DEPARTMENT OF
ECOLOGY
State of Washington

Guidance for Effectiveness Monitoring of Total Maximum Daily Loads in Surface Water

May 2013

Publication No. 13-03-024

Publication and Contact Information

This report is available on the Department of Ecology's website at
www.ecy.wa.gov/biblio/1303024.html

The Activity Tracker Code for this study is 12-024.

For more information contact:

Publications Coordinator
Environmental Assessment Program
P.O. Box 47600, Olympia, WA 98504-7600
Phone: (360) 407-6764

Washington State Department of Ecology - www.ecy.wa.gov/

o Headquarters, Olympia	(360) 407-6000
o Northwest Regional Office, Bellevue	(425) 649-7000
o Southwest Regional Office, Olympia	(360) 407-6300
o Central Regional Office, Yakima	(509) 575-2490
o Eastern Regional Office, Spokane	(509) 329-3400

Any use of product or firm names in this publication is for descriptive purposes only and does not imply endorsement by the author or the Department of Ecology.

If you need this document in a format for the visually impaired, call 360-407-6764.
Persons with hearing loss can call 711 for Washington Relay Service.
Persons with a speech disability can call 877-833-6341.

Guidance for Effectiveness Monitoring of Total Maximum Daily Loads in Surface Water

by

Scott Collyard and George Onwumere

Environmental Assessment Program
Washington State Department of Ecology
Olympia, Washington 98504-7710

This page is purposely left blank

Table of Contents

List of Figures

Page

List of Tables

Page

Abstract

The federal Clean Water Act gives states the primary responsibility for implementing programs to protect and restore water quality, including monitoring and assessing the nation's waters and reporting on their quality. In Washington State, the Department of Ecology is the agency primarily responsible for implementing the requirements and provisions of the Clean Water Act, including monitoring the effectiveness of water pollution cleanup plans.

Effectiveness monitoring uses a combination of monitoring types to evaluate whether specified activities have achieved the desired effect. It is an essential component to the adaptive management process when best management practices are implemented to control anthropogenic pollution. It is also one of the several required components when (1) we develop Total Maximum Daily Loads (TMDLs) or other watershed-based pollution control plans, or (2) state and federal funds are used to implement nonpoint-source pollution control strategies.

This is a living document that presents a strategy for monitoring the effectiveness of established TMDLs and other nonpoint-source and point-source pollution control plans.

Acknowledgements

The authors of this report thank the following people for their contribution to this report:

- Carol Smith, Washington State Conservation Commission, and Leska Fore, Puget Sound Partnership, for technical assistance and review of the final report.
- Jill Gable and Jayne Carlin, U.S. Environmental Protection Agency, for funding and guidance with developing Ecology’s effectiveness monitoring Program.
- Laura Blake, Cadmus Group Inc., for support through development of additional reports and guidance documents.
- Washington State Department of Ecology staff:
  - Jessica Archer, Jim Ross, and Mindy Roberts for reviewing the final report.
  - Markus Von Prause for technical assistance.
  - Helen Bresler for assistance with Water Quality policy.
  - Jean Maust and Joan LeTourneau for formatting and editing the final report.

Introduction

The Washington State Department of Ecology (Ecology) is required, under Section 303(d) of the federal Clean Water Act and implementing regulations, to periodically prepare a list of waterbodies that are out of compliance with the state water quality standards. After the preparation of this list and the subsequent approval by the U.S. Environmental Protection Agency (EPA), Ecology is responsible for preparing and implementing Total Maximum Daily Loads (TMDLs) on these watersheds as well as evaluating the effectiveness of the cleanup plans to achieve the needed improvement in water quality.

TMDLs are tools for implementing water quality standards under the federal Clean Water Act and are based on the relationship between pollution sources and instream water quality conditions. A TMDL is the summation of the individual wasteload allocations for permitted discharges (i.e., point sources), load allocations for nonpoint sources of pollution (including natural background conditions), a margin of safety, and future growth. The TMDL process establishes the allowable loadings or other quantifiable parameters for a waterbody. It also provides the framework that establishes water quality-based controls designed to bring the waterbody into compliance with applicable water quality standards.

Ecology is currently working under a memorandum of agreement (MOU) with EPA to address all polluted waters on the State Section 303(d) list. The MOU requires Ecology to develop implementation plans for all approved TMDLs. These TMDL implementation plans must include “feedback loops” to evaluate TMDL effectiveness. Specifically, Ecology must determine:

- Whether the required source controls have been put into place.
- Whether those source controls are effective as measured against relevant TMDL and implementation plan targets.
- Whether the TMDL or implementation plan needs revision.

This guidance document attempts to provide a strategy for effectiveness monitoring that is consistent with supporting watershed-based adaptive management efforts. Specifically, this document lays the initial framework for the development of a program for monitoring the effectiveness of TMDL implementation plans while providing information to make higher level management decisions. Although this guidance document does not provide instructions on how to conduct specific effectiveness monitoring studies, a framework and links to resources and references are provided. This document is intended to be updated periodically as our knowledge about this process improves.

TMDL Process

The TMDL process begins with the development of a scientific study to identify the pollution sources and the load allocations needed to bring the waterbody into compliance with state water quality standards (Figure 1). Once the study is completed, Ecology scientists generate a technical report analyzing the pollution parameters identified in the Section 303(d) list of impaired waterbodies.

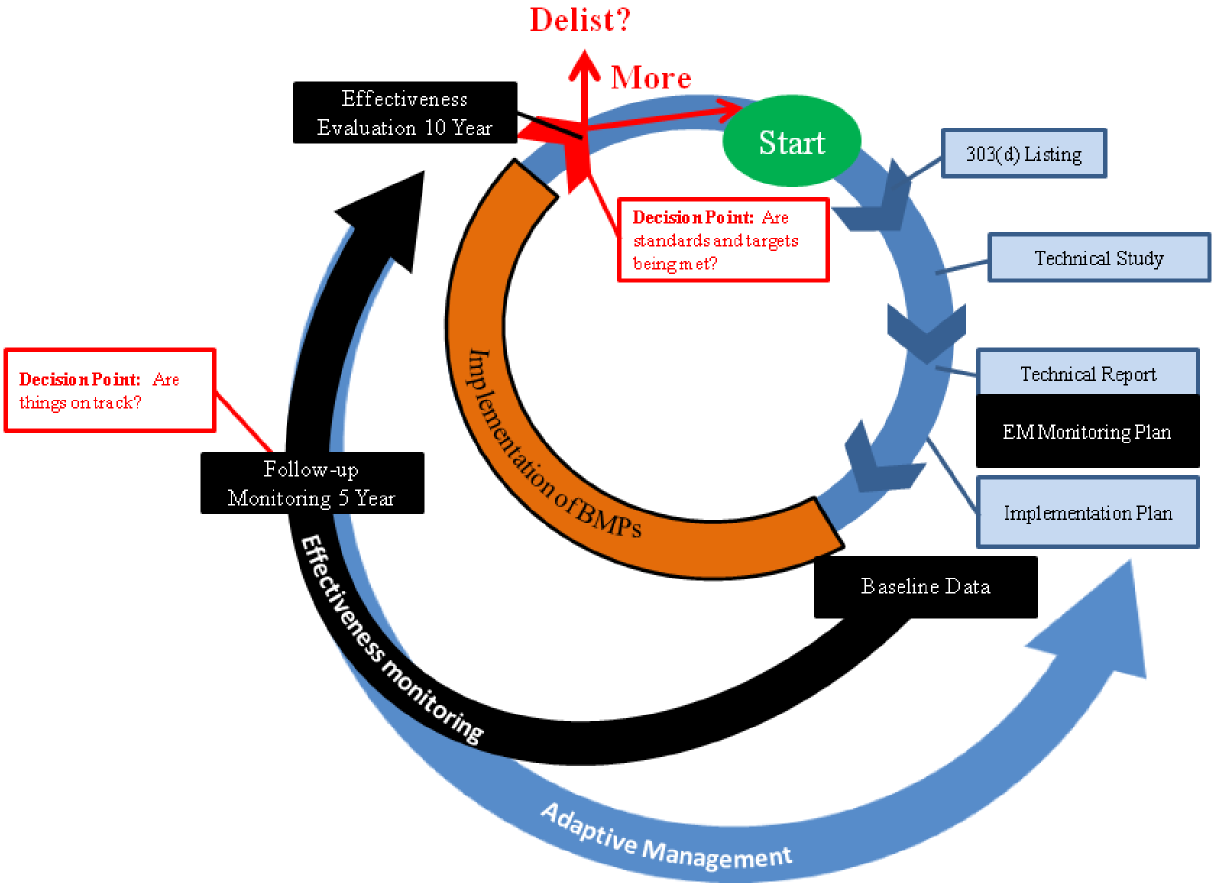


Figure 1. Conceptual framework for implementing pollution control strategies in Washington State.

Following the technical report an implementation plan is developed, which outlines technology-based best management practices (BMPs) needed to meet load allocations developed for the watershed (Baldwin et al., 2007). Implementation activities continue until periodic follow-up monitoring indicates compliance with state water quality standards. A final evaluation of the strategies effectiveness occurs once all activities outlined in the implementation plan have been addressed.

Adaptive Management

Adaptive management is defined as a formal process for continually improving management policies and practices by learning from their outcomes (Murray and Marmorek, 2003). For TMDLs, adaptive management requires an explicitly scientific approach to managing pollution controls which also improves our understanding of how land use practices and nonpoint pollution can be controlled.

For the TMDL process, this definition can be expanded as follows:

- **Testing assumptions** is about implementing specific BMPs identified in the implementation plan to achieve a desired outcome (i.e., improvement in water quality). Once implemented, the outcomes are monitored to see how they compare to the ones predicted by assumptions outlined in the TMDL (e.g., expect pollution load reductions). Monitoring plans must be designed not only to determine which actions worked but also which did not and why.
- **Adaptation** is about taking action to improve implementation performance based on the outcome of the monitoring. If the implemented actions did not achieve expected results, it is because assumptions were either incorrect or poorly executed, conditions in the TMDL area have changed, some sources of pollutants were not identified, the monitoring design was inadequate, or a combination of the above. Adaptation involves changing or updating assumptions and intervening to respond to the new information obtained through monitoring efforts.
- **Learning** is about systematically documenting the process and the results that have been achieved. This documentation will help to avoid repeating undesirable outcomes in other TMDLs and will also enable the broader conservation community to benefit from lessons learned. Furthermore, it will also enable the scientific community to better understand the relationships between land use practices and nonpoint pollution to waterways.

All Ecology TMDL implementation plans as well as most state and federal watershed restoration plans in Washington State make reference to adaptive management. However, many of these plans fail to achieve the desired outcomes because:

- Plans fail to identify a formal process for implementing an adaptive management strategy.
- Monitoring efforts are inadequate, disconnected, or not considered in the adaptive management process.
- Plans and program policies are often changed (or remain in place) without sufficient evaluation of why goals and objectives were not being meet.

For Ecology’s TMDL program to achieve success through adaptive management, we must integrate project level monitoring efforts to help answer large scale programmatic questions (i.e., is the TMDL program effective?). Additionally, we must develop and integrate a comprehensive monitoring program to support adaptive management at all levels. The goal of Ecology’s effectiveness monitoring program is to provide this feedback by integrating a monitoring schedule through the life of the TMDL.

What is Effectiveness Monitoring?

TMDL effectiveness monitoring evaluates whether management activities have achieved the desired effect. Rather than monitoring the effectiveness of a particular project, Ecology's TMDL effectiveness monitoring program measures the cumulative effect of all activities in the watershed.

Effectiveness monitoring is a fundamental component of any TMDL implementation activity. It is an important tool in the adaptive management process because it informs and allows restoration strategies to be adjusted if project goals are not being achieved. If implemented thoughtfully, it will increase the likelihood that activities to control pollution will succeed.

The benefits of effectiveness evaluation include:

- More efficient allocation of funding.
- Optimization in planning/decision-making (program benefits).
- Watershed recovery status (how much restoration has been achieved, how much more effort is required).
- Adaptive management or technical feedback to refine restoration treatment design and implementation.

The effectiveness evaluation addresses four fundamental questions about restoration or implementation activity:

- Is the restoration or implementation work achieving the desired goal of significant improvement?
- How can restoration or implementation techniques be improved?
- Is the improvement sustainable?
- How can the work become more cost-effective?

Levels of Effectiveness

An effectiveness monitoring program that is efficient should also be able to answer management questions at multiple levels in order to meet the goals of the Clean Water Act (Figure 2).

The monitoring efforts required to inform each level are not mutually exclusive. Since clean water comes directly from events that occur on the ground, the actions of all preceding levels must result in decisions that can be practically implemented. This requires integrating of monitoring efforts and defining a clear process for implementing adaptive management.

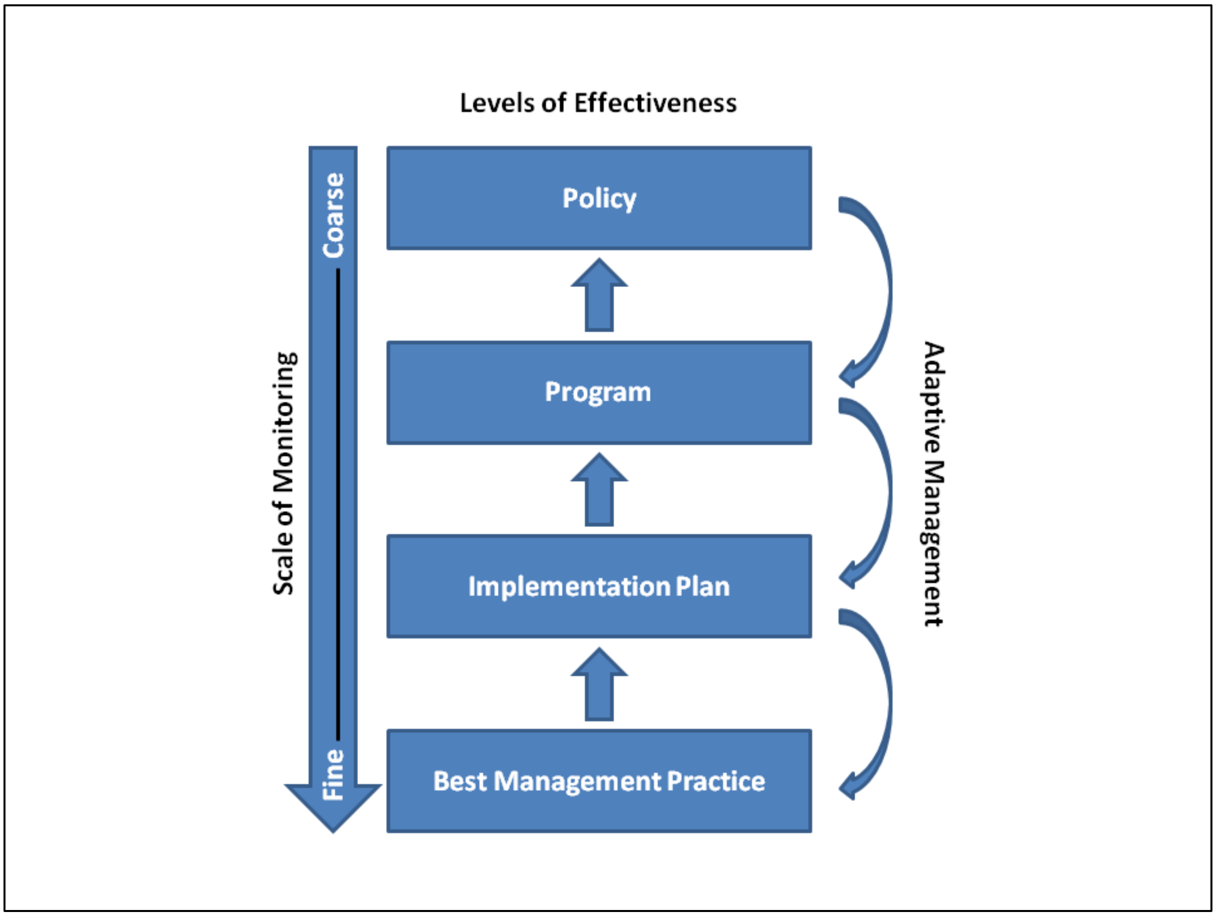


Figure 2. Management levels in the TMDL process that an effectiveness monitoring program should address.

Leveraging Existing Monitoring Programs

Effectiveness monitoring will use multiple types of monitoring to achieve objectives. Ecology currently uses many monitoring types to determine if an implementation plan is effective (Table 1). A more detailed definition of monitoring types is provided in Appendix B.

Ecology manages several monitoring programs that can provide data which can be used to answer many of these questions at a programmatic scale as well as provide supplement project-level data for effectiveness monitoring. Data collected from these monitoring programs are also used to develop water quality indicators and track changes in water quality over time (Hallock, 2011). These programs include:

- River and Stream Water Quality Monitoring Program
- River and Stream Flow Monitoring
- Intensively Monitored Watersheds
- Stream Biological Monitoring
- Status and Trends for Watershed Health and Salmon Recovery
- Aquatic Plant Monitoring

- BEACH Program
- Marine Water Quality Monitoring
- Marine Sediment Monitoring
- Washington State Toxics Monitoring
- Washington's State Effectiveness Monitoring

These programs make up the statewide monitoring and assessment program used to help EPA meet the prerequisites of the federal Clean Water Act. The document entitled *Clean Water Act Monitoring Strategy for Washington State* describes these programs and explains how they fit into Washington State's overall monitoring strategy (Ecology, 2013).

In addition to Ecology's statewide monitoring programs, Washington State has other monitoring efforts at the local level. Leveraging these local monitoring efforts or programs is integral to the success and efficiency of any effectiveness monitoring study.

Table 1. Types of monitoring used to evaluate the effectiveness of TMDLs.

Question	Type of monitoring
What are the current water quality conditions?	Baseline
What is the overall status of water in the watershed?	Status
Are conditions changing over time?	Trend
Are water quality standards and TMDL/NPDES targets being met?	Compliance
Where BMPs are installed, are they being maintained?	Implementation
Are additional source controls needed?	Source Identification
Are the original assumptions of the water quality model correct?	Validation
Are changes in water quality linked to implementation of pollution control measures?	Effectiveness

NPDES: National Pollutant Discharge Elimination System

Effectiveness Monitoring Strategy

There is a need for a single, overarching, and consistent approach for monitoring the effects of pollution control actions on surface waters. The strategy, by design, should be flexible enough to develop monitoring plans specifically suited to the objectives and strategies of a multitude of efforts. It should also include a process which evaluates progress and an approach which collects and analyzes data that can provide practical information to stakeholders and watershed planners.

The strategy outlined here attempts to provide the building blocks of a monitoring framework that complements the development and successful implementation of TMDLs and other pollution control plans.

The strategy was developed in accordance with requirements of the Clean Water Act which are outlined in the EPA publication, *Elements of a State Water Monitoring Assessment Program* (EPA, 2003). Also, the strategy fulfills requirements of both state and federal grants used to develop and implement TMDLs and other pollution control plans (Ecology, 2012a). In short, these plans must have a monitoring strategy to evaluate the effectiveness of the implementation effort over time to assure progress is being made toward attaining water quality standards.

The monitoring strategy should be (1) fully integrated with the established schedule and interim milestone criteria, (2) designed to assess progress in achieving loading reductions and meeting water quality standards, and (3) conducted at a watershed-wide scale to measure the effects of multiple programs, projects, and trends over time.

The effectiveness monitoring strategy specifically addresses the Clean Water Act requirements through development and implementation of the following components:

- **Effectiveness Monitoring Project Plan:** Identifies specific monitoring goals and objectives and describes the process for generating and analyzing data to meet the goals and objectives.
- **Quality Assurance Plan:** Outlines additional details related to field sampling and laboratory protocols, data management, and the proposed process for addressing data quality issues that arise during the course of the project.
- **Reporting:** Summarizes information, provides for adaptive management, and informs managers and the public.

Key Components

Goals and Objectives

Although goals and objectives may vary with each effectiveness monitoring study, some of them should be consistent among all effectiveness monitoring studies. The broad-based goals of effectiveness monitoring for watershed-based pollution control plans are to determine if (1) water quality standards and targets are being met (2) progress is being made towards meeting standards and targets, (3) water quality improvements are linked to water cleanup activities, and (4) the current implementation strategy is sufficient.

The core objectives for assessing the effectiveness of Ecology's TMDLs are to:

- Review and analyze historic data.
- Define and map land uses and list potential pollutants.
- Collect data consistent with sampling requirements for delisting, as specified in Washington's Water Quality Program Policy 1-11 (Ecology, 2013).
- Document implementation efforts.
- Analyze and interpret data to determine if changes in water quality are significant.
- Measure link between implementation efforts and changes in water quality.

Timeline for Studies

Implementation of BMPs at a watershed scale is an iterative process. Also, surface water response time to BMPs varies widely (Figure 3). Some parameters may respond more quickly to BMPs than others. For example, construction of BMPs such as a livestock exclusion fence to reduce fecal coliform levels is effective immediately. In contrast, the results of tree planting to reduce stream temperature are not likely to be completely effective for decades. An effectiveness monitoring strategy that tracks the project timeline assures an adequate feedback mechanism is in place to help guide management decisions.

Ecology has implemented a 10-year timeline for reporting the progress of TMDLs and other pollution control measures. Monitoring is conducted as part of the original technical study, then again after 10 years or after sufficient implementation occurred in the watershed. Although this approach was sufficient for some smaller point source TMDLs, it has been inadequate for TMDLs with multiple sources and impairments. Therefore, the monitoring strategy presented here calls for monitoring every 5 years.

Although the monitoring efforts will vary between these monitoring periods, this approach is necessary to assure that progress is being made throughout the process. For example, follow-up monitoring may be scaled back to only a limited number of key sampling stations or tracking of implementation. Failure to implement this multi-year monitoring approach decreases the likelihood of determining if the BMPs for the cleanup project were indeed effective.

