, and number of samples needed explained. The expected change in water quality and estimated response time of the various BMPs or TMDL implementation measures should also be described along with the error variance in the water quality parameter of interest. These two measures determine the estimated sample size necessary to detect a statistically significant change.

Demonstrating Effectiveness

The objective of a TMDL effectiveness monitoring plan is to demonstrate incremental improvements in the relevant physical, chemical, and/or biological elements of the impaired water resource. While monitoring is vitally important, implementation of watershed management measures to restore and protect the resource is equally critical, as incremental improvements will not be seen without these actions. Effectiveness monitoring results help to target watershed management actions through an adaptive management approach. Ultimately, effectiveness will be determined by compliance with water quality standards and attainment of designated uses. This is the overriding goal that all implementation and monitoring activities should be designed to achieve. If an activity does not help to achieve this goal, its status as a program priority should be reevaluated.