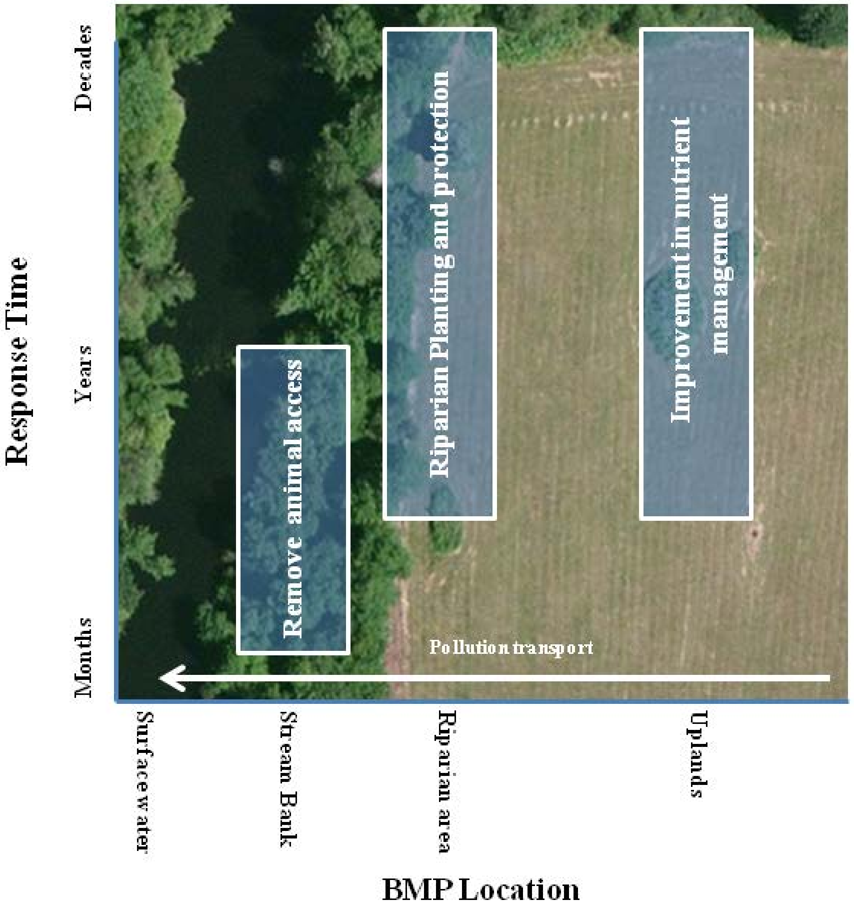


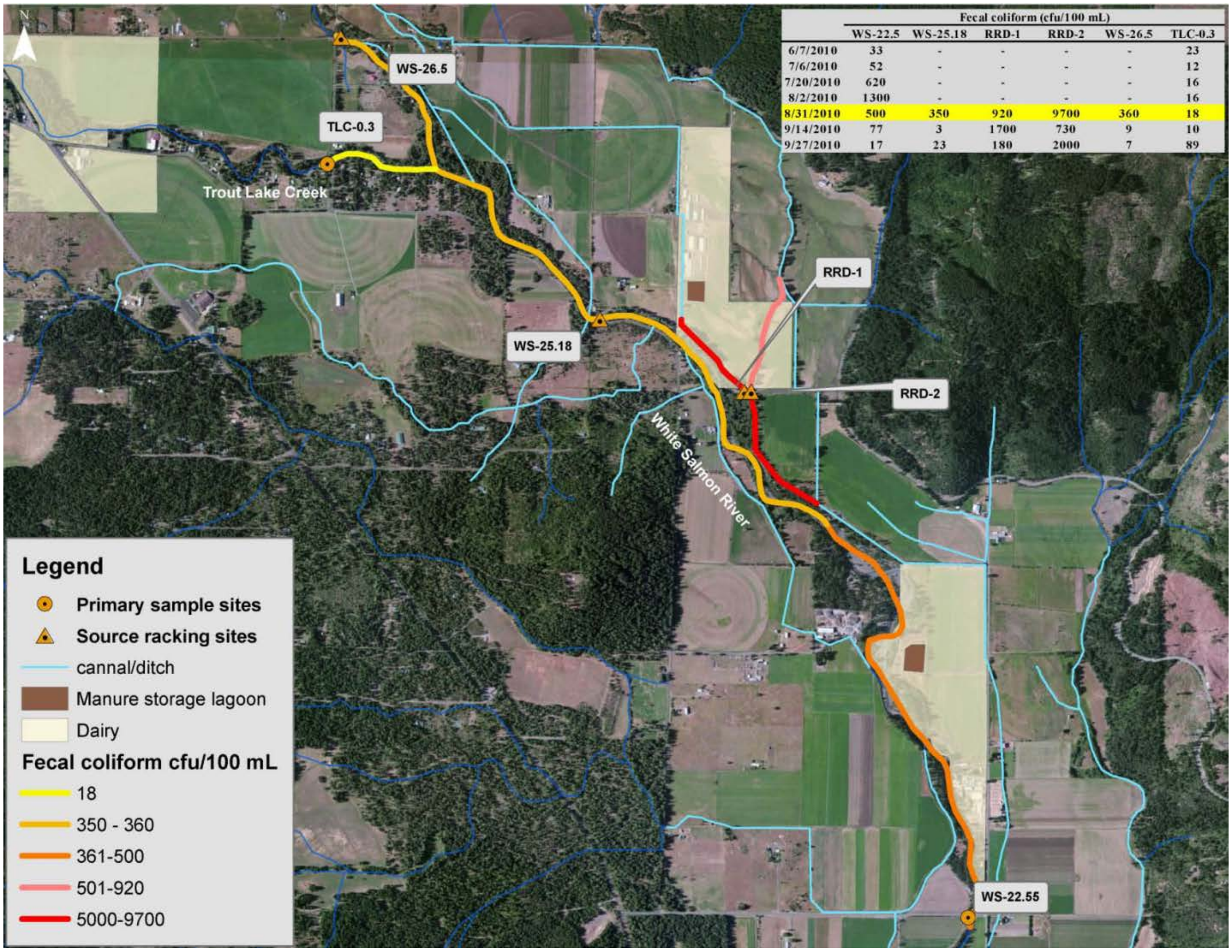
Figure 3. The scale of common BMPs and their typical response times.
Modified from Mesner, 2011.

Source Identification

During the project, additional sites and/or samples should be added or sampled, allowing for immediate identification and resolution of pollution sources. Including a source identification component in the study can be an important component in the adaptive management process.

The source tracking of a pollutant is usually triggered after a number of violations of water quality standards are observed during routine sampling of a parameter. Sources or hot spots of pollution may be identified using a combination of desktop reconnaissance and bracketed sampling of parameters upstream of primary sampling locations. Source tracking stations can be chosen by bracketing between land uses, land use activities, or individual property parcels. Before source tracking occurs, a process to report the progress and results should be in place so results can be immediately conveyed to watershed planners.

Figure 4 provides an example of using bracketed sampling to identify sources of fecal coliform pollution in a rural watershed (Collyard, 2011).



	Fecal coliform (cfu/100 mL)					
	WS-22.5	WS-25.18	RRD-1	RRD-2	WS-26.5	TLC-0.3
6/7/2010	33	-	-	-	-	23
7/6/2010	52	-	-	-	-	12
7/20/2010	620	-	-	-	-	16
8/2/2010	1300	-	-	-	-	16
8/31/2010	500	350	920	9700	360	18
9/14/2010	77	3	1700	730	9	10
9/27/2010	17	23	180	2000	7	89

Figure 4. Example of using bracketed fecal coliform sampling and GIS to identify pollution sources.

Fecal coliform sampling locations were bracketed between land parcels after two consecutive high samples were collected at a primary sampling location (WS-22.5) (Collyard, 2011).

Biological and habitat data can also be used to identify pollution sources. The causal analysis/ diagnosis decision information system (CADDIS) can be used to make causal assessments of watershed data to help with stressor identification (EPA, 2010). This process can be especially useful for identifying pollution types in watershed-scale studies with mixed land uses (Haake et al., 2010; Wiseman et al., 2009).

Implementation Monitoring

Implementation monitoring assesses whether activities were carried out as planned (EPA, 1991). The most common use of implementation monitoring is to determine whether BMPs were implemented as specified in TMDLs. Typically this is carried out as a review or site inspection and does not involve any water quality measurements.

Implementation monitoring is the most cost-effective means to reduce nonpoint-source pollution because it can provide immediate feedback to managers on whether the BMP process is being carried out as intended. Implementation monitoring itself cannot directly link management activities to water quality changes. However, it is a critical part of an evaluation of effectiveness and is necessary to meet many of the objectives outlined in this document. Figure 5 provides an example of using Geographic Information Systems (GIS) to track implementation of BMPs installed to reduce fecal coliform levels in a stream (Collyard, 2011).

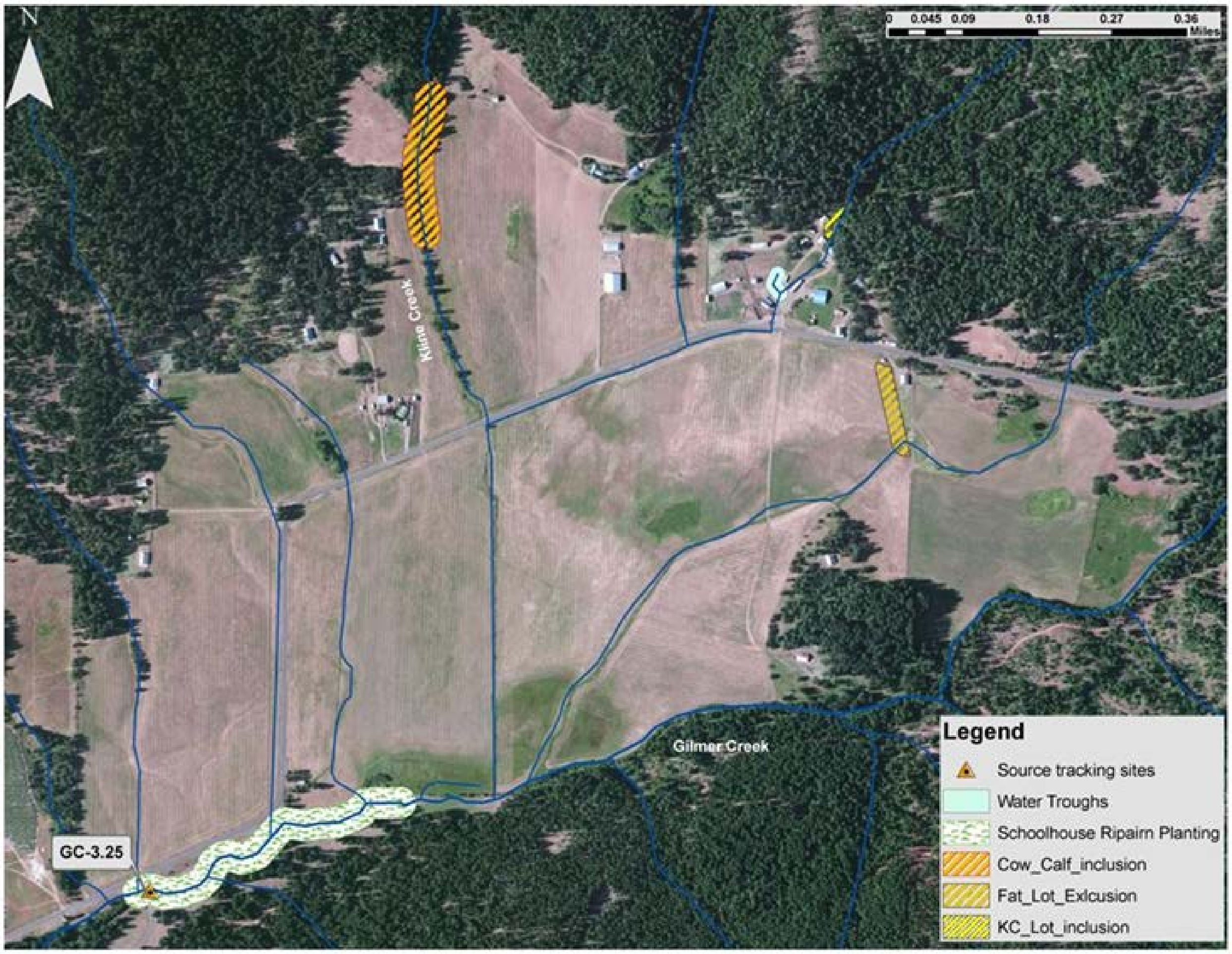
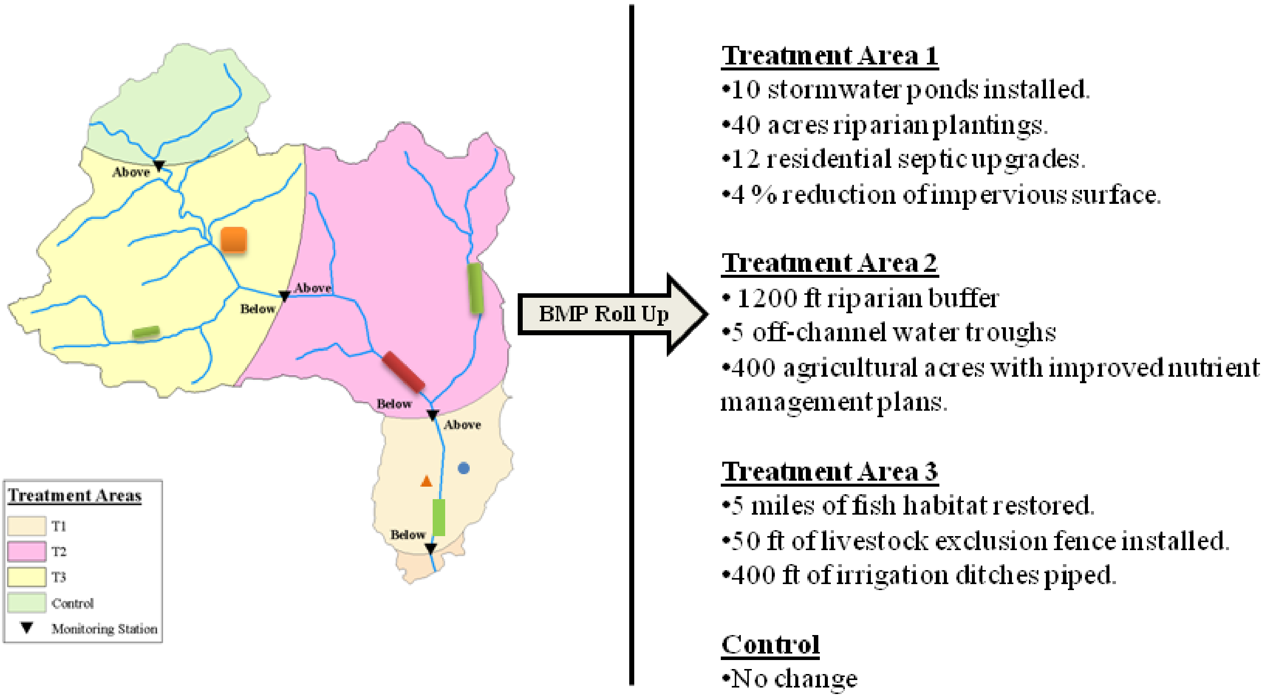


Figure 5. Use of GIS to document the implementation of BMPs to reduce fecal coliform levels in Gilmer Creek (Collyard, 2011).

Linking Implementation Activities to Results

Although implementation monitoring tracks individual BMPs at a site scale, the goal for TMDL effectiveness monitoring is to measure the cumulative effect of all BMPs on water quality (see *Study Design* section). This approach will require a “roll up” of all BMPs within treatment areas and will require a set of implementation metrics to be developed (Figure 6). These metrics, in conjunction with water quality measurements, can be used to link management activities to water quality improvements (Figure 7).



Treatment Area 1
- •10 stormwater ponds installed.
- •40 acres riparian plantings.
- •12 residential septic upgrades.
- •4 % reduction of impervious surface.

Treatment Area 2
- • 1200 ft riparian buffer
- •5 off-channel water troughs
- •400 agricultural acres with improved nutrient management plans.

Treatment Area 3
- •5 miles of fish habitat restored.
- •50 ft of livestock exclusion fence installed.
- •400 ft of irrigation ditches piped.

Control
- •No change

Figure 6. Example of rolling up individual BMPs within treatment areas for purposes of reporting and data analysis.

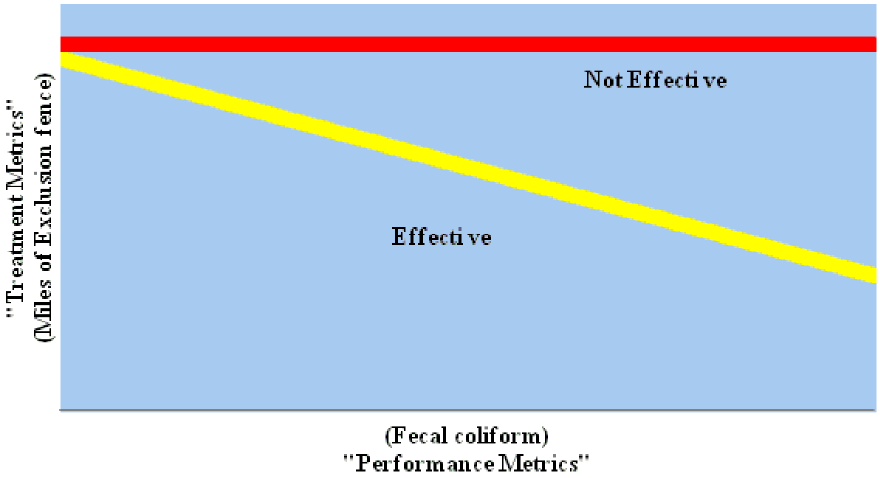


Figure 7. Implementation efforts are the treatment metrics while water quality indicators are the performance metrics.

Using Results to Adapt Management Strategies

Information generated from effectiveness monitoring evaluations can be used to adapt and improve management, planning, accountability, and overall implementation plans. Results from effectiveness monitoring studies can be applied in adaptive management strategies in the following way:

- Highlight the progress of implementation actions.
- Assist in identifying and setting new priorities for future actions.
- Promote accountability.
- Increase stakeholder awareness, participation, and support.

Table 2 provides examples of adaptive management that could be expected in TMDL implementation process.

Table 2. Hypothetical use of effectiveness monitoring results to inform adaptive management.

Monitoring Determines	Adaptive Management Response
Implementation milestones are on schedule and water quality trends continue to improve.	Convey results to stakeholder.
Implementation milestones have not been met.	In consultation with primary stakeholders, consider how management actions can be modified to improve progress.
Water quality continues to exceed standards and no other sources have been identified.	Review BMP implementation at a project scale and make adjustments as needed.
Additional sources of pollution exist that have not previously been identified.	Implement source identification monitoring and report results to appropriate state or local agency. When applicable, modify the implementation plan or TMDL to include additional sources.
Data analysis suggests water quality improvements were driven by non-anthropogenic factors.	Continue following implementation plan and monitoring schedule until assurances can be made that water quality will remain consistent.

Project Monitoring Plan

Effectiveness monitoring studies need a plan which details the process for generating, analyzing, and reporting data to meet identified monitoring goals and objectives. The plan must be developed before implementation of pollution controls and should be incorporated together with the implementation plan.

The effectiveness monitoring plan should contain the following components:

- Characterization of study area
- Site location
- Parameter selection
- Study design
- Data analysis procedures

Once the monitoring plan is complete, a Quality Assurance (QA) Project Plan is developed. The QA Project Plan outlines additional details related to field sampling and laboratory protocols, data management, and the proposed process for addressing data quality issues that arise during the course of the project.

Additional guidance for designing an effectiveness monitoring plan is presented in Appendix D.

Characterization of Study Area

Characterizing the study area includes reviewing current and historical water quality data and implementation efforts as well as evaluating current land uses. Considerable data may already exist for a particular waterbody and should be evaluated with the monitoring objectives in mind.

In many cases, existing data can be used as “pre-implementation” or baseline data. Also, locating monitoring stations at historic or discontinued monitoring stations can improve detection of changes over time and also decrease the time it takes to detect meaningful trends. This information will improve understanding of current and historical water quality conditions, including short- and long-term variability as well as type, location, and timing of implementation actions within the watershed.

GIS has become a powerful tool that can provide many types of information to support the development of a monitoring plan. Specifically, GIS can be used to effectively evaluate land use and pollution sources within the study area (Figure 8). Information collected can be used to evaluate qualitative relationships to changes in water quality. It may also used to determine land uses or land uses practices causing pollutant load and can assistance in the development of other implementation strategies and BMP development.

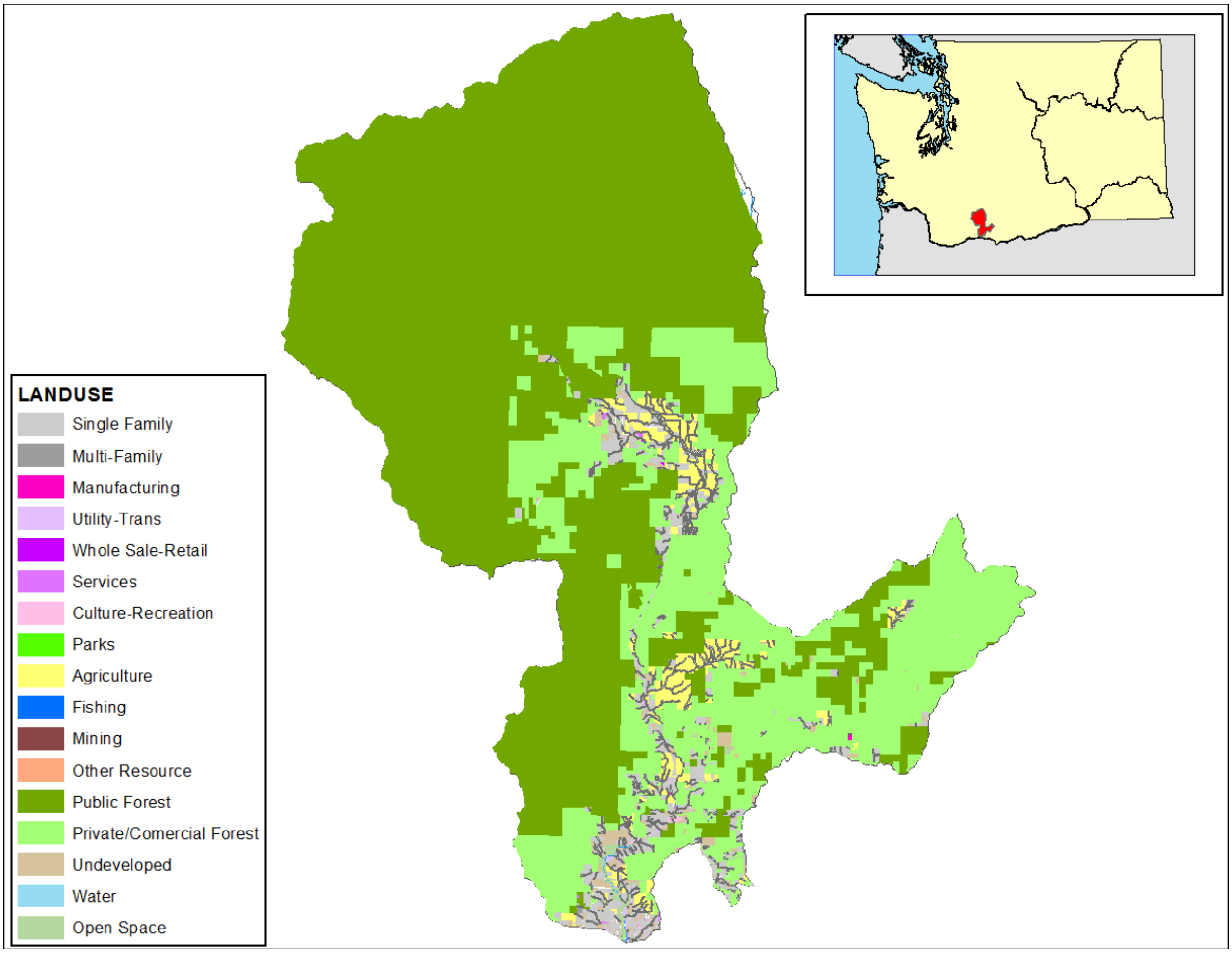


Figure 8. Use of GIS to evaluate land use in relation to high fecal coliform levels on the White Salmon River (Collyard, 2011).

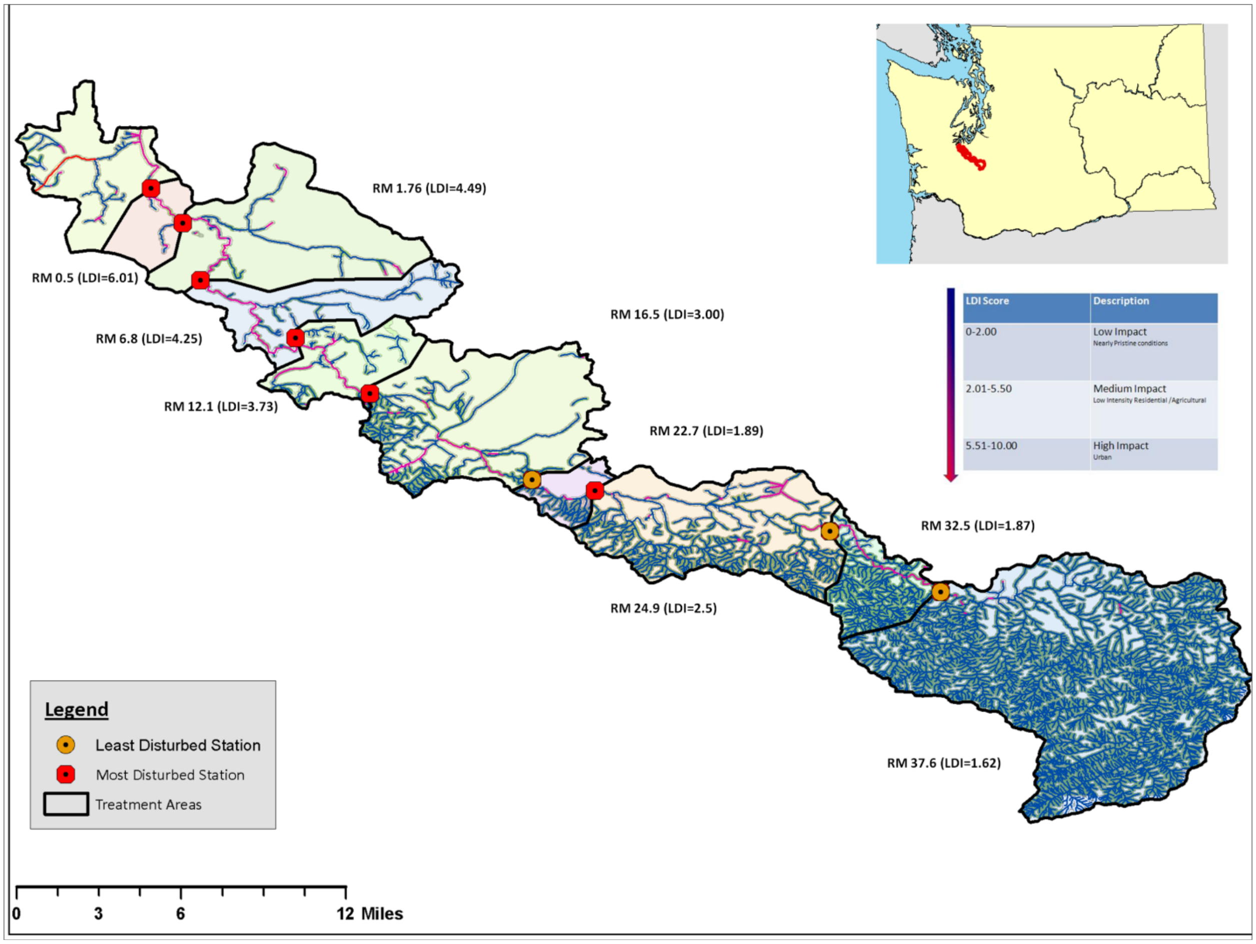
Site Selection

In general, a monitoring site should be located where implementation is expected to have discernible water quality effects. This includes sites on impaired or degraded waterbodies that are downstream of pollution sources or discharges or pollution control measures. Additional sites may also be added by choosing sample locations based on breaks in land use throughout the watershed. An even finer scale evaluation can be conducted at the parcel scale for purposes of identifying obvious sources and for measuring changes in land use over time. This approach is useful to:

- Develop “treatment areas” for measuring the effects of BMPs on water quality within those treatment areas.
- Establish baseline data for land use and land use practices which can be used to measure changes across the watershed.

Ecology is currently developing a GIS tool that uses a human disturbance index developed by Brown and Visas (2005) to predict potential impacts from land use activities (Collyard and Von Prause 2009; Collyard et al., 2011).

Using existing land use data and a development-intensity measure based on energy use per area, a Landscape Development Intensity Index (LDI) can be calculated at the river, stream, or watershed scale. Resulting data can be used to establish treatment areas in between sampling stations (Figure 9). Because treatment areas are based on a disturbance index, information can also be used to prioritize implementation activities in high impact areas.



LDI Score	Description
0-2.00	Low Impact Nearly Pristine conditions
2.01-5.50	Medium Impact Low Intensity Residential /Agricultural
5.51-10.00	High Impact Urban

Figure 9. Evaluation of land use using a Landscape Development Intensity Index (LDI) to establish sampling locations on the Deschutes River watershed (Collyard et al., 2012).

Monitoring Parameters

Primary monitoring should provide data to determine if water quality standards and targets are being achieved and if BMP implementation has improved water quality. Surrogate parameters that relate to water quality impairments such as biological and habitat parameters should also be considered.

Biological and physical habitat sampling for the purposes of TMDL effectiveness monitoring is a new and developing facet to TMDL effectiveness monitoring, but it is expected to become an important component in the future. In addition, the diagnostic capabilities of this type of assessment can identify problems related to water quality and habitat and are a less expensive alternative to intense water quality sampling.

Biological and physical habitat parameters are typically responsive to water quality impairments and can be effective for evaluating water quality improvements. Generally, these parameters integrate the effects of different pollutant stressors and provide an overall measure of the aggregate impact of stressors (EPA, 1997).

Although surrogate data cannot be used to directly demonstrate compliance with numeric water quality standards, they can be used as part of a weight of evidence approach to show that narrative standards are being met (aquatic life use). An example of biological sampling for the purpose of effectiveness monitoring can be found in Collyard and Von Prause (2009).

Covariates

It is important to consider parameters external to the implementation process that influence water quality over time. These parameters or *covariates* can be environmental or anthropogenic. They influence the fate and transport of nonpoint-source pollution to surface waters. This can result in an erroneous interpretation of the data. A list of some potential covariates is presented in Table 3. Because numerous covariates affect any given water quality parameter, it may be best to prioritize them based on their strength of association with the target water quality parameter.

Table 3. Examples of environmental and anthropogenic covariates that can influence water quality data.

Environmental	Anthropogenic
Flow	Land use changes
Temperature	Irrigation practices
Precipitation	Grazing
Weather patterns	Urbanization

Study Design

The overall design of an effectiveness monitoring study is largely determined by the study objectives and the parameters to be monitored. Figure 10 presents a study design that can be used to support the broad-based objectives for Ecology's effectiveness monitoring studies. Water quality monitoring locations are bracketed between treatments areas (typically land use or land uses activities) while BMP implementation occurs within treatment areas. This type of monitoring approach incorporates the principles of two typical statistical designs (EPA, 1997):

Before/after study designs are useful for assessing the effectiveness of BMPs to improve water quality in a waterbody or small basin. Monitoring is conducted in the waterbody for several years before implementation of BMPs and for several years following the implementation of BMPs, depending on the budget. Baseline monitoring in the waterbody is required to establish the baseline data which will then be used for statistical comparisons with post-BMP monitoring data collected in the waterbody. The success of before/after studies depends on the ability to account for temporal variability (e.g., annual or seasonal variation in weather, flow variability, and land use).

Nested paired study designs are very useful for assessing the effectiveness of BMPs to improve water quality in a waterbody in a short period of time. A nested paired watershed design is sometimes referred to as an "above and below" design where one monitoring station is located above a treatment area and one station is located below the treatment area. The main advantage of this approach is that the variation due to non-anthropogenic influences can be statistically controlled provided enough baseline data has been collected.

One important attribute between both designs is the ability to account for the influence of variables (covariates) not related to implementation of BMPs. Establishing "control" sites where parameters are monitored below areas that are untreated or have few anthropogenic activities provides a basis for separating the treatment effect from other extraneous factors (Figure 11). In the case of upstream/downstream comparisons, the upstream sites usually act as the control, while the downstream site serves as the treatment.

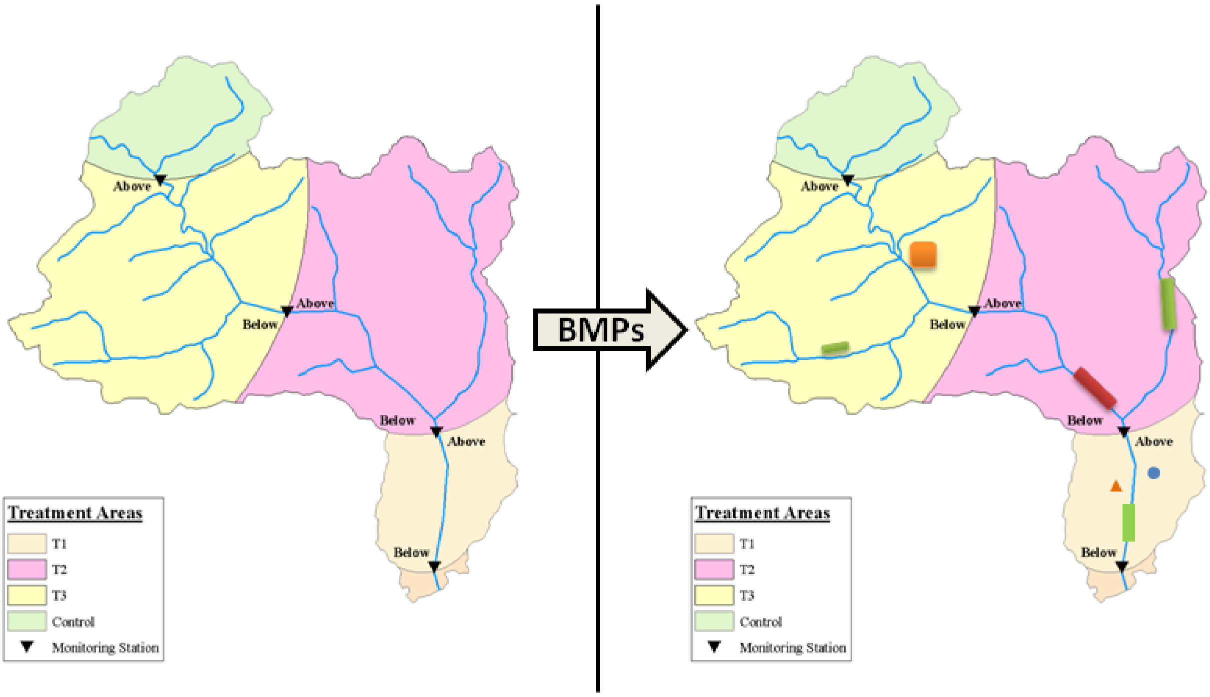


Figure 10. Example of study design used for TMDL effectiveness monitoring in Washington State.

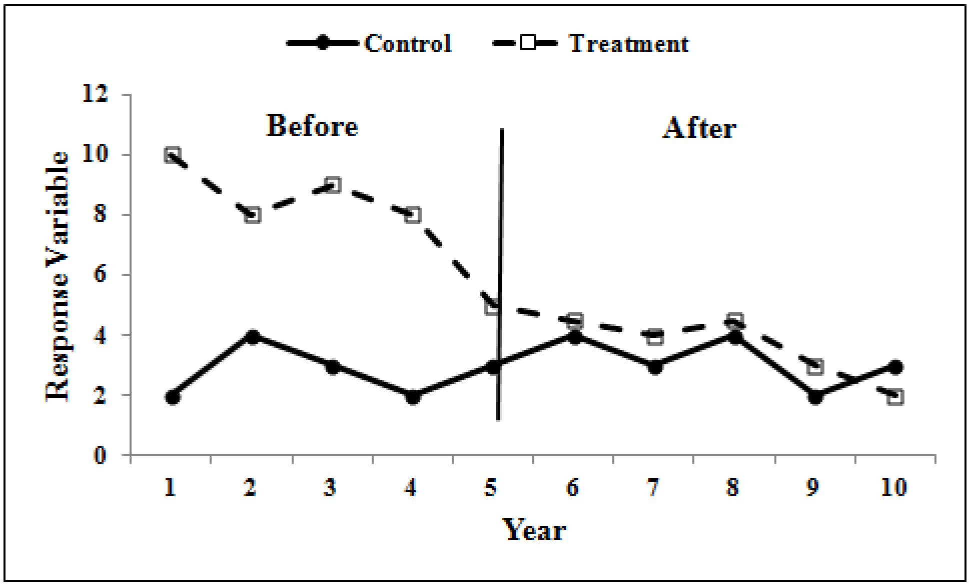


Figure 11. Control sites are monitored concurrently with treatment sites to identify extraneous factors.

Data Analysis

The TMDL effectiveness monitoring plan needs to describe the methodology for assessing information on TMDLs to evaluate attainment of water quality standards. The methodology includes criteria for compiling, analyzing, and integrating ambient conditions with project implementation information. The methodology describes integration of data collected with other data sources, possibly collected for other purposes.

It is important to identify the methods to be used to analyze the data before the data are collected. Efficient development of the data flow process, from project goals and objectives through sample design, scheme, and data management requires a thorough understanding of the data analysis.

A statistical review should be incorporated as part of the monitoring plan to ensure that the statistical testing is appropriate for meeting the monitoring objectives. In addition, a quantified estimate of the accuracy and precision that can be expected from the monitoring effort should be established using a pilot study, historic data, or some other method. This information is necessary to determine if the efforts will answer management questions critical to the higher level decision-making.

This means that before data are collected the following questions should be addressed (EPA, 1991):

- How many samples are likely to be needed to characterize a parameter?
- How many samples are likely to be needed to measure change over time or differences between samples?
- How will the precision and accuracy of the data be assured?

Typical data analysis procedures usually begin with screening and graphical methods, followed by evaluating statistical assumptions, computing summary statistics, and comparing groups of data. Figure 12 provides a diagram for some statistical tests used to support the effectiveness monitoring study described above.

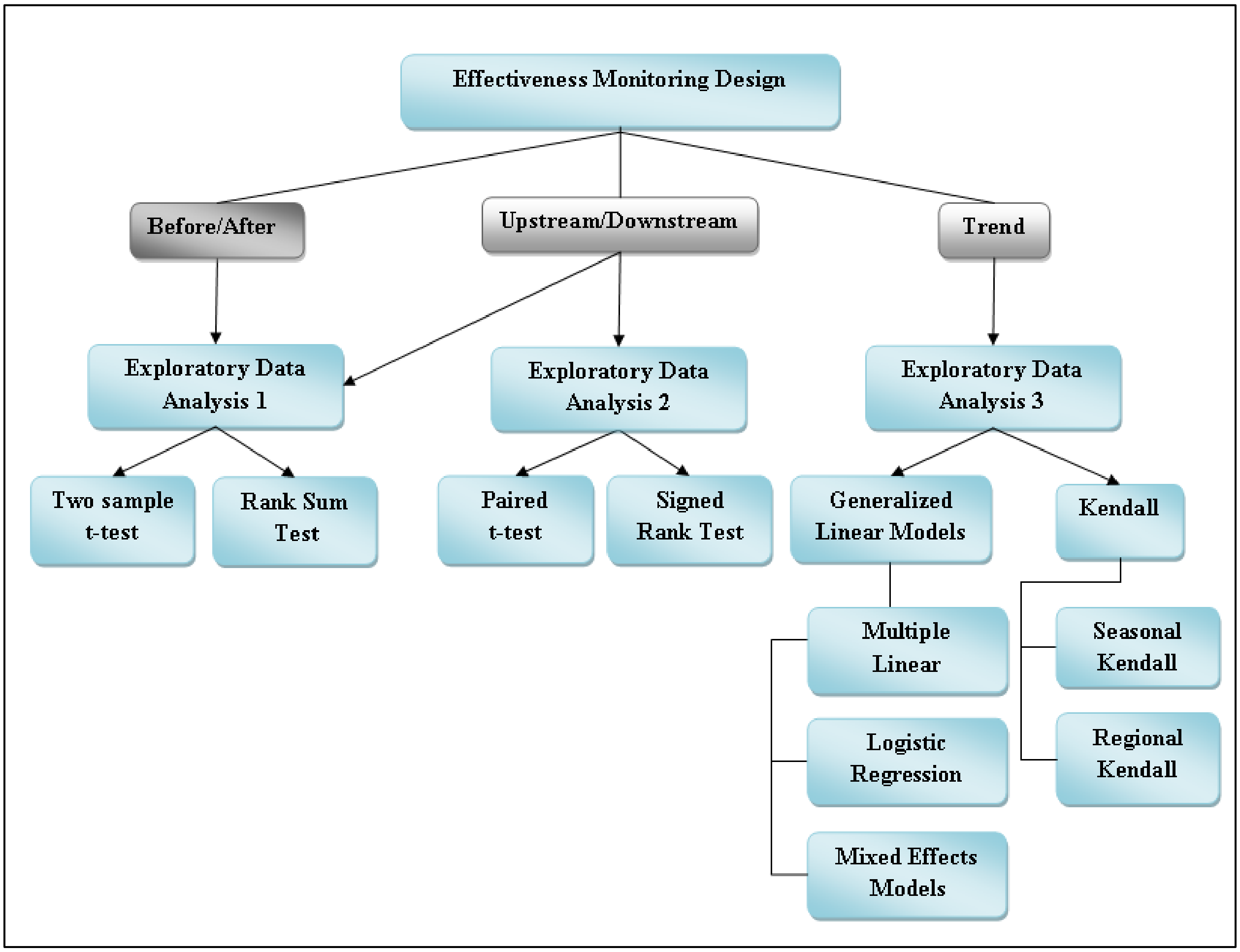


Figure 12. Statistical analysis commonly used for monitoring effectiveness of treatment metrics.

Weight of Evidence Approach

Designing a robust statistical sampling design for assessing effectiveness can be expensive. This is especially true when attempting to measure the cumulative effectiveness of restoration activities on water quality at a watershed scale. Because of this, a surrogate approach to assigning a level of certainty to final results is needed (Diefenderder et al., 2011). Currently, Ecology is exploring a weight of evidence approach to assess data collected under an effectiveness monitoring plan.

A weight of evidence approach relies on correlative data to suggest causal inference. The following criteria are often used to demonstrate causation from the levels-of-evidence (correlative) approach:

- Strength of the Association
- Consistency of Association
- Specificity of Association
- Temporality
- Biological or Ecological Gradient

Although correlative approaches may not carry the same level of certainty of inference that is attached to design-based approaches, they nevertheless can provide strong evidence of treatment effects if they demonstrate the same result.

Quality Assurance

Most of the monitoring activities conducted by Ecology identify the primary use of the data in a QA Project Plan. Ecology's Executive Policy 1-21 states that "A Quality Assurance Project Plan is prepared for each environmental study/activity that acquires or uses environmental measurement data." It further states: "This policy applies to environmental data collection studies/activities conducted or funded by Ecology."

The Guidelines for Preparing Quality Assurance Project Plans for Environmental Studies (Ecology, 2004) describes 14 elements to be addressed in a plan and provides supporting information and examples relevant to the content of each element. QA and quality control responsibilities for management and staff are described in the *Quality Management Plan* (Ecology, 2000). EPA's approval of the Quality Management Plan delegates to Ecology the authority to review and approve QA Project Plans prepared in the agency.

The TMDL effectiveness monitoring plan requires development of QA Project Plans for all monitoring to ensure the scientific validity of monitoring and laboratory activities and also to ensure that reporting requirements are met. QA Project Plans document the planning, implementation, and assessment procedures for a particular project, as well as any specific quality assurance and quality control activities. These plans should reflect the level of data quality that is appropriate for the specific uses of the data, such as listing of impaired waters or pollution control effectiveness. Data quality and quantity needs are expected to vary according to the consequences of the resulting water quality decisions.

TMDL effectiveness monitoring must implement identified quality assurance and control practices to ensure the data are scientifically valid. These quality assurance practices consist of policies, procedures, specifications, standards, and documentation sufficient to produce data of adequate quality to meet project objectives and minimize loss of data due to out-of-control conditions or malfunctions.

The QA Project Plan document is a description of how:

- Each monitoring project objective is defined in specific qualitative and quantitative terms and linked to an environmental management decision.
- Selected indicators offer the most direct means of assessing the environmental condition being evaluated.
- The uncertainty associated with estimates and conclusions drawn are quantified or discussed.
- The sampling scheme will yield data representative of the environmental conditions, based on consideration of statistical probabilities associated with sampling.
- The quality of the data will be assessed and validated to ensure that the data quality objectives of the QA Project Plan were met.

Reporting

Post-project water quality data, compared to pre-project and the data for the technical analysis, form the basis of the effectiveness monitoring report. The report should include a discussion on the link between the pollution controls implemented and the effect they have on water quality.

The following outline may be used to report on effectiveness evaluations:

1. Introduction
2. Background Information
3. Historical Water Quality
4. Discussion of Implemented Pollution Controls
5. Current Water Quality
6. Results of Statistical Analysis
7. Results of Weight of Evidence Analysis
8. Conclusions and Recommendations

Reporting should not be limited to a final report after all implementation has occurred. A process for reporting results to water quality managers in "real time" should be employed throughout the course of the effectiveness monitoring project. This ensures that information used for adaptive management purposes is current and applicable.

Programmatic Roles

This format for TMDL effectiveness monitoring requires coordination and close cooperation between Ecology's Environmental Assessment Program (EAP), Ecology's Water Quality Program (WQP), and the stakeholders.

These planning efforts require the cooperation between EAP and WQP in both content and conclusion of the final TMDL effectiveness monitoring reports. At the programmatic level for developing the report, the responsibilities are to:

- Develop the specific study design – EAP
- Report contents – WQP and EAP
- Prepare the final report – EAP

For the final report, the responsibilities are:

- Introduction – WQP and EAP
- Background Information – EAP
- Project Information – WQP
- Historical Water Quality – EAP
- Current Water Quality – EAP
- Discussion of Implemented Pollution Controls – WQP and EAP
- Conclusion and Recommendations – EAP

Resources

Publications

Table 4 provides a list of publications that can provide additional details on effectiveness monitoring of pollution controls in receiving waters.

Table 4. Publications relevant to developing and implementing effectiveness monitoring plans.

Reference	Title	Topic
EPA, 1991	Monitoring Guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska	Guidance for designing water quality monitoring projects and selecting monitoring parameters
EPA, 1997	Monitoring guidance for determining the effectiveness of nonpoint source control	General study design and data analysis
Hoornbeek et al., 2011	Measuring Water Quality Improvements	Nationwide assessment of progress in implementing Total Maximum Daily Loads
Mesner and Ginger, 2011	Best management practices (BMPs) monitoring guide for stream systems. University of Wyoming.	General guidance on establishing nonpoint-source effectiveness monitoring program.
Ecology, 2012b	Water Quality Standards for Surface Waters of the State of Washington	Water quality standards and designated uses and criteria for Washington State.
Cadmus, 2010	State Approaches and Needs for Measuring, Tracking, and Reporting on Water Quality Improvements.	A summary and overview of how other states are measuring effectiveness of water quality improvement efforts.
Helsel and Hirsch, 2002	Statistical Methods in Water Resources	General statistical methods for analyzing water quality data.

Internet

General Effectiveness Monitoring

Washington State Effectiveness Monitoring Program:
http://www.ecy.wa.gov/programs/eap/tem/index.html#reports

University of Wyoming BMP Monitoring Guidance: http://www.uwyo.edu/bmp-water/

Washington State Implementation

King County Habitat Restoration Projects:
http://www.kingcounty.gov/environment/animalsAndPlants/restoration-projects.aspx

The Puget Sound Nearshore Ecosystem Restoration Project:
http://www.ecy.wa.gov/programs/spills/community_outreach/restoration_projects.html

Washington State Recreation and Conservation Office Project Information System:
http://www.rco.wa.gov/prism_app/about_prism.shtml

EPA's Grants Reporting and Tracking System:
http://iaspub.epa.gov/pls/grts/f?p=110:199:5167501601552::NO:::

City of Bellingham Habitat Restoration Projects:
http://www.cob.org/services/environment/restoration/projects.aspx

Wild Fish Conservancy Projects
http://wildfishconservancy.org/projects

Federal Puget Sound Grant Program
http://www.ecy.wa.gov/puget_sound/grants_fed.html

Washington State Parks Puget Sound-Hood Canal cleanup projects
http://www.parks.wa.gov/cleanwaterprojects/

Washington Water Trust
http://www.washingtonwatertrust.org/

Nisqually Delta Restoration
http://nisquallydeltarestoration.org/

NOAA Pacific Northwest Salmon Habitat Restoration Project Tracking Database
https://www.webapps.nwfsc.noaa.gov/pnshp/

Summary of Key Concepts

Key concepts of effectiveness monitoring programs are listed below.

- TMDL effectiveness monitoring measures the cumulative effect of all restoration activities on water quality.
- Monitoring effectiveness requires the use of multiple types of monitoring.
- Implementation monitoring is crucial to the process.
- Effectiveness monitoring plans need to be developed before implementation of BMPs.
- Effectiveness monitoring plans need to include:
  - Specific goals and objectives.
  - A long-term monitoring schedule that is followed through the life of the TMDL.
  - A description of what will be monitored and a plan for how data will be analyzed.
  - A strategy for adaptive management.
- Developing local partnerships and leveraging existing monitoring efforts helps meet the goals of the study.
- Results of effectiveness monitoring evaluations must have some level of precision (i.e., statistical design).
- A weight of evidence approach should be employed when disseminating results of effectiveness monitoring studies.
- Results should be reported throughout the process in order to inform the adaptive management process.
- Reports must convey results in such a way they can be used by a diverse group of people to make relevant decisions.

References

Baldwin, K, D. Gray, and J. Jones, 2008. Walla Walla Watershed PCBs, Chlorinated Pesticides, Fecal Coliform, Temperature, pH and Dissolved Oxygen Total Maximum Daily Load: Water Quality Implementation Plan. Washington State Department of Ecology, Olympia, WA. Publication No. 08-10-094.
https://fortress.wa.gov/ecy/publications/SummaryPages/0810094.html

Brown, M.T., and M. B. Vivas. 2005. Landscape development intensity index. Environmental Monitoring and Assessment (2005) 101:289-309.

The Cadmus Group, 2010. State approaches and needs for measuring, tracking, and reporting on water quality improvements. Washington State Department of Ecology, Olympia, WA. Publication No. 10-03-017.
https://fortress.wa.gov/ecy/publications/SummaryPages/1003017.html

Collyard, S., 2011. White Salmon River Watershed Fecal Coliform bacteria Attainment Monitoring Study. Washington State Department of Ecology, Olympia, WA. Publication No. 11-03-046. https://fortress.wa.gov/ecy/publications/summarypages/1103046.html

Collyard, S., and M. Von Prause, 2009. Deschutes River Multi-Parameter Total Maximum Daily Load Effectiveness Monitoring Pilot Project. Water Quality Study Design (Quality Assurance Project Plan). Washington State Department of Ecology, Olympia, WA. Publication No. 09-03-133. https://fortress.wa.gov/ecy/publications/SummaryPages/0903133.html

Collyard, S., M. Von Prause, and G. Onwumere, 2012. Using biological and habitat metrics to determine the effectiveness of TMDLs: A Case Study. Washington State Department of Ecology, Olympia, WA. Publication No. 12-03-033.
https://fortress.wa.gov/ecy/publications/summarypages/1203033.html

Diefenderder, H.L., R.M. Thom, G.E. Johnson, J.R. Skalski, K.A. Vogt, B.D. Ebberts, G.C. Roegner, and E.M. Dawley, 2011. A Level of Evidence Approach for Assessing Cumulative Ecosystem Response to Estuary and River Restoration Programs. Ecological Restoration, Vol. 29, Nos. 1-2.

Dzinbal, K. and S. Butkus, 2005. Clean Water Act Monitoring Strategy For Washington State. Washington State Department of Ecology, Olympia, WA. Publication No. 05-03-034.
https://fortress.wa.gov/ecy/publications/summarypages/0503034.html

Ecology, 2000. Quality Management Plan. Washington State Department of Ecology, Olympia, WA. Publication No. 00-03-012.
https://fortress.wa.gov/ecy/publications/summarypages/0003012.html

Ecology, 2004. Guidelines for Preparing Quality Assurance Project Plans for Environmental Studies. Washington State Department of Ecology, Olympia, WA. Publication No. 04-03-030.
http://www.ecy.wa.gov/biblio/0403030.html

Ecology, 2012a. Funding for Water Quality Improvement and Protection. Washington State Department of Ecology, Olympia, WA. Publication No. 08-10-013. https://fortress.wa.gov/ecy/publications/summarypages/0810013.html

Ecology, 2012b. Water Quality Standards for Surface Waters of the State of Washington. Chapter 173-201A WAC. Washington State Department of Ecology, Olympia WA. Publication No. 06-10-091. https://fortress.wa.gov/ecy/publications/summarypages/0610091.html

Ecology, 2013. DRAFT Clean Water Act Monitoring Strategy for Washington State. Environmental Assessment Program, Washington State Department of Ecology, Olympia, WA.

EPA, 1991. Monitoring Guidelines to Evaluate Effects of Forestry Activities on Stream in the Pacific Northwest and Alaska. U.S. Environmental Protection Agency, Water Division, EPA/910/9-91-001.

EPA, 1997. Monitoring guidance for determining the effectiveness of nonpoint source controls. U.S. Environmental Protection Agency, Office of Water, EPA 841-B-96-004.

EPA, 2003. Elements of a State Water Monitoring and Assessment Program. Assessment and Watershed Protection Division Office of Wetlands, Oceans and Watershed. U.S. Environmental Protection Agency. EPA 841-B-03-003.

EPA, 2010a. Causal Analysis/Diagnosis Decision Information System (CADDIS). U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC. http://www.epa.gov/caddis.

EPA, 2010b. Identification of Most Probable Stressors to Aquatic Life in the Touchet River, Washington (Final). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-08/145, 2010.

Dzinbal, K. and S. Butkus, 2005. Clean Water Act Monitoring Strategy For Washington State. Washington State Department of Ecology, Olympia, WA. Publication No. 05-03-034. https://fortress.wa.gov/ecy/publications/summarypages/0503034.html

Haake, D.M., T. Wilton, K. Krier, A.J. Stewart, and S.M. Cormier, 2010. Causal assessment of biological impairment in the Little Floyd River, Iowa, USA. Human and Ecological Risk Assessment 16(1):116-148.

Hallock, D., 2011. River and Stream Water Quality Monitoring Report, Water Year 2011. Washington State Department of Ecology, Olympia, WA. Publication No. 12-03-039. https://fortress.wa.gov/ecy/publications/SummaryPages/1203039.html

Helsel, D.R., and R.M. Hirch, 2002. Statistical methods in Water Resources. Techniques of Water-Resources Investigations of the United States Geological Survey. Book 4, Hydrologic Analysis and Interpretation. http://water.usgs.gov/pubs/twri/twri4a3/

Hoornbeek, J., E. Hansen., A. Hereford, J. Filla, and S. Satpathi, 2011. Measuring Water Quality Improvements: TMDL Implementation Progress, Indicators, and Tracking. Kent State University, the Center for Public Administration and Public Policy. http://www.kent.edu/cpapp/water-environment-energy/our-research.cfm

McBride, R., 2000. Responsiveness Summary: Fiscal Year 2000 TMDL Priority List. Washington State Department of Ecology, Olympia, WA. 23 p. Publication No. 00-10-055. https://fortress.wa.gov/ecy/publications/summarypages/0010055.html

Mesner, N. and P. Ginger, 2011. Best management practices monitoring guide for stream systems. University of Wyoming. http://www.uwyo.edu/bmp-water/docs/bmp%20mon%20guide%20streams%20web.pdf

Murray, C. and D. Marmorek. 2003. Adaptive Management and Ecological Restoration. Chapter 24, in: Freiderici, P. (ed.). 2003. Ecological Restoration of Southwestern Ponderosa Pine Forests. Island Press (Washington ,Covelo CA, London), pp. 417-428.

Wiseman, C.D., M. LeMoine, R. Plotnikoff, J. Diamond, A. Stewart, S. Cormier, 2009. Identification of Most Probable Stressors to Aquatic Life in the Touchet River, Washington. U.S. Environmental Protection Agency, Cincinnati OH. EPA/600/R 08/145.

Appendices

This page is purposely left blank

Appendix A. Glossary, Acronyms, and Abbreviations

Glossary

Anthropogenic: Human-caused.

Clean Water Act: A federal act passed in 1972 that contains provisions to restore and maintain the quality of the nation's waters. Section 303(d) of the Clean Water Act establishes the TMDL program.

National Pollutant Discharge Elimination System (NPDES): National program for issuing, modifying, revoking and reissuing, terminating, monitoring, and enforcing permits, and imposing and enforcing pretreatment requirements under the Clean Water Act. The NPDES program regulates discharges from wastewater treatment plants, large factories, and other facilities that use, process, and discharge water back into lakes, streams, rivers, bays, and oceans.

Nonpoint source: Pollution that enters any waters of the state from any dispersed land-based or water-based activities, including but not limited to atmospheric deposition, surface-water runoff from agricultural lands, urban areas, or forest lands, subsurface or underground sources, or discharges from boats or marine vessels not otherwise regulated under the NPDES program. Generally, any unconfined and diffuse source of contamination. Legally, any source of water pollution that does not meet the legal definition of "point source" in section 502(14) of the Clean Water Act.

Parameter: Water quality constituent being measured (analyte). A physical, chemical, or biological property whose values determine environmental characteristics or behavior.

Point source: Sources of pollution that discharge at a specific location from pipes, outfalls, and conveyance channels to a surface water. Examples of point source discharges include municipal wastewater treatment plants, municipal stormwater systems, industrial waste treatment facilities, and construction sites that clear more than 5 acres of land.

Pollution: Contamination or other alteration of the physical, chemical, or biological properties of any waters of the state. This includes change in temperature, taste, color, turbidity, or odor of the waters. It also includes discharge of any liquid, gaseous, solid, radioactive, or other substance into any waters of the state. This definition assumes that these changes will, or are likely to, create a nuisance or render such waters harmful, detrimental, or injurious to (1) public health, safety, or welfare, or (2) domestic, commercial, industrial, agricultural, recreational, or other legitimate beneficial uses, or (3) livestock, wild animals, birds, fish, or other aquatic life.

Stormwater: The portion of precipitation that does not naturally percolate into the ground or evaporate but instead runs off roads, pavement, and roofs during rainfall or snow melt. Stormwater can also come from hard or saturated grass surfaces such as lawns, pastures, playfields, and from gravel roads and parking lots.

Total Maximum Daily Load (TMDL): Water cleanup plan. A distribution of a substance in a waterbody designed to protect it from not meeting (exceeding) water quality standards. A TMDL is equal to the sum of all of the following: (1) individual wasteload allocations for point sources, (2) the load allocations for nonpoint sources, (3) the contribution of natural sources, and (4) a Margin of Safety to allow for uncertainty in the wasteload determination. A reserve for future growth is also generally provided.

Watershed: A drainage area or basin in which all land and water areas drain or flow toward a central collector such as a stream, river, or lake at a lower elevation.

Acronyms and Abbreviations

BMP	Best management practice
EAP	Environmental Assessment Program
Ecology	Washington State Department of Ecology
EPA	U.S. Environmental Protection Agency
GIS	Geographic Information System software
NPDES	(See Glossary above)
QA	Quality assurance
TMDL	(See Glossary above)
WQP	Water Quality Program

Appendix B. Types of Monitoring Used in Effectiveness Monitoring Evaluations (EPA, 1991)

Baseline monitoring: Baseline monitoring is used to characterize existing water quality conditions and to establish a database for planning or future comparisons. The intent of baseline monitoring is to capture much of the temporal variability of the constituent(s) of interest, but there is no explicit endpoint at which continued baseline monitoring becomes trend monitoring.

Status monitoring: The purpose of this type of monitoring is a population scale assessment used to estimate the overall status of environmental conditions. One of the most often asked questions is: *What is the overall status of water quality in my watershed?* This is the same question the state is required to report on statewide conditions for the report to Congress under the federal Clean Water Act (Section 305(b)). It is impossible to conduct a full census of conditions by monitoring every surface water in the state to answer this question. The approach instead is to randomly sample a subset of surface waters to infer conditions over the scale of the assessment (i.e., statewide or at a watershed scale). This approach, known as *sample survey monitoring design*, provides a statistically representative view of sampled conditions over a broad spatial scale. This monitoring provides fundamental information on baseline conditions and complements other types of monitoring.

Trend Monitoring: Trend monitoring tracks changes in environmental conditions over time. Trends can be determined at the site scale or on the population scale depending on the monitoring design. Data are collected on a routine basis over time and analyzed with statistical tests. Formal statistical trend analysis provides a rational, scientific basis for addressing issues that can be confused by natural variations in watershed conditions. The length of time required to detect a trend is based on the variation of the environmental indicator being measured. The frequency of required monitoring depends on the statistical independence of the monitoring data. The presence or absence of trends is a good indication of the degree to which watershed health indicators are responding to changes in the watershed.

Compliance Monitoring: This type of monitoring tracks compliance with established laws, rules, or benchmarks. Compliance of monitoring determines whether specified criteria (e.g., state water quality standards) are being met. The criteria can be numeric or descriptive. Generally, regulations associated with individual criterion specify the location, frequency, and method of measurement. Washington State is required under the federal Clean Water Act (Section 303(d)) to periodically assess water quality and prepare a list of waters with impaired beneficial uses. To meet this requirement, Ecology routinely evaluates water quality data compiled from numerous sources to determine compliance with the state water quality standards. Waters impaired by human-caused pollutants require further pollution controls.

Implementation Monitoring: Implementation monitoring assesses whether activities were carried out as planned. The most common use of implementation monitoring is to determine whether best management practices (BMPs) were implemented as specified in an environmental assessment, environmental impact statement, other planning document, or contract. Typically, this type of monitoring is carried out as an administrative review and does not involve any water quality measurements. *Implementation monitoring* is one of the few terms which has a relatively widespread and consistent definition. Many believe that implementation monitoring is the most cost-effective means to reduce nonpoint source pollution because it provides immediate feedback to the managers on whether the BMP process is being carried out as intended. On its own, however, implementation monitoring cannot directly link management activities to water quality, because no water quality measurements are being made.

Source Identification Monitoring: Source identification monitoring is the use of water quality measurements or other technologies to identify sources of pollution to waterbodies. It can help watershed managers determine where additional pollutant controls need to be implemented if standards or targets are not meet. Integrating source identification into an effectiveness monitoring evaluation provides a process of adaptive management.

Validation Monitoring: This refers to the quantitative evaluation of proposed water quality model. The data set used for validation should be different from the data set used to construct and calibrate the model. This separation helps ensure that the validation data will provide an unbiased evaluation of the overall performance of the model. The intensity and type of sampling for validation monitoring should be consistent with the output of the model being validated.

Effectiveness Monitoring: While implementation monitoring is used to assess whether a particular activity was carried out as planned, effectiveness monitoring is used to evaluate whether the specified activities had the desired effect. Confusion arises over whether effectiveness monitoring should be limited to evaluating individual BMPs or whether it also can be used to evaluate the total effect of an entire set of practices. The problem with this broader definition is that the distinction between effectiveness monitoring and other terms, such as project or compliance monitoring, becomes blurred. Monitoring the effectiveness of individual BMPs, such as the spacing of water bars on skid trails, is an important part of the overall process of controlling nonpoint source pollution. However, in most cases the monitoring of individual BMPs is quite different from monitoring to determine whether the cumulative effect of all the BMPs results in adequate water quality protection. Evaluating individual BMPs may require detailed and specialized measurements best made at the site of, or immediately adjacent to, the management practice. Thus, effectiveness monitoring often occurs outside of the stream channel and riparian area, even though the objective of a particular practice is intended to protect the designated uses of a waterbody. In contrast, monitoring the overall effectiveness of BMPs usually is done in the stream channel, and it may be difficult to relate these measurements to the effectiveness of individual BMPs.

Ambient Monitoring: Ambient water quality monitoring is a data-gathering tool used for almost all water quality assessments. Monitoring programs serve to identify waters needing TMDLs, quantify loads, verify models, and evaluate effectiveness of water quality controls (including BMP effectiveness). Once TMDLs have been established for a given waterbody, follow-up monitoring is recommended to document improvement or lack of improvement. Since the TMDL process is iterative, monitoring data can provide the information for updating and revising current TMDLs. Ambient monitoring is used for setting permit conditions, compliance, and enforcement, and detecting new problems and trends.

Appendix C. Summary of EPA's Nine Key Elements for Section 319 Funding

Introduction

All projects that apply for Section 319 funding under the Clean Water Act and which are administered through the Washington State Department of Ecology must include nine key elements in their watershed-based plans. These elements are listed in the Nonpoint Source Guidance Document by EPA (http://www.epa.gov/polwaste/nps/cwact.cfm).

Element a.

Identification of causes of impairment and pollutant sources or groups of similar sources that need to be controlled to achieve needed load reductions, and any other goals identified in the watershed plan.

Sources that need to be controlled should be identified at the significant subcategory level along with estimates of the extent to which they are present in the watershed (e.g., X number of dairy cattle feedlots needing upgrading, including a rough estimate of the number of cattle per facility; Y acres of row crops needing improved nutrient management or sediment control; or Z linear miles of eroded streambank needing remediation).

What does this mean?

Your watershed-based plan (WBP) source assessment should encompass the watershed of any impaired waterbody being restored, and include map of the watershed that locates the major causes and sources of impairment in the planning area. To address these impairments, you will set goals to (at a minimum) meet the appropriate water quality standards for pollutants that threaten or impair the physical, chemical, or biological integrity of the watershed covered in the plan.

This element will usually include an accounting of the significant point and nonpoint sources in addition to the natural background levels that make up the pollutant loads causing problems in the watershed. If TMDLs exist for the waters under consideration, this element may be adequately addressed in those documents. If not, you will need to conduct a similar analysis (which may involve mapping, modeling, monitoring, and field assessments) to make the link between the sources of pollution and the extent to which they cause the water to exceed relevant water quality standards.

Element b.

An estimate of the load reductions expected from management measures.

What does this mean?

On the basis of the existing source loads estimated for element *a*, you will similarly determine the reductions needed to meet water quality standards. After identifying the various management measures that will help to reduce the pollutant loads (see element *c* below), you will estimate the load reductions expected as a result of implementing these management measures, recognizing the difficulty in precisely predicting the performance of management measures over time. Estimates should be provided at the same level as that required in the scale and scope described in element *a* (e.g., the total load reduction expected for dairy cattle feedlots, row crops, eroded streambanks, or implementation of a specific stormwater management practice). For waters for which TMDLs have been approved or are being developed, the plan should identify and incorporate the TMDLs; the plan needs to be designed to achieve the applicable load reductions in the TMDLs. Applicable loads for downstream waters should be included so that water delivered to a downstream or adjacent segment does not exceed the water quality standards for the pollutant of concern at the water segment boundary. The estimate should account for reductions in pollutant loads from point and nonpoint sources identified in the TMDL as necessary to attain the applicable water quality standards.

Element c.

A description of the nonpoint source management measures that will need to be implemented to achieve load reductions in element b, and a description of the critical areas in which those measures will be needed to implement this plan.

What does this mean?

The plan should describe the management measures that need to be implemented to achieve the load reductions estimated under element b, as well as to achieve any additional pollution prevention goals outlined in the watershed plan (e.g., habitat conservation and protection). Pollutant loads will vary even within land use types, so the plan should also identify the critical areas in which those measures will be needed to implement the plan. This description should be detailed enough to guide needed implementation activities throughout the watershed and can be greatly enhanced by developing an accompanying map with priority areas and practices. Thought should also be given to the possible use of measures that protect important habitats (e.g. wetlands, vegetated buffers, and forest corridors) and other non-polluting areas of the watershed. In this way, waterbodies would not continue to degrade in some areas of the watershed while other parts are being restored.

Element d.

Estimate of the amounts of necessary technical and financial assistance, associated costs, and/or support of sources and authorities to implement this plan.

What does this mean?

You should estimate the financial and technical assistance needed to implement the entire plan. This includes implementation and long-term operation and maintenance of management measures, information/education (I/E) activities, monitoring, and evaluation activities. You should also document which relevant authorities might play a role in implementing the plan. Plan sponsors should consider the use of federal, state, local, and private funds or resources that might be available to assist in implementing the plan. Shortfalls between needs and available resources should be identified and addressed in the plan.

Element e.

An information and education component used to enhance public understanding of the plan and encourage their early and continued participation in selecting, designing, and implementing the nonpoint source management measures that will be implemented.

What does this mean?

The plan should include an I/E component that identifies the education and outreach activities that will be used to implement the plan. These I/E activities may support the adoption and long-term operation and maintenance of management practices and support stakeholder involvement efforts.

Element f.

A schedule for implementing the nonpoint source management measures identified in this plan that is reasonably expeditious.

What does this mean?

You should include a schedule for implementing the management measures outlined in your watershed plan. The schedule should reflect the milestones you develop in *g* and begin implementation as soon as possible. Activities that can start right away include: conducting baseline monitoring and outreach for implementing water quality projects. It is important that schedules not be “shelved” for lack of funds or program authorities; instead they should identify steps toward obtaining needed funds, as feasible.

Element g.

A description of interim measurable milestones for determining whether nonpoint source management measures or other control actions are being implemented.

What does this mean?

The WBP should include interim, measurable implementation milestones to measure progress in implementing the management measures. These milestones will be used to track implementation of the management measures, such as whether they are being implemented according to the schedule outlined in element *f*, whereas element *h* (see below) will develop criteria to measure the effectiveness of the management measures by, for example, documenting improvements in water quality.

Element h.

A set of criteria that can be used to determine whether loading reductions are being achieved over time and substantial progress is being made toward attaining water quality standards.

What does this mean?

As projects are implemented in the watershed, you will need water quality benchmarks to track progress toward attaining water quality standards. The *criteria* in element *h* (not to be confused with *water quality criteria* in state regulations) are the benchmarks or waypoints to measure against through monitoring. These interim targets can be direct measurements (e.g., fecal coliform concentrations, nutrient loads) or indirect indicators of load reduction (e.g., number of beach closings). These criteria should reflect the time it takes to implement pollution control measures, as well as the time it would take for water quality indicators to respond, including lag times. For example, consider water quality response to slowly moving groundwater sources or the extra time for sediment-bound pollutants to break down, degrade, or otherwise be isolated from the water column. Appendix B of these guidelines, "Measures and Indicators of Progress and Success," although intended as measures for program success, may provide some useful examples. You should also indicate how you will determine whether the WBP needs to be revised if interim targets are not met. These revisions could involve changing management practices, updating the loading analyses, and reassessing the timeframe for pollution concentrations to respond to treatment.

Element i.

A monitoring component to evaluate the effectiveness of the implementation efforts over time, measured against the criteria established under element h.

What does this mean?

The WBP should include a monitoring component to determine whether progress is being made toward attaining or maintaining the applicable water quality standards for each waterbody addressed in the plan. The monitoring program should be fully integrated with the established schedule and interim milestone criteria identified above. The monitoring component should be designed to assess progress in achieving loading reductions and meeting water quality standards. Watershed-scale monitoring can be used to measure the effects of multiple programs, projects, and trends over time. Instream monitoring does not have to be conducted for individual BMPs unless that type of monitoring is particularly relevant to the project.

For more detailed information on developing watershed-based plans, see *A Handbook for Developing Watershed Plans to Restore and Protect Our Waters*, U.S. EPA, EPA 841-B-08-002 March 2008, www.water.epa.gov/polwaste/nps/handbook_index.cfm
Other resources for watershed planning are available on the Watershed Central website - including the Watershed Central Wiki and Plan Builder tool at
water.epa.gov/type/watersheds/datait/watershedcentral/index.cfm

Appendix D. Technical Guidance for Designing a TMDL Effectiveness Monitoring Plan

This page is purposely left blank

Technical Guidance for Designing a TMDL Effectiveness Monitoring Plan

1. Introduction

Water quality impairments in lakes, rivers, and other water bodies are often addressed through the Total Maximum Daily Load (TMDL) process. Included in the TMDL process is the determination of the maximum pollutant amount a water body can receive without exceeding water quality standards, and the allocation of this amount between point and nonpoint sources of pollution. Additionally, the TMDL process can include the development of guidelines for TMDL implementation, or water quality management initiatives that are needed to achieve the TMDL target.

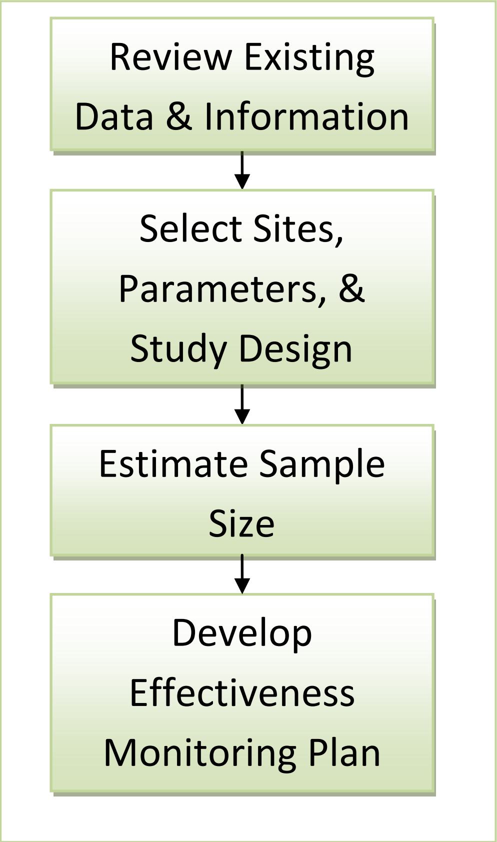


Figure 1. Recommended steps for TMDL effectiveness monitoring planning.

A fundamental yet often overlooked component of TMDL implementation is ***TMDL Effectiveness Monitoring***. The primary goal of TMDL effectiveness monitoring is to identify water quality improvements (or lack thereof) that result from TMDL implementation. This information serves as an important source of feedback for refining and optimizing management approaches. Like any project involving data collection, the value of TMDL effectiveness monitoring can be greatly enhanced if monitoring activities are preceded by thorough and detailed project planning. Without proper planning, effectiveness monitoring may not produce the type and quantity of data needed to detect water quality changes. Oftentimes, the need for planning is not apparent until a monitoring project is underway or complete. In a review of stream restoration 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 an improved scientific evaluation of project effectiveness (Rumps, et al., 2007). This document is intended to serve as a guide for water quality practitioners planning a TMDL effectiveness monitoring project. Several steps are outlined that inform the design of a detailed TMDL effectiveness monitoring plan (Figure 1).

Specific topics covered in this guide include:

- The benefits of watershed scale TMDL effectiveness monitoring;
- Sources for data review and data review outcomes;
- Selecting sampling parameters and monitoring sites;
- Estimating sample size requirements using power analysis; and
- Integration of these and other considerations into a final TMDL effectiveness monitoring plan.

A key recommendation in this guide is for planners to make data-driven planning decisions using a pilot dataset that captures the general characteristics of the data to be collected through TMDL effectiveness monitoring. A discussion of pilot data exploration methods can be found in a companion document titled *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*. Additionally, an Excel-based TMDL effectiveness monitoring planning Tool has been provided as a complement to this document to facilitate data-driven planning.

2. Watershed Scale TMDL Effectiveness Monitoring

The watershed approach to water resource management is becoming the standard model for maintaining and improving the quality of surface waters in the U.S. The watershed approach considers linkages among landscape conditions, ecological resources, and multiple interconnected water bodies throughout a land area to guide adaptive water resource management strategies.

The U.S. Environmental Protection Agency (EPA) has promoted the watershed approach for TMDL development rather than single-segment TMDLs as a means to streamline and improve the TMDL process. A watershed TMDL incorporates watershed-wide information on pollutant types and sources to simultaneously complete TMDLs for all impaired waters within the watershed. Specific benefits of using the watershed approach for TMDLs include (U.S. Environmental Protection Agency, 2008):

- The ability to conduct a broad assessment of pollutant sources;
- The ability to capture the interaction between upstream and downstream sources and impacts;
- The restoration of unimpaired but threatened waters and/or protection of high-quality waters;
- Reduced per-TMDL costs;
- The involvement of multiple stakeholders; and
- The opportunity to integrate TMDLs with other watershed programs.

Each of these benefits is relevant to watershed scale TMDL effectiveness monitoring. Relative to a single-segment effectiveness monitoring project, watershed monitoring has the potential to be more efficient and informative due to the integration of a broad assessment of pollutant sources, and knowledge of interactions between upstream and downstream pollutant sources, impacts, and controls into project planning. Further, watershed scale effectiveness monitoring can demonstrate water quality improvements in waters that are at risk for impairment, and can take advantage of partnerships with interested stakeholders and existing data collection programs.

It is highly recommended that practitioners design a TMDL effectiveness monitoring project using the watershed framework. Note that the term "watershed" itself does not explicitly denote the actual scale of the project (i.e., the size of the study area), as a watershed can range from a small headwater drainage to a major river basin. The project scale should be decided upon by project planners using information on the number and extent of impaired or threatened waters in the region of interest, project resources, and project partners. One important point to consider is that water quality improvements that are demonstrated for projects that include one or more watersheds at the 12-digit hydrologic unit code (HUC-12) scale can be reported under EPA's National Water Program Guidance Measure SP-12[1]. Measure SP-12 is intended to measure and track incremental improvements in water quality at the HUC-12 watershed scale that are attributable to implementation measures using a watershed approach. Measure SP-12 is a useful framework on which to build an effectiveness monitoring program because it requires use of a watershed approach and demonstration of statistically significant changes or multiple lines of evidence indicating improvement.

3. Review Existing Data & Information

The design of a TMDL effectiveness monitoring project begins with a thorough review of all available information that may direct the process, including:

- TMDL report(s) and implementation plan(s);
- Water quality reports;
- Locations of best management practices (BMPs) and other TMDL implementation actions;
- Timing of TMDL implementation;
- Existing watershed management plans and stakeholder groups;
- Existing water quality monitoring sites and data;
- TMDL effectiveness monitoring project goals and resources (staff, funding, etc.); and
- Other sources describing watershed features and conditions.

---

[1] http://water.epa.gov/resource_performance/planning/def_wq11.cfm#SP-12

The above information should provide an understanding of: current and historic water quality conditions, including short- (daily/seasonal) and long- (annual) term variability; and the type(s), location, and timing of TMDL implementation actions within the watershed. Planners should have a firm grasp of the ecological resources in the watershed and how these resources have been impacted by degraded water quality. Finally, planners should understand how upstream management activities can potentially influence downstream water quality conditions throughout the watershed.

Ultimately, the review of existing data and information will provide direction on selecting TMDL effectiveness monitoring sites. In general, a monitoring site should be located where TMDL implementation is expected to have discernible water quality effects. This includes sites on impaired or degraded water bodies that are located downstream of:

- Point sources with new or revised wasteload allocations;
- Discontinued illicit discharges;
- Nonpoint sources that are managed through BMPs;
- Stream channel restoration projects;
- Improved onsite wastewater management or expansion of sanitary sewer service; and
- Other TMDL-specific pollution control measures.

Knowledge of pollutant types/sources, TMDL implementation, ecological conditions, and watershed characteristics gained from the information review will also inform decisions regarding the selection of monitoring parameters and methods for demonstrating water quality improvements. These steps, and further guidance for site selection, are presented in the following section.

4. Select Monitoring Sites, Parameters, and Study Design

As discussed in Section 3, the selection of TMDL effectiveness monitoring sites will be based on a review of existing information to identify locations where water quality improvements are expected to occur. Potential monitoring locations can be further refined based on project goals and resources. A central goal of watershed scale TMDL effectiveness monitoring should be the demonstration of watershed-wide water quality improvements. Two general methods are available to meet this goal. A basic approach is to perform water quality monitoring at the watershed outlet (the ***pour point*** method) (Figure 2). Under this option, the cumulative effect of all TMDL implementation actions in the watershed is evaluated using data collected at the watershed outlet. The pour point method is well-suited for projects in which water quality is known to be degraded at the watershed outlet and where limited monitoring resources are available.

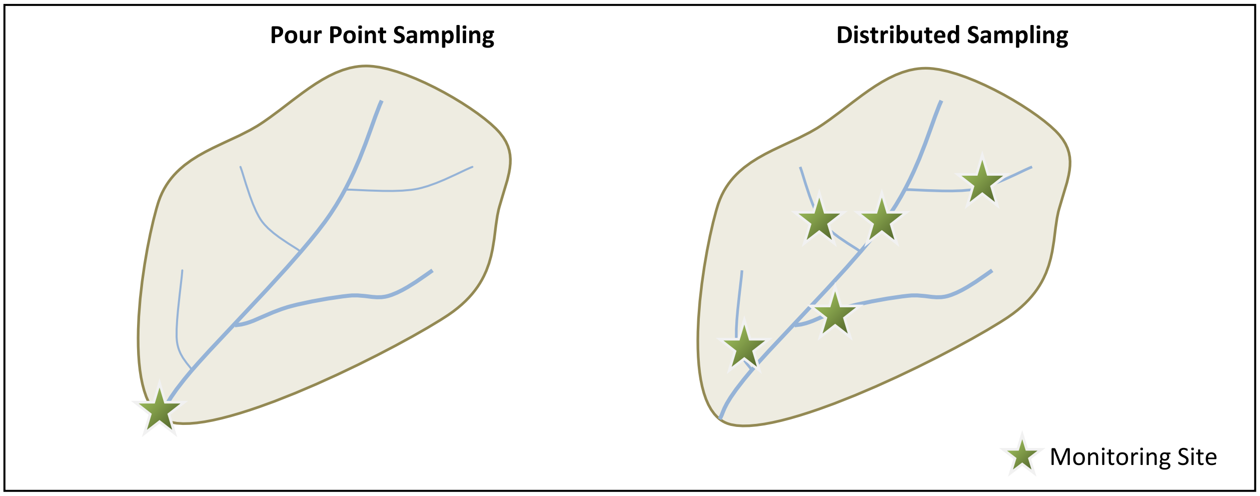


Figure 2. Illustration of the pour point and distributed sampling methods.

A more intensive approach for demonstrating watershed-wide water quality improvements is through the collection of water quality data at multiple sites within the watershed (the ***distributed sampling*** method) (Figure 2). Under this approach, conclusions drawn for individual sites can be pieced together to infer watershed-wide water quality improvements. Distributed sampling can be applied when project resources or partnerships allow for a rigorous sampling effort. Benefits of distributed sampling include the ability to assess the effects of individual TMDL implementation actions. Such information is key to adaptive, cost-effective management. A distributed sampling approach may also be preferred when project goals include monitoring to justify changes to the impairment status of multiple water body segments in the watershed (e.g., from an impaired water body with an approved TMDL to a water body that meets water quality standards).

In many cases, the location of potential monitoring sites will coincide with an existing water quality monitoring network. In such cases, existing monitoring sites should be reviewed to determine the past and present status of monitoring activities, parameters monitored, data quality, and monitoring organization(s). If existing monitoring activities agree with the needs/goals of the TMDL effectiveness monitoring project, these sites and existing data can be integrated into the project.

If newly-established monitoring sites are needed for TMDL effectiveness monitoring, planners are faced with the additional task of identifying their precise location. The following are a number of site-specific characteristics to consider when identifying sampling sites.

Site Accessibility and Safety

- Consider road access, parking availability, and ownership of adjacent land.
- Ensure that access to the site is not obstructed by natural or manmade features.
- Avoid steep banks and/or other hazardous features (wire or other debris).
- Be aware of the presence of poisonous plants or poisonous animal habitat.

- Note hazardous flow and channel conditions (rapidly flowing water, sudden changes in water depth).
- Consider how each of the above may change seasonally.

Physical Conditions

- Be sure that discharges from tributaries, point sources, or groundwater are well-mixed with upstream flows.
- Avoid areas with a high potential for damage/loss of sampling equipment.

Logistical Considerations

- Note travel time relative to other monitoring sites.
- Consider travel time and maximum holding time for lab analysis.

In addition to the selection of sampling locations, planners must also determine which water quality parameters will be monitored. In general, this will include the pollutant(s) addressed by the TMDL(s) developed for water bodies in the watershed. Project resources may allow for monitoring of additional parameters that relate to water body impairments. These can include stressor variables (e.g., nutrients, bacteria, sediment) and/or response variables (e.g., macroinvertebrates, fish counts, riparian habitat). Monitoring of response variables can be highly informative, as improved biological diversity, habitat, etc. may be observed prior to detectable improvements in water quality and can provide supporting evidence for watershed-wide water quality improvement under Measure SP-12.

Monitoring of additional water quality parameters can be useful if those parameters are statistically associated with the primary pollutant(s) of interest. Such ***covariates*** can improve the power of subsequent statistical analyses (see *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*). A common covariate for pollutants in streams and rivers is flow magnitude (streamflow). Streamflow monitoring should be conducted in conjunction with water quality monitoring whenever possible, as stream flow data allow for the calculation of pollutant loads, and can improve the analysis of water quality data. To reduce costs, planners may be able to situate monitoring sites near existing streamflow gaging stations.

The collection of water quality data under a TMDL effectiveness monitoring program should be completed in the context of a specific study design. The study design formally outlines how water quality improvements will be demonstrated. Haphazard study design decisions can derail an otherwise well-planned TMDL effectiveness monitoring program and result in the collection of data that is unfit for achieving program goals. Potential study designs for TMDL effectiveness monitoring are depicted in Figure 3 and include:

- Before/After Study;
- Upstream/Downstream Study;
- Paired Watersheds Study; and
- Trend Monitoring Study.

Study design selection is dependent on multiple factors. Planners need to consider the type(s) of TMDL implementation actions, implementation schedule, the availability and quality of previously collected data, project resources, and the existence of suitable reference sites. Monitoring duration and sample size requirements are also key pieces of information to consider. Sample size requirements for a particular study design can be estimated using Power Analysis (see Section 5).

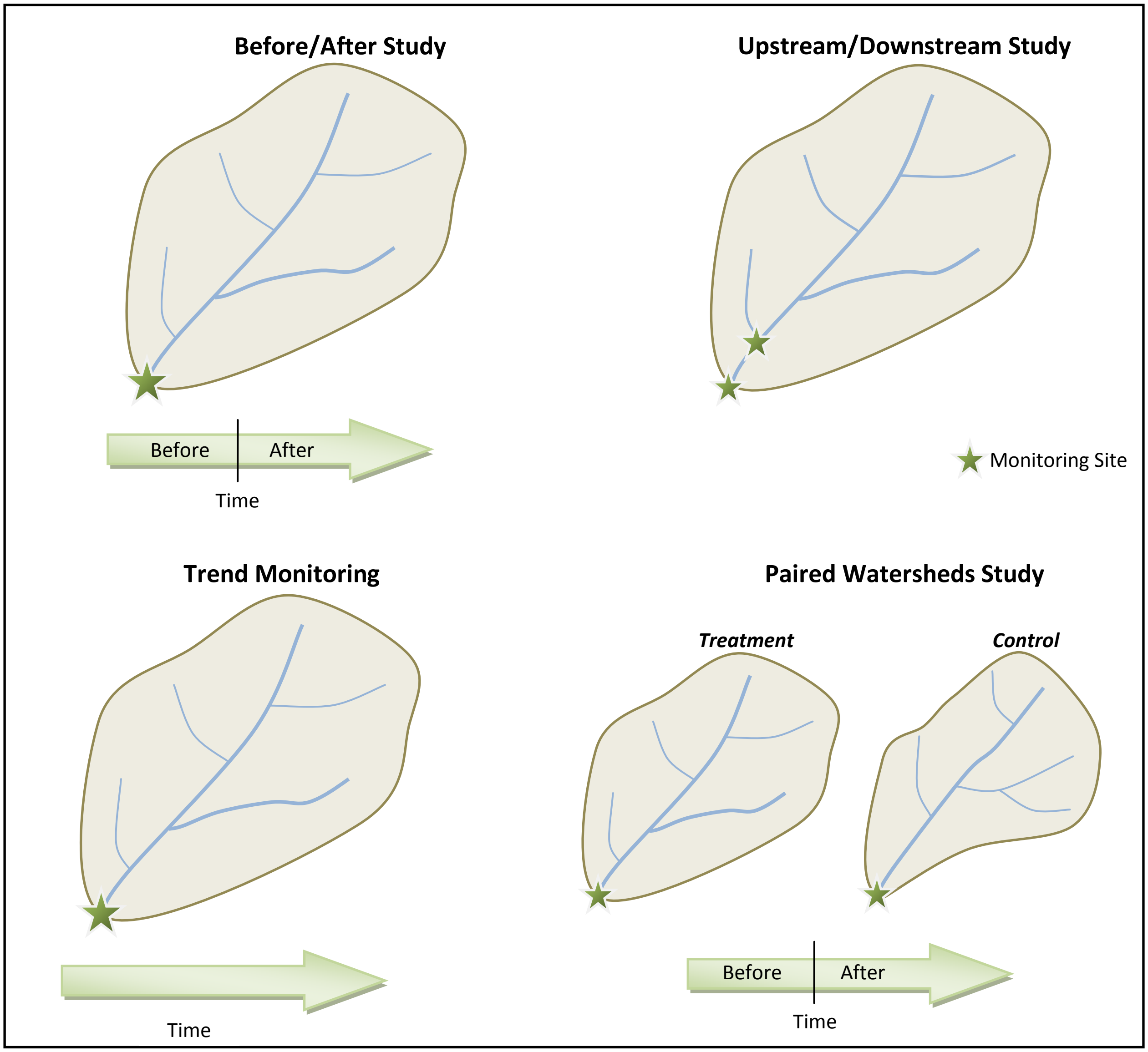


Figure 3. Study designs used for demonstrating TMDL effectiveness. Note that these diagrams reflect monitoring at the watershed outlet only (pour-point monitoring). If distributed sampling is undertaken, different study designs can be employed at each monitoring location.

Before/After Study

In a before/after design, a single monitoring site is established and water quality is sampled before and after TMDL implementation. Following data collection, data from each time period is compared. This design requires minimal data collection effort due to the use of a single site and can take advantage of existing water quality data. A before/after study cannot be employed if TMDL effectiveness monitoring begins after TMDL implementation has begun and if historical data are not available. Further, existing datasets should be reviewed to determine the number of observations, the date/time of sampling, and sampling frequency. These characteristics are especially important for water quality parameters that fluctuate over daily or seasonal scales. Annual variability in climate and hydrologic conditions between the two time periods can also interfere with the identification of water quality improvements if they are not accounted for. Local stream flow data can greatly enhance the power of this type of design. The two sample t-test or rank sum test are common statistical techniques for evaluating data from this type of monitoring design (see *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*).

Upstream/Downstream Study

The upstream/downstream study uses monitoring sites located upstream and downstream of TMDL implementation activities. Water quality monitoring is conducted during/after TMDL implementation and data from each location are compared. This study design is more resource intensive than the before/after study, but the variability of climate and hydrologic conditions over time is better accounted for.

The upstream/downstream study is one type of control/impact study, where control (upstream) and impact (downstream) sites are compared. In such studies, careful attention must be paid to the selection of the control site to ensure that water quality data from each location are comparable. For example, the upstream/downstream study is generally not appropriate where a tributary enters the reach of interest between potential upstream and downstream locations.. The two-sample t-test, paired t-test, or signed rank test are common statistical techniques for evaluating data from this type of monitoring design (see *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*).

Paired Watersheds Study

The paired watersheds study combines the before/after and control/impact study designs and is a specific type of before-after-control-impact (BACI) study. In the paired watersheds study, two monitoring sites are established, one located at the outlet of a treatment watershed and one located at the outlet of a control watershed where no management activities take place. The control site can be within an adjacent or nearby watershed, or upstream of the treatment area. Water quality is sampled at both sites before and after TMDL implementation. Data from the pre-implementation period is used to develop a statistical relationship between sites and this relationship is used to evaluate post-implementation water quality data.

The paired watersheds study is a highly rigorous study design that accounts for the variability of water quality in time and space. As such, it requires a major investment in resources. Additionally, watershed conditions in the control watershed must remain stable throughout the duration of the study, or the results will be compromised. Therefore, it may not be the most practical option. The paired t-test, signed rank test, or analysis of covariance (ANCOVA) are common statistical techniques for evaluating data from this type of monitoring design (see *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*).

Trend Monitoring

Trend monitoring typically uses water quality data from a single monitoring station to demonstrate TMDL effectiveness. Water quality data are collected over an extended time period (at least 5 – 10 years) and changes over time are assessed. The main advantage of trend monitoring versus a before/after study is the ability to evaluate TMDL implementation that occurs over an extended time period using regularly collected data. This is important for documenting incremental improvements, as full restoration of an impaired water body typically takes many years or decades. Further, trend monitoring in a watershed with multiple monitoring sites can be used to estimate the "regional" trend or overall trend for the watershed. Collection of streamflow and other covariate data (e.g., turbidity, temperature) is particularly valuable under this study design due to its ability to improve the statistical analyses performed on the resulting data. Linear regression and the Mann-Kendall test are common statistical techniques for evaluating data from this type of monitoring design (see *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*). More advanced methods are required for evaluating regional trends for multiple monitoring sites within a watershed.

5. Estimate Sample Size

Water quality data are often collected without considering the number of samples needed to demonstrate statistically significant changes. Statistically significant results are important for communicating improvements to decision makers and stakeholders with a specified level of confidence. They are also important for demonstrating improvement under Measure SP-12. Objective and informed sample size decisions can be made using a statistical method known as ***Power Analysis***. A power analysis uses information from pilot data to determine the optimal number of samples needed to identify statistically significant changes or trends. This ensures that sufficient data are available to identify future changes or trends where they actually exist, without sampling more than necessary, bringing improved efficiency to the monitoring program.

The power of a statistical test is defined as the likelihood of detecting a change that actually occurred and it depends on the magnitude of the change to be detected by the test (the effect size), the number of samples included in the test, the level of significance (the likelihood that the water quality change identified by the test actually occurred), and the variability of the parameter being tested. Power increases with sample size, effect size, and/or significance level, and decreases with sample variability.

Since the sample variability and effect size can be estimated from pilot data or set at minimum threshold levels, the sample size needed to achieve a pre-defined level of power and significance can be calculated. For TMDL effectiveness monitoring, planners can take advantage of this type of power analysis to estimate sample size requirements and plan for associated sampling costs.

Sample size calculations can be complex and vary depending on the type of statistical test that will be used to analyze the monitoring data. Therefore, a Microsoft Excel-based TMDL effectiveness monitoring planning tool that facilitates sample size estimation is available as a complement to this document. Note that sample size analysis that is carried out before a study is conducted (an *a priori* analysis) requires certain assumptions regarding the parameter variability, distribution, and the statistical test that will be applied. Such assumptions may not be valid for the final dataset or statistical approach. For example, the inclusion of covariates in trend analysis or presence of autocorrelation in sample data can require statistical methods that are more sophisticated than originally planned. For this reason, *a priori* sample size estimates may be biased high or low. Methods for exploring these assumptions are discussed in *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*, and several of these are featured in the Excel-based effectiveness monitoring planning tool.

6. Develop TMDL Effectiveness Monitoring Plan

The preceding sections have described recommended steps for planning a TMDL effectiveness monitoring project. These steps include:

- Review Existing Data and Information;
- Site Selection;
- Parameter Selection;
- Study Design Selection; and
- Sample Size Estimation.

The results of each step should be documented in a TMDL effectiveness monitoring plan. The planning document should include relevant background information and clearly spell out the goals of the project, where and when monitoring will occur, and what will be monitored. A preliminary discussion of intended data analysis methods should also be presented, including the selected level of significance. A TMDL effectiveness monitoring plan can be incorporated into a Quality Assurance Project Plan (QAPP). The QAPP should outline additional details related to field sampling and laboratory protocols, data management, and the proposed process for addressing data quality issues that arise during the course of the project. By spending adequate time on the development of a TMDL effectiveness monitoring plan, and following the plan's direction, those tasked with designing and managing a TMDL effectiveness monitoring program will maximize its ecological, economic, and social value.

7. References

Rumps, J., Katz, S., Barnas, K., Morehead, M., Jenkinson, R., Clayton, S., et al. (2007). Stream Restoration in the Pacific Northwest: Analysis of Interviews with Project Managers. *Restoration Ecology* , Vol. 15, No. 3.

U.S. Environmental Protection Agency. (2002). *Guidance on Choosing a Sampling Design for Environmental Data Collection EPA QA/G-5S*. Available online at: http://www.epa.gov/quality/qs-docs/g5s-final.pdf.

U.S. Environmental Protection Agency. (2008). *Handbook for developing watershed TMDLs.* Available online at: www.epa.gov/owow/tmdl/pdf/draft_handbook.pdf.

This page is purposely left blank

Appendix E. Technical Guidance for Exploring TMDL Effectiveness Monitoring Data

This page is purposely left blank

Technical Guidance for Exploring TMDL Effectiveness Monitoring Data

1. Introduction

Effectiveness monitoring is a critical step in the Total Maximum Daily Load (TMDL) process for addressing water quality impairments in lakes, streams, and other water bodies. Since the overall goal of TMDL effectiveness monitoring is to identify water quality improvements that result from TMDL implementation, high quality data analysis is needed throughout all project phases. During project planning, water resource practitioners will find themselves making decisions that can be informed through statistical analysis of pilot data using ***Exploratory Data Analysis*** (EDA) techniques. EDA provides insight into the basic characteristics of individual datasets, as well as general relationships between two or more data groups.

EDA methods used to address project planning questions are also relevant when conducting analysis for interim and final reporting using ***Confirmatory Data Analysis*** (CDA) techniques. CDA is generally applied to confirm or reject assumed relationships between two or more data groups with a quantifiable level of confidence. A number of statistical tests are available for CDA. Although the selection of an appropriate test is largely based on the study design and objectives, characteristics of the data at hand must also be considered. Failure to do so can result in violation of basic assumptions of the statistical test and lead to conclusions that are, at best, poorly justified and, at worst, invalid. EDA techniques provide a straightforward means to evaluate such assumptions.

The primary purpose of this guide is to direct readers to EDA techniques that influence: 1) the selection of water quality parameters to include as part of TMDL effectiveness monitoring; 2) sampling frequency, timing, and duration; and 3) statistical test selection. Several EDA methods are demonstrated using stream monitoring data collected by the Ambient River and Stream Monitoring Program of the Washington Department of Ecology and screen captures from Microsoft Office Excel 2007 spreadsheet software. Specific topics covered in this guide include:

- Use of histograms to investigate data distributions and identify outliers;
- Transforming data with a skewed distribution;
- The use of time series plots to identify missing data and daily/seasonal fluctuations;
- Use of boxplots to further explore daily/seasonal fluctuations;
- Use of scatterplots to identify covariates;
- Use of lag plots to explore autocorrelation; and
- Analysis of censored data.

This guide offers just a small glimpse into exploratory and confirmatory statistical analysis methods. Readers interested in learning more about these topics are encouraged to consult the references listed at the end of this document. Additional information on TMDL effectiveness monitoring can also be found in a companion document titled *Technical Guidance for Designing a TMDL Effectiveness Monitoring Plan*.

2. Data Exploration

In any study involving data analysis, the integrity of reported results is directly tied to how thoroughly the basic characteristics of study data are explored and understood. For a TMDL effectiveness monitoring project, exploration of monitoring data that are collected following project implementation is essential. Further, exploration of data used to inform planning decisions (pilot data) can greatly enhance the informative value of the project. Pilot data can include data collected as part of a pilot water quality study, data collected from study sites under other water quality monitoring programs, or existing data collected from comparable sites.

Data exploration answers a number of questions that are relevant to the selection of water quality parameters for TMDL effectiveness monitoring; sampling frequency, timing, and duration; and statistical test selection. These questions can be answered by examining individual data points and by developing a few, readily producible graphs.

Data Distributions & Histograms

Consider any water quality parameter, such as a sample site's total phosphorous concentration. Over time, the concentration of total phosphorous will vary between a minimum and maximum value. Some values will occur regularly while others only in rare instances. If samples were collected continuously at all points in time (i.e., the entire population were known), a complete assessment of the ***probability distribution*** of total phosphorous could be performed. A parameter's probability distribution describes the range of values that can occur and the relative likelihood that they will occur. Since it is not feasible to measure each and every occurrence of total phosphorous, the probability distribution of total phosphorous (or any other water quality parameter) must be inferred from the distribution of a limited number of data points, the ***sampling distribution***. The sampling distribution can be thought of as an estimate of the otherwise-unknown probability distribution.

The sampling distribution of a variable can be displayed graphically using a ***histogram***. A histogram consists of individual data values, or ranges of data values (called bins), on the x-axis. The y-axis contains the number of times each data (or bin) value appears in the dataset (the frequency of the value). Step-by-step instructions for generating a histogram in Microsoft Excel 2007 are provided on page 5.

When viewing a histogram, a principal question to ask is "Do the data display the ***normal distribution***?" The normal distribution is a specific type of probability distribution, with values clustered symmetrically around a common value (the mean) (Figure 1). The normal distribution serves as the basis for ***parametric statistics***. A fundamental assumption of parametric statistical methods is that data are normally distributed, and the accuracy of these methods is greatly diminished if they are applied to data that are not normally distributed. This is a key point for TMDL effectiveness monitoring since a number of parametric statistical tests are available to evaluate monitoring data. These tests frame study questions in the form of a prior assumption, or hypothesis, about the data. As such, they are known as ***hypothesis tests.*** Alternatives to parametric tests exist in the form of ***nonparametric tests***. Nonparametric tests include no assumption of a normal (or any other) distribution (i.e., they are robust). When applied to normally distributed data however, nonparametric tests always have less power than an equivalent parametric test.

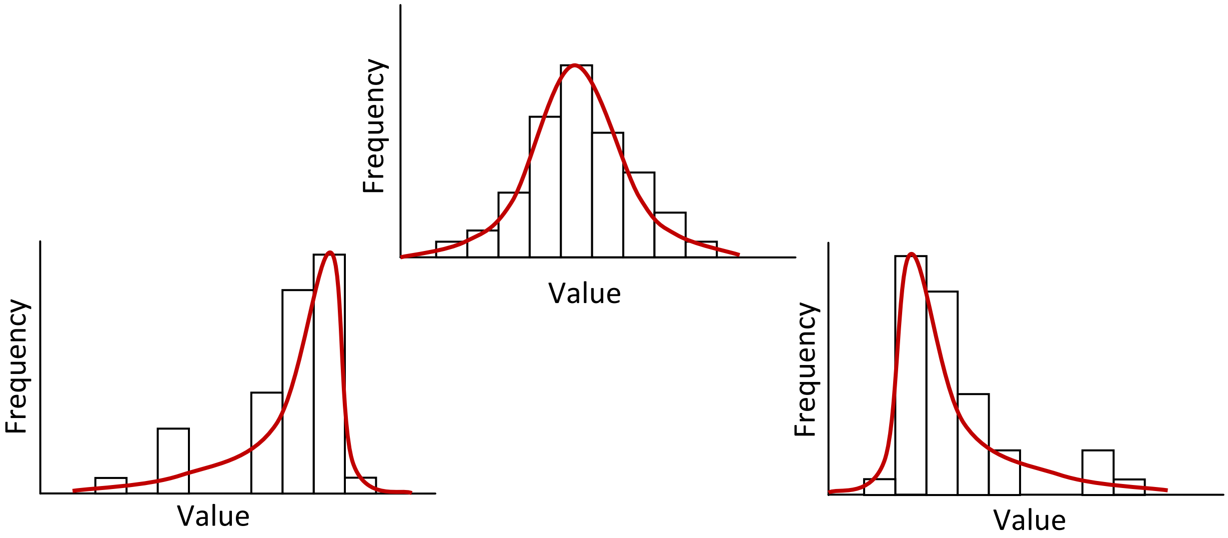
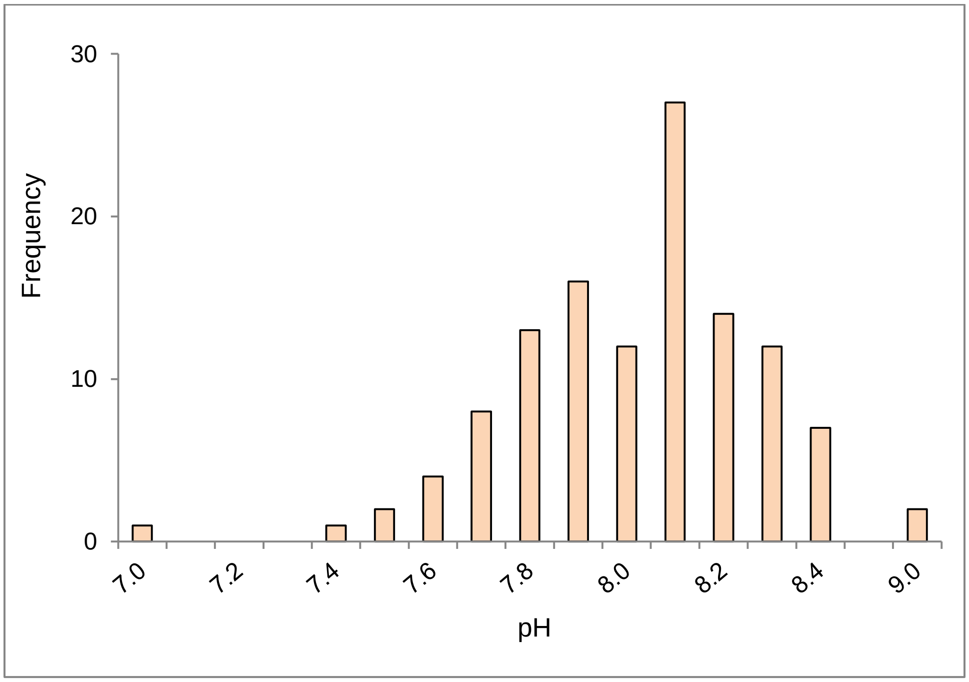


Figure 1. Example histograms demonstrating data with the normal distribution (top), left-skewed distribution (bottom left), and right-skewed distribution (bottom right).

Figure 2 illustrates histograms for two sample datasets from the Spokane River, WA. The pH histogram conveys an approximate normal distribution, while the phosphorous histogram appears right-skewed. Keep in mind that the classic, bell-shape of a histogram of normally-distributed data will rarely be observed. This should not lead one to reject the prospect of a normal distribution. Rather, characteristics such as sample size, the degree of skewness, and the parameter considered should be evaluated.



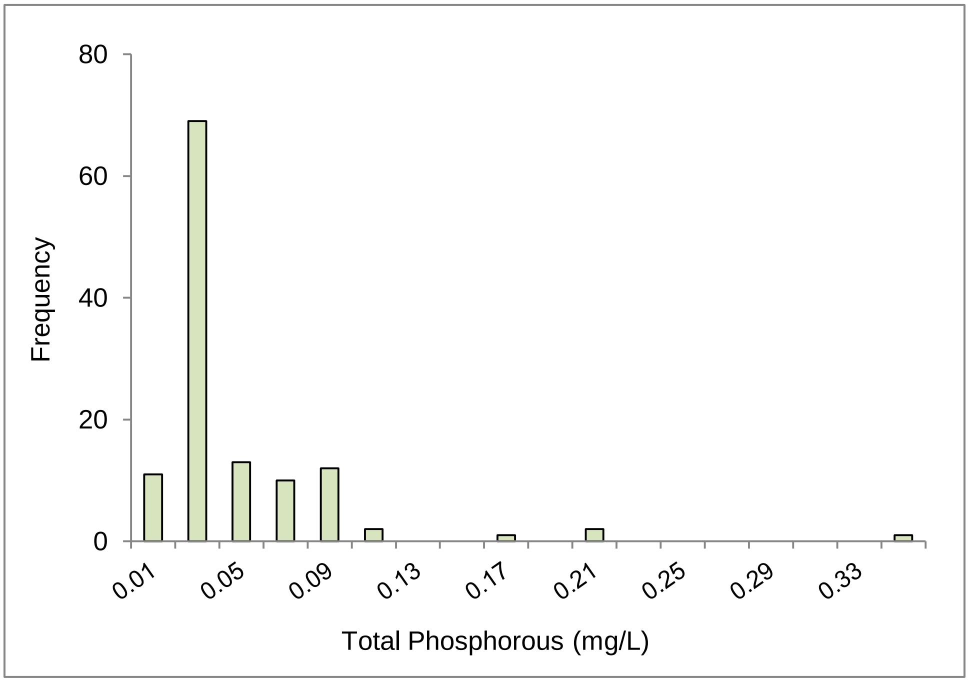
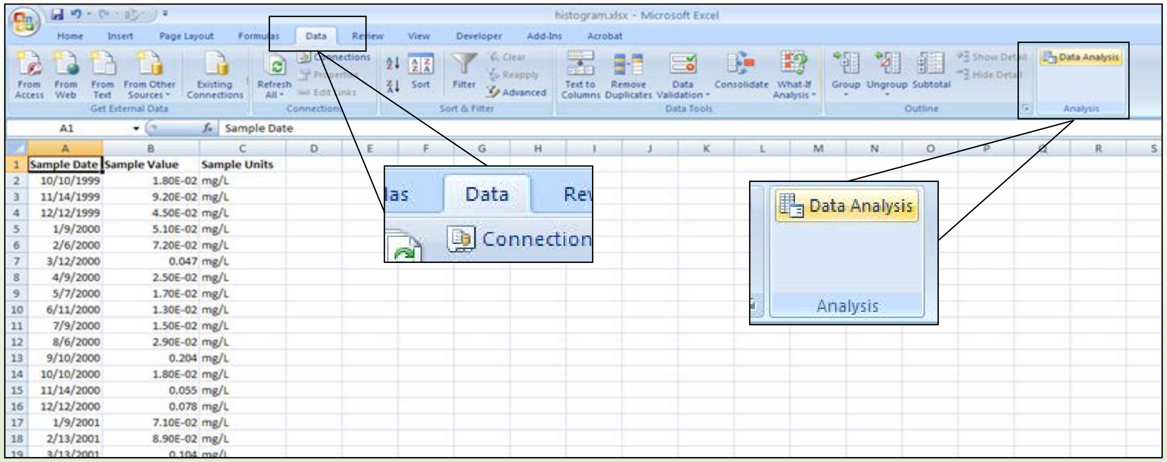


Figure 2. Histograms of Spokane River (at Riverside State Park) pH (top) and total phosphorous concentration (bottom) data collected between 10/1999 and 9/2009. Note differences in the degree of skewness (symmetry) for each parameter.

EDA in Excel. Creating a Histogram

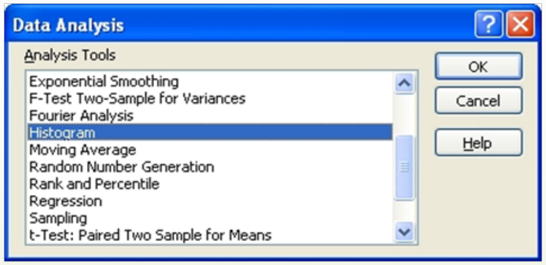
In Excel 2007, use the ***Data Analysis*** tool to create a histogram. In the example below, a histogram is created for Spokane River (at Riverside State Park) total phosphorous concentration using monthly samples collected over the period of October 1999 through September 2009.

Step 1. Arrange the data so that one column contains all measured values (see below example). Click the ***Data*** tab. In the Analysis section, click ***Data Analysis***. If the Data Analysis button is not displayed, see Appendix A for instructions on installing the Analysis Toolpak.

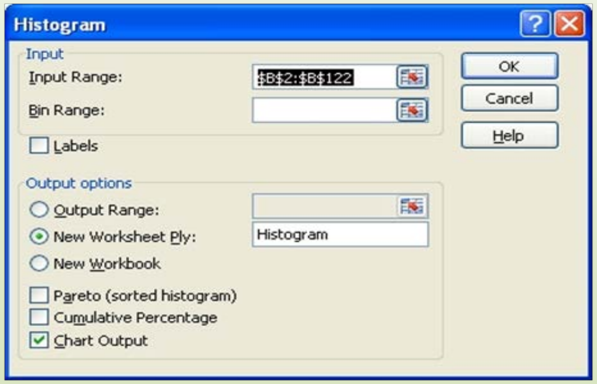


	A	B	C
1	Sample Date	Sample Value	Sample Units
2	10/10/1999	1.80E-02	mg/L
3	11/14/1999	9.20E-02	mg/L
4	12/12/1999	4.50E-02	mg/L
5	1/9/2000	5.10E-02	mg/L
6	2/6/2000	7.20E-02	mg/L
7	3/12/2000	0.047	mg/L
8	4/9/2000	2.50E-02	mg/L
9	5/7/2000	1.70E-02	mg/L
10	6/11/2000	1.30E-02	mg/L
11	7/9/2000	1.50E-02	mg/L
12	8/6/2000	2.90E-02	mg/L
13	9/10/2000	0.204	mg/L
14	10/10/2000	1.80E-02	mg/L
15	11/14/2000	0.055	mg/L
16	12/12/2000	0.078	mg/L
17	1/9/2001	7.10E-02	mg/L
18	2/13/2001	8.90E-02	mg/L
19	3/13/2001	0.104	mg/L

Step 2. Select ***Histogram*** from the Data Analysis dialog box.



Step 3. In the histogram interface, specify the Input & Output Range. Make sure the ***Chart Output*** option is checked.



Input Range: $B$2:$B$122

Bin Range:

☐ Labels

○ Output Range:

● New Worksheet Ply: Histogram

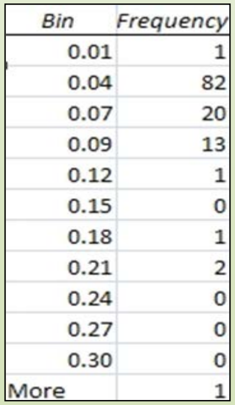
○ New Workbook

☐ Pareto (sorted histogram)

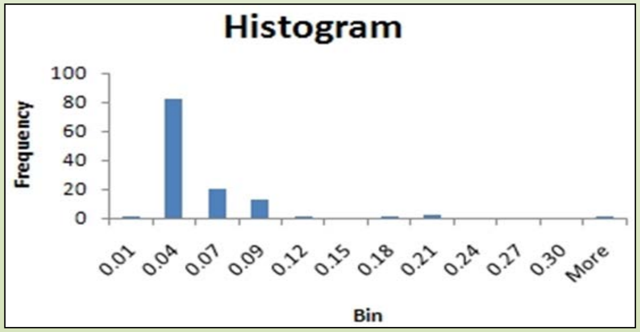
☐ Cumulative Percentage

☑ Chart Output

Histogram output includes bin ranges and frequency data in table and graph form. Bin ranges can be edited and input to the histogram interface by repeating steps 1 through 3.



Bin	Frequency
0.01	1
0.04	82
0.07	20
0.09	13
0.12	1
0.15	0
0.18	1
0.21	2
0.24	0
0.27	0
0.30	0
More	1



Data Transformation

If a sampling distribution is heavily skewed, a ***data transformation*** can often be applied to produce symmetrical data that better fit the normality assumption of parametric statistical tests. In theory, any transformation can be used as long as it is applied to all data points and symmetry is achieved. Log or power transformations are the most common transformation types. These transformations involve calculation of the logarithm of each data point or the value of each data point raised to some power (such as the square or square-root of each data point).

Figure 3 contains histograms of untransformed and log-transformed total phosphorous data collected from the Spokane River (at Riverside State Park). The untransformed data are right-skewed. Transformation using the logarithm provides a dataset that is approximately normally distributed. These data can subsequently be used for parametric statistical testing.

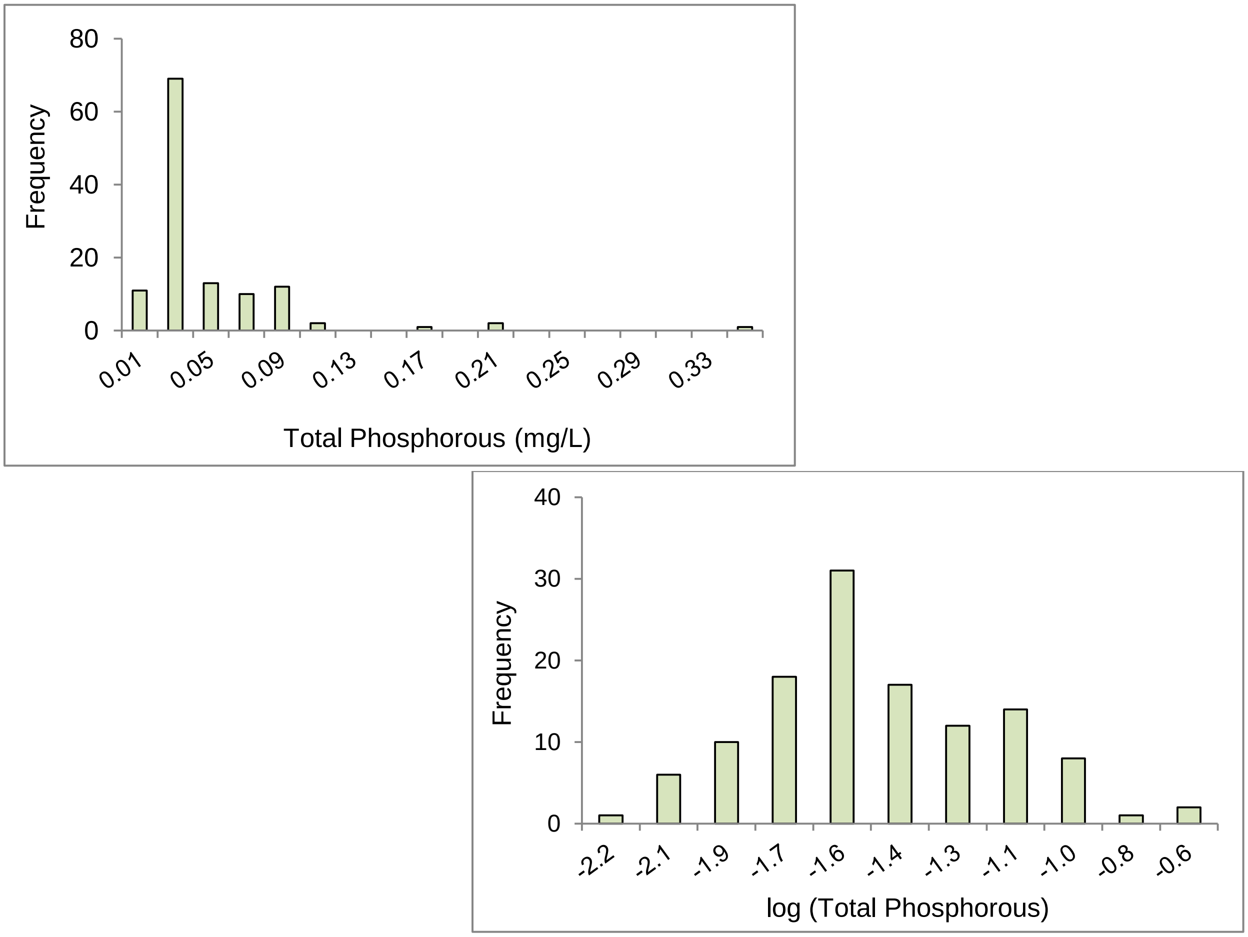


Figure 3. Histograms of total phosphorous concentration (top) and log-transformed total phosphorous concentration (bottom) data for samples collected from the Spokane River (at

Riverside State Park) between 10/1999 and 9/2009. Note that log-transformed data appear normally distributed.

When dealing with skewed data, one should not become preoccupied with finding the "right transformation." Start with the common transformations (log, square, square-root, etc.) and compare histograms. If one of these does not produce symmetrical data, consult sources specific to the water quality parameter being analyzed. Keep in mind that transformations can provide values of zero and negative numbers and that some transformations cannot be applied if a dataset contains one or more zero values. Remember that if two groups of data are compared (e.g., before and after TMDL implementation), the same transformation must be applied to each group. A comparison of pre-TMDL log-transformed total phosphorous and post-TMDL untransformed total phosphorous, for example, would not be appropriate. Also, results generated using transformed data are often transformed back to original units for reporting (back-transformation). Readers are advised to use caution when back-transforming results to prevent calculation and reporting errors.

A final note to consider is that even though a transformation may provide data that meet the normality assumption of parametric statistical tests, other assumptions specific to the test(s) applied should also be explored. Refer to Helsel & Hirsch (2006) and U.S. Environmental Protection Agency (2006) for a detailed discussion of the assumptions of common statistical tests.

Outliers

It is common for a dataset to contain one or more values that are far removed from the remaining observations. These values, called *outliers*, can arise from a variety of sources. An outlier may simply reflect the occurrence of a rare event or be the product of a measurement or data entry error. Histograms are useful for identifying outliers (e.g., see the top histogram in

Figure 3, which shows an outlier of 0.35 mg/L).

Outliers should not be removed or deleted from a dataset (unless they are known to be measurement/recording errors). Instead, the presence of outliers should be noted, as these values can influence the selection of a parametric vs. nonparametric statistical test. This is due to the use of the mean value of a dataset for parametric tests and bias in the mean that is caused by outliers. If one or more outliers are identified, data transformation options should be explored to determine if an outlier-free dataset can be produced. If transformation does not produce symmetry in the data (or if transformation is not desirable), nonparametric methods should be used for subsequent analysis. Should a parametric test be performed on a dataset that includes outliers, practitioners can evaluate the influence of the outliers by performing the test twice, once using the full dataset (including the outliers) and again on the reduced dataset (excluding the outliers). If the results are different from one another, a nonparametric test should be used.

Time Series & Missing Data Points

Water quality monitoring often involves the collection of samples at regular intervals over time (hourly, weekly, monthly, etc.). Plotting these data in sequence (in the order they were collected) can display important information, including shifts in the central tendency (mean or median) or variability over time, and the existence of large gaps in coverage over time. These plots are known as ***time series*** plots. Time series plots contain the sample date on the x-axis and sample value on the y-axis. Instructions for producing time series plots in Microsoft Excel 2007 are provided on page 10.

A time series plot of fecal coliform concentration in Burnt Bridge Creek, WA over a six-year period is provided in

Figure 4. Readers may notice one of several data characteristics conveyed in this plot. These include the presence of outliers, seasonality of measured values (lower values during the winter months), and the large gap in data points from October 2005 to October 2007.

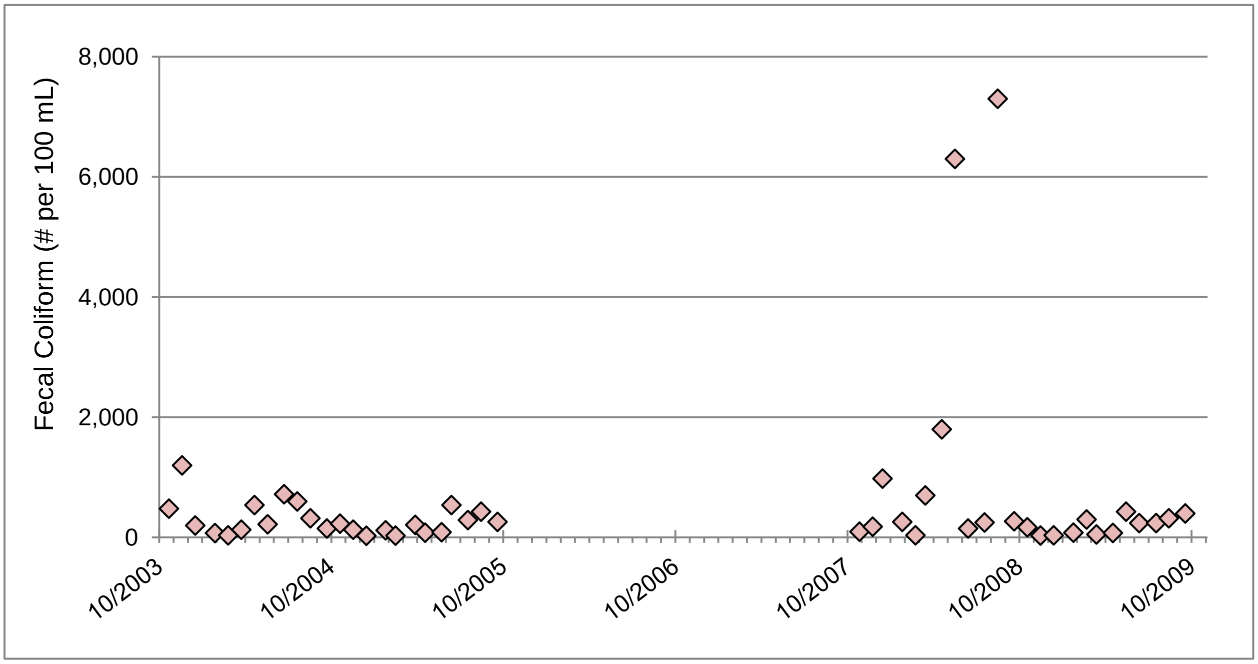


Figure 4. Time series plot for monthly fecal coliform concentration in Burnt Bridge Creek, WA. The plot displays several important pieces of information, including the data gap in 2005-06, the presence of 2 outliers in May and August of 2008, and seasonality in measured fecal coliform concentrations (lower values during the winter months).

The large data gap apparent in

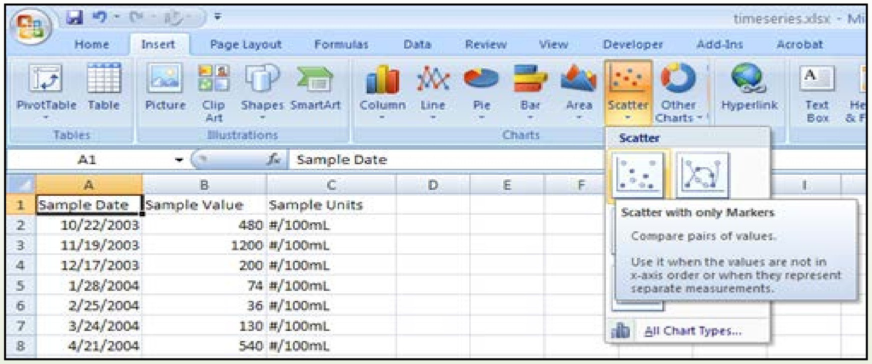
Figure 4 can be problematic if these data are included in a trend monitoring study. Trend monitoring data are used to assess changes in a parameter over time (are the values increasing or decreasing with time?). A complete data record is ideal for such analysis. However, data gaps often occur, and researchers must decide whether the data gap will affect trend analysis results. An objective method for evaluating the significance of data gaps, proposed by Helsel & Hirsch (2002), includes 3 steps:

1. Divide the entire study period into three separate periods of equal length.
2. Calculate the percent coverage in each period (the ratio of actual observations to potential observations, as a percentage).
3. Discard the data if percent coverage is less than 20% in any of the three periods.

EDA in Excel. Creating Time Series Plots

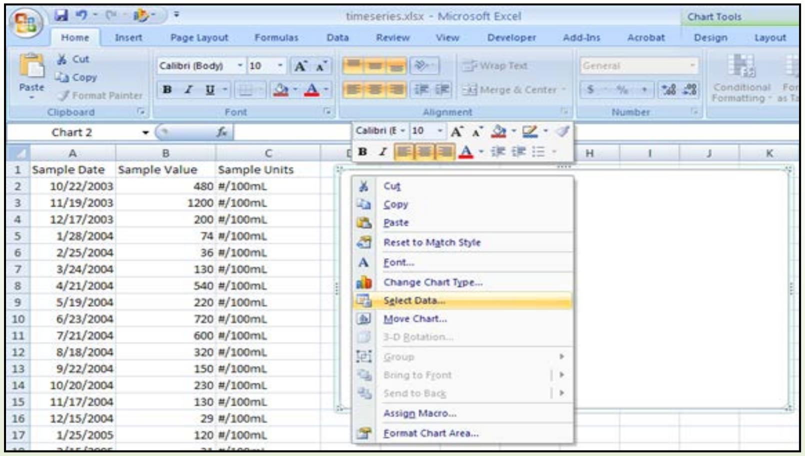
In Excel 2007, use the ***Scatter Chart*** tool to create a time series plot. In the example below, a time series plot is created for Burnt Bridge Creek fecal coliform concentration using monthly sample data collected over the period October 2003 through September 2009.

Step 1. Arrange the data so that one column contains all measured values and one column contains ordered sample dates (see below example). Click the ***Insert*** tab. In the ***Charts*** section, click ***Scatter***. Select ***Scatter with only Markers***.



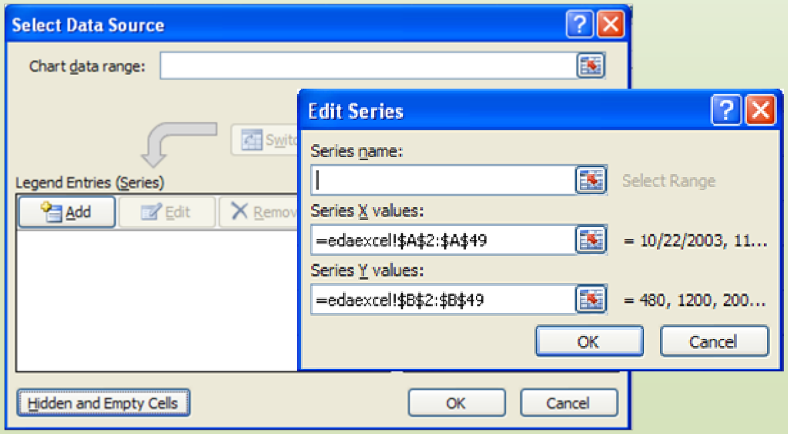
	A	B	C
1	Sample Date	Sample Value	Sample Units
2	10/22/2003	480	#/100mL
3	11/19/2003	1200	#/100mL
4	12/17/2003	200	#/100mL
5	1/28/2004	74	#/100mL
6	2/25/2004	36	#/100mL
7	3/24/2004	130	#/100mL
8	4/21/2004	540	#/100mL

Step 2. A blank chart is created. Right-click on the chart area and choose ***Select Data***.



	A	B	C
1	Sample Date	Sample Value	Sample Units
2	10/22/2003	480	#/100mL
3	11/19/2003	1200	#/100mL
4	12/17/2003	200	#/100mL
5	1/28/2004	74	#/100mL
6	2/25/2004	36	#/100mL
7	3/24/2004	130	#/100mL
8	4/21/2004	540	#/100mL
9	5/19/2004	220	#/100mL
10	6/23/2004	720	#/100mL
11	7/21/2004	600	#/100mL
12	8/18/2004	320	#/100mL
13	9/22/2004	150	#/100mL
14	10/20/2004	230	#/100mL
15	11/17/2004	130	#/100mL
16	12/15/2004	29	#/100mL
17	1/25/2005	120	#/100mL

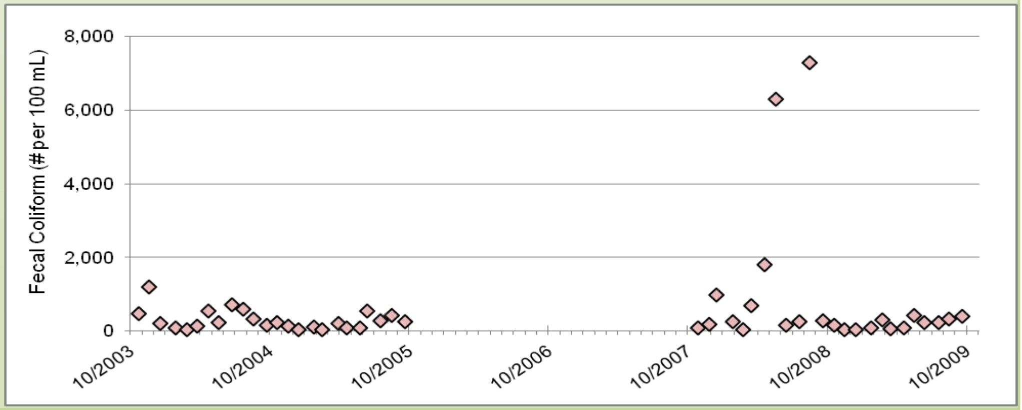
Step 3. In the ***Select Data Source*** dialog box, click ***Add.*** In the Edit Series dialog box, enter the range of cells containing sample dates below ***Series X values***. Enter the range of cells containing sample values below ***Series Y values***. Click *OK*.



Series X values: =edaexcel!$A$2:$A$49 = 10/22/2003, 11...

Series Y values: =edaexcel!$B$2:$B$49 = 480, 1200, 200...

Step 4. Edit axis limits, titles, and data markers as needed.



Daily/Seasonal Fluctuation and Box Plots

Just as weather data (air temperature, rainfall, etc.) regularly display distinct daily and seasonal patterns, water quality parameters often fluctuate over daily and/or seasonal cycles. Daily, or ***diel***, fluctuation refers to regular, cyclic variation in the distribution of a parameter over a 24-hour period. ***Seasonality*** refers to the phenomenon of regular, cyclic variation over a one-year period. Diel and seasonal fluctuations can be detected from time series plots and are important for analysis of TMDL effectiveness monitoring data. For example, a comparison of pre- and post-TMDL data for a parameter showing strong seasonality is of little use if pre-TMDL data consist of observations from the month of December and post-TMDL data consist of July observations only. Similarly, changes in a parameter with large diel fluctuation over time can be clouded if continuous data are not available and individual grab samples are not collected at the same time of day.

In addition to time series plots, cyclic fluctuation can be explored using ***box plots***. A box plot conveys five key pieces of information on the distribution of a parameter: the minimum, the maximum, the median (50th percentile), the 25th percentile, and the 75th percentile. A box plot contains summary statistic values on the y-axis and parameter groups on the x-axis. For each parameter group, the five summary statistics are displayed using the following format:

- A box with lower and upper limits at the 25th and 75th percentile values, respectively.
- A horizontal line or point within the box representing the median value.
- Vertical lines (often called whiskers) extending outside of the box to the minimum and maximum observed values.

A daily or seasonal boxplot can be constructed to compare summary statistics between seasons or periods of the day (day vs. night). An example is shown in

Figure 5 and instructions for preparing boxplots in Microsoft Excel 2007 are provided on page 13.

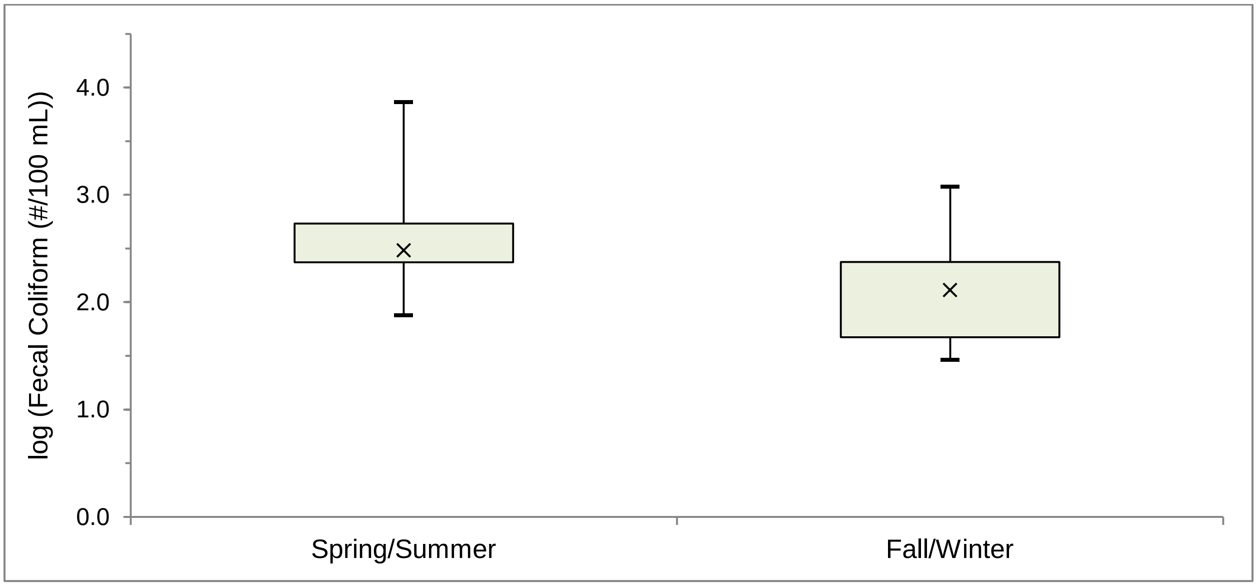


Figure 5. Box plot comparing spring/summer and fall/winter fecal coliform concentration in Burnt Bridge Creek, WA over the period 10/2003 through 9/2009. The box plot provides a graphical display of five key summary statistics: the median (the X symbol); 25th and 75th percentile values (the lower and upper sides of the box); and minimum and maximum observed values (the lines outside the box). Here, the box plot points to seasonal differences in the distribution of fecal coliform concentrations.

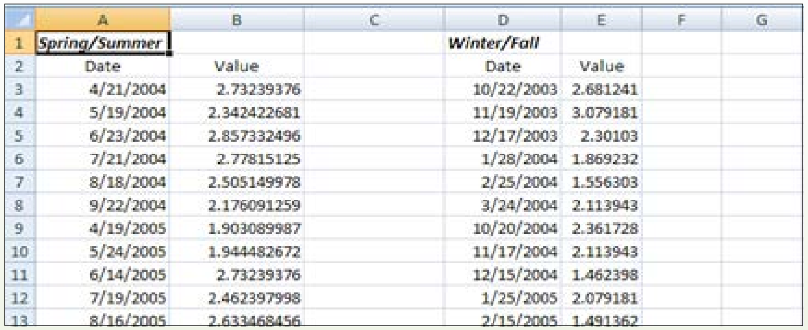
Knowledge of diel and seasonal fluctuation in parameters included in TMDL effectiveness monitoring should be integrated into project planning and analysis decisions. For example, it may be desirable to sample a parameter with large seasonal fluctuation only during the portion of the year when it is known to be at risk for exceeding water quality standards (e.g., summer bacteria counts). Alternatively, several statistical methods that take seasonal variability into account are available for analysis of TMDL effectiveness monitoring data. For example, the Seasonal Kendall test, a modified version of the nonparametric Kendall test for trend, is a common method for including the effect of seasonal variability in trend analysis. Detailed information on the Seasonal Kendall test, and other statistical methods that account for seasonality effects, can be found in Helsel & Hirsch (2002) and Helsel et al. (2006).

Parameters with a strong diel pattern may require continuous water quality monitoring, in which measurements are collected and recorded several times throughout the day. These data can be used to generate daily statistics (mean, maximum, minimum, etc.) for use in further statistical analysis. If continuous monitoring is not a viable option, parameters with large diel fluctuation should be sampled at roughly the same time of day. Examples of parameters with potential diel cycles include dissolved oxygen, pH, water temperature, and streamflow. In some water bodies, the concentration of nutrients can also vary dramatically throughout the day.

EDA in Excel. Creating a Boxplot

Excel 2007 does not include a built-in box plot tool. Box plots instead must be created using the ***Line Chart Tool***. In the example below, a seasonal box plot is created for Burnt Bridge Creek fecal coliform concentration using monthly sample data collected over the period October 2003 through September 2009.

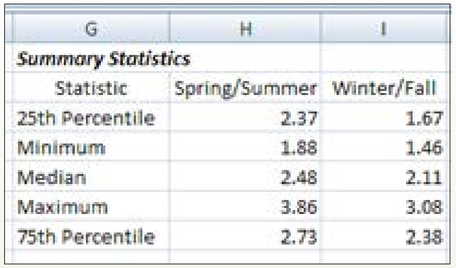
Step 1. Arrange the data so that one column contains all measured values and one column contains the sampling date for each seasonal group.



	A	B	C	D	E	F	G
1	*Spring/Summer*			*Winter/Fall*			
2	Date	Value		Date	Value		
3	4/21/2004	2.73239376		10/22/2003	2.681241		
4	5/19/2004	2.342422681		11/19/2003	3.079181		
5	6/23/2004	2.857332496		12/17/2003	2.30103		
6	7/21/2004	2.77815125		1/28/2004	1.869232		
7	8/18/2004	2.505149978		2/25/2004	1.556303		
8	9/22/2004	2.176091259		3/24/2004	2.113943		
9	4/19/2005	1.903089987		10/20/2004	2.361728		
10	5/24/2005	1.944482672		11/17/2004	2.113943		
11	6/14/2005	2.73239376		12/15/2004	1.462398		
12	7/19/2005	2.462397998		1/25/2005	2.079181		
13	8/16/2005	2.633468456		2/15/2005	1.491362		

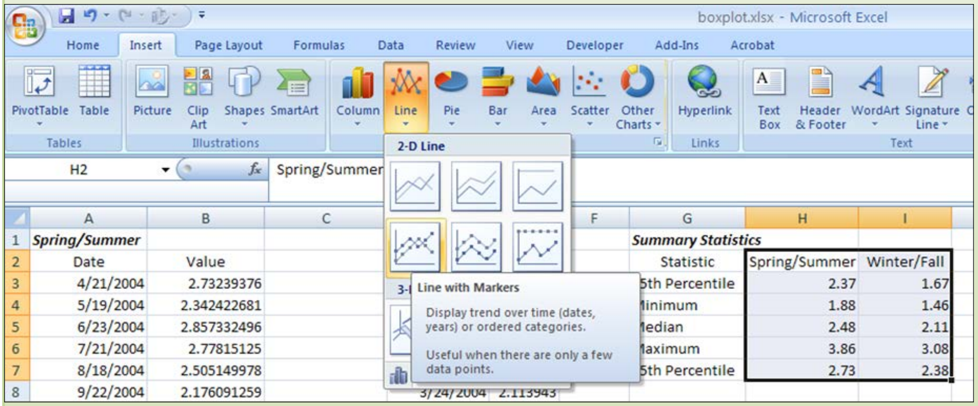
Step 2. Create a table of the five summary statistics for each seasonal group in the following order: 25th percentile, Minimum, Median, Maximum, 75th percentile. Formulas for calculating summary statistics are provided below:

- 25th percentile: =PERCENTILE(Data Range, 0.25)
- Minimum: =MINIMUM(Data Range)
- Median: =MEDIAN(Data Range)
- Maximum: =MAXIMUM(Data Range)
- 75th percentile: =PERCENTILE(Data Range,0.75)



G	H	I
Summary Statistics		
Statistic	Spring/Summer	Winter/Fall
25th Percentile	2.37	1.67
Minimum	1.88	1.46
Median	2.48	2.11
Maximum	3.86	3.08
75th Percentile	2.73	2.38

Step 3. Highlight summary data and click the ***Insert*** tab. In the ***Charts*** section, click ***Line*** and select ***Line with Markers***.



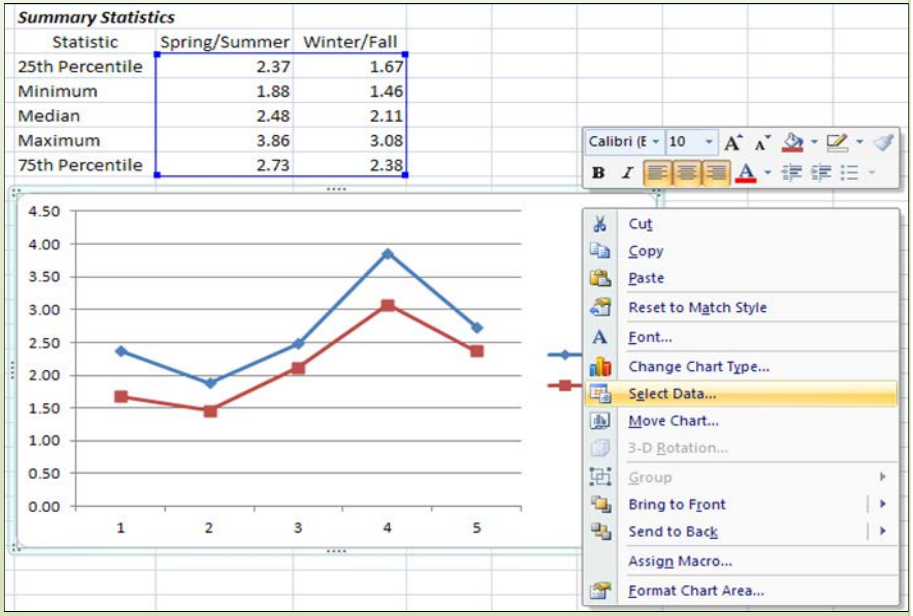
Line with Markers

Display trend over time (dates, years) or ordered categories.

Useful when there are only a few data points.

	A	B	C	F	G	H	I
1	*Spring/Summer*				*Summary Statistics*		
2	Date	Value			Statistic	Spring/Summer	Winter/Fall
3	4/21/2004	2.73239376			5th Percentile	2.37	1.67
4	5/19/2004	2.342422681			inimum	1.88	1.46
5	6/23/2004	2.857332496			edian	2.48	2.11
6	7/21/2004	2.77815125			aximum	3.86	3.08
7	8/18/2004	2.505149978			5th Percentile	2.73	2.38
8	9/22/2004	2.176091259					

Step 4. Right-click on the new chart and go to ***Select Data…***



Summary Statistics		
Statistic	Spring/Summer	Winter/Fall
25th Percentile	2.37	1.67
Minimum	1.88	1.46
Median	2.48	2.11
Maximum	3.86	3.08
75th Percentile	2.73	2.38

Continued on next page…

EDA in Excel. Creating a Boxplot (cont'd.)

Step 5. In the Select Data Source box, click *Switch Row/Column*. Click *OK*.

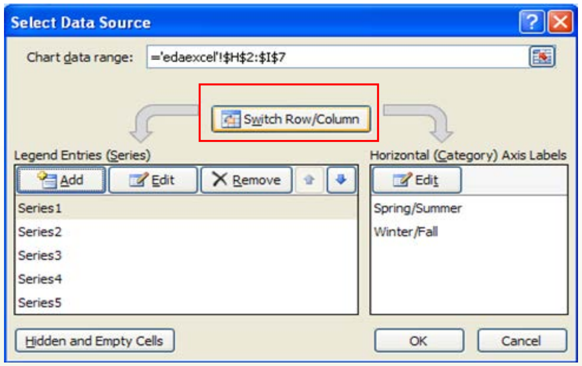
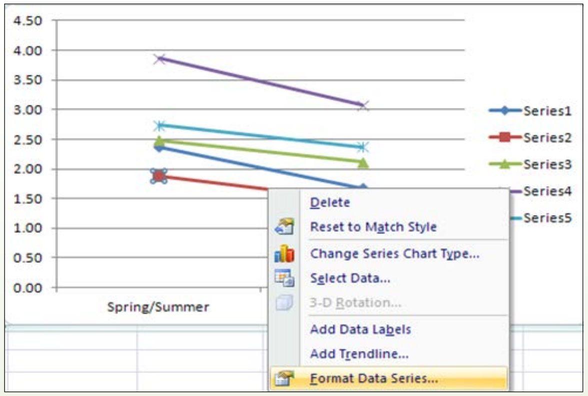
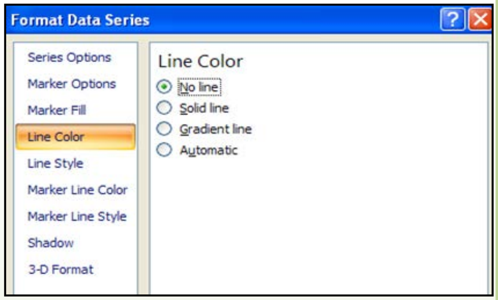


Chart data range: ='edaexcel'!$H$2:$I$7

Step 6. Right-click on a data series and select *Format Data Series…*. In the Format Data Series box, select the *Line Color* menu. Select *No Line* and click *Close*. Repeat for all 5 data series.





Line Color: ● No line

Step 7. Under *Chart Tools*, select the *Layout* tab. In the *Analysis* section, select *Lines*. Click *High-Low Lines*. Repeat for *Up/Down Bars*.

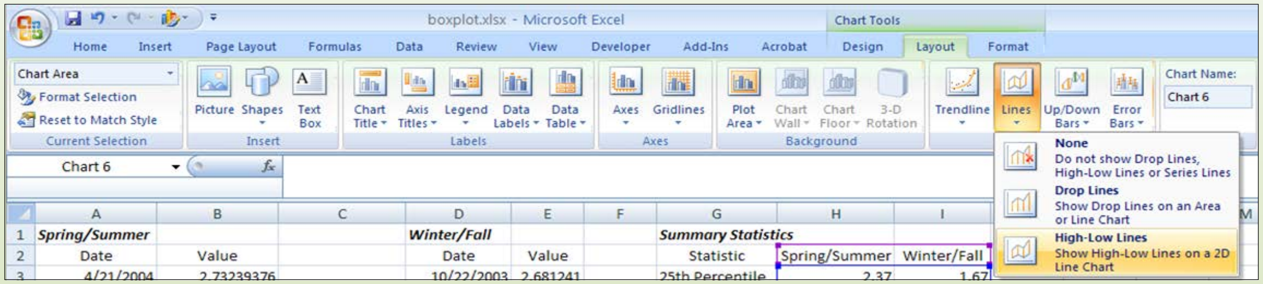


Chart Name: Chart 6

	A	B	C	D	E	F	G	H	I
1	Spring/Summer			Winter/Fall			Summary Statistics		
2	Date	Value		Date	Value		Statistic	Spring/Summer	Winter/Fall
3	4/21/2004	2.73239376		10/22/2003	2.681241		25th Percentile	2.37	1.67

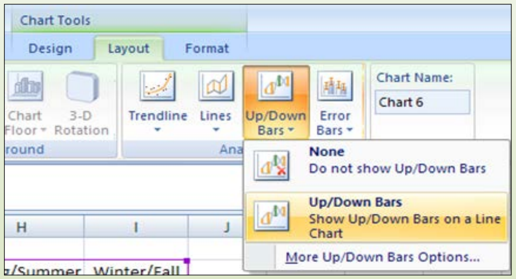
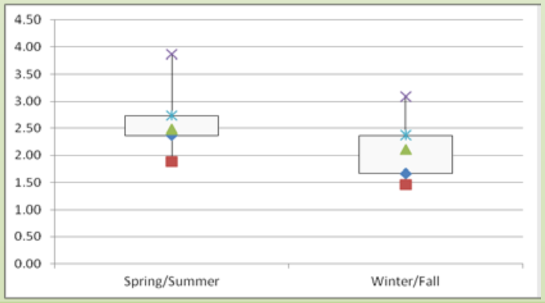
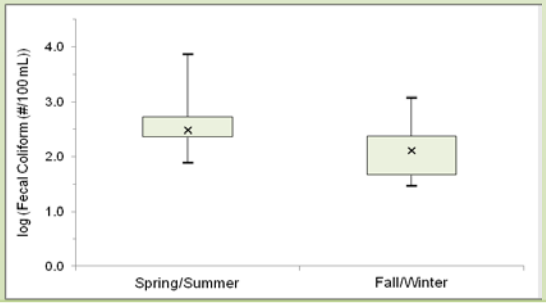


Chart Name: Chart 6

Your chart will now contain markers for each summary statistic, a vertical line connecting the minimum and maximum value, and a box connecting the 25th and 75th percentile values. Edit chart formatting as desired.



→



Autocorrelation and Lag Plots

Seasonal and daily cycling in time series data are two specific examples that reflect the statistical phenomenon known as ***autocorrelation***. Autocorrelation, or serial correlation, is defined as correlation between the elements of a series with other elements of the same series separated by some time interval. In other words, autocorrelation describes how similar (positive autocorrelation) or different (negative autocorrelation) observed values are with past or future values. For a parameter with strong seasonal fluctuation, observations separated by one year will exhibit positive autocorrelation, and observations separated by six months may demonstrate negative autocorrelation. Similarly, a parameter with large diel fluctuation will exhibit positive autocorrelation among observations separated by 24 hours and negative autocorrelation among observations separated by 12 hours.

When exploring autocorrelation, the time interval separating correlated observations is referred to as the time lag. A common tool for evaluating autocorrelation at a particular time lag is the *lag plot*. The lag plot is a collection of points, each representing a matched pair between the observed value at time, t, and the observed value at time, t minus lag. Random scatter in the lag plot indicates minimal autocorrelation at the selected lag time, while data with strong autocorrelation will produce a more structured lag plot. Example lag plots are shown in

Figure 6 and instructions for creating a lag plot in Microsoft Excel 2007 are provided on page 17.

The importance of autocorrelation that reflects regular, periodic cycling of water quality conditions (e.g., diel/seasonal cycling) to TMDL effectiveness monitoring is discussed in the preceding section. Autocorrelation that indicates short-term persistence in water quality conditions from one observation to the next must also be reviewed. Persistence refers to the tendency of similar values to follow one another (large values follow large values and small values follow small values). Short-term persistence can be evaluated by exploring autocorrelation at small time lags, generally at the sampling frequency (i.e., time lag equal to 1 sampling interval, or lag-1 autocorrelation).

Lag-1 autocorrelation is noteworthy for analysis of TMDL effectiveness monitoring data since a basic assumption of many standard statistical tests is that the dataset is comprised of independent or random observations. The presence of strong lag-1 autocorrelation in a dataset implies that observed values are not random. Ultimately, this reduces the amount of information contained within the data, as a given sample provides minimal "new" information beyond what was already known from the previous observation. While lag-1 autocorrelation is generally strongest for samples collected at a high frequency, it can also be present in data collected at monthly or even annual intervals.

Strong lag-1 autocorrelation can affect the accuracy of long-term trend detection and increase the frequency of "false positives" reported by statistical tests (i.e., the actual confidence level is less than that specified under a classical statistical test). Methods to "correct" for autocorrelation generally involve modifying the dataset to remove the autocorrelation effect or the use of statistical tests that have been specifically developed for autocorrelated data. Simple options include performing a classical statistical test using a subset of the original dataset or grouped values (such as monthly or quarterly means). If lag plots point to the presence of strong short-term persistence, readers are encouraged to consult advanced statistical resources (such as Helsel & Hirsch (2002) or Kirchner (2001)) or a statistician to gain insight into methods specific to their analysis objectives.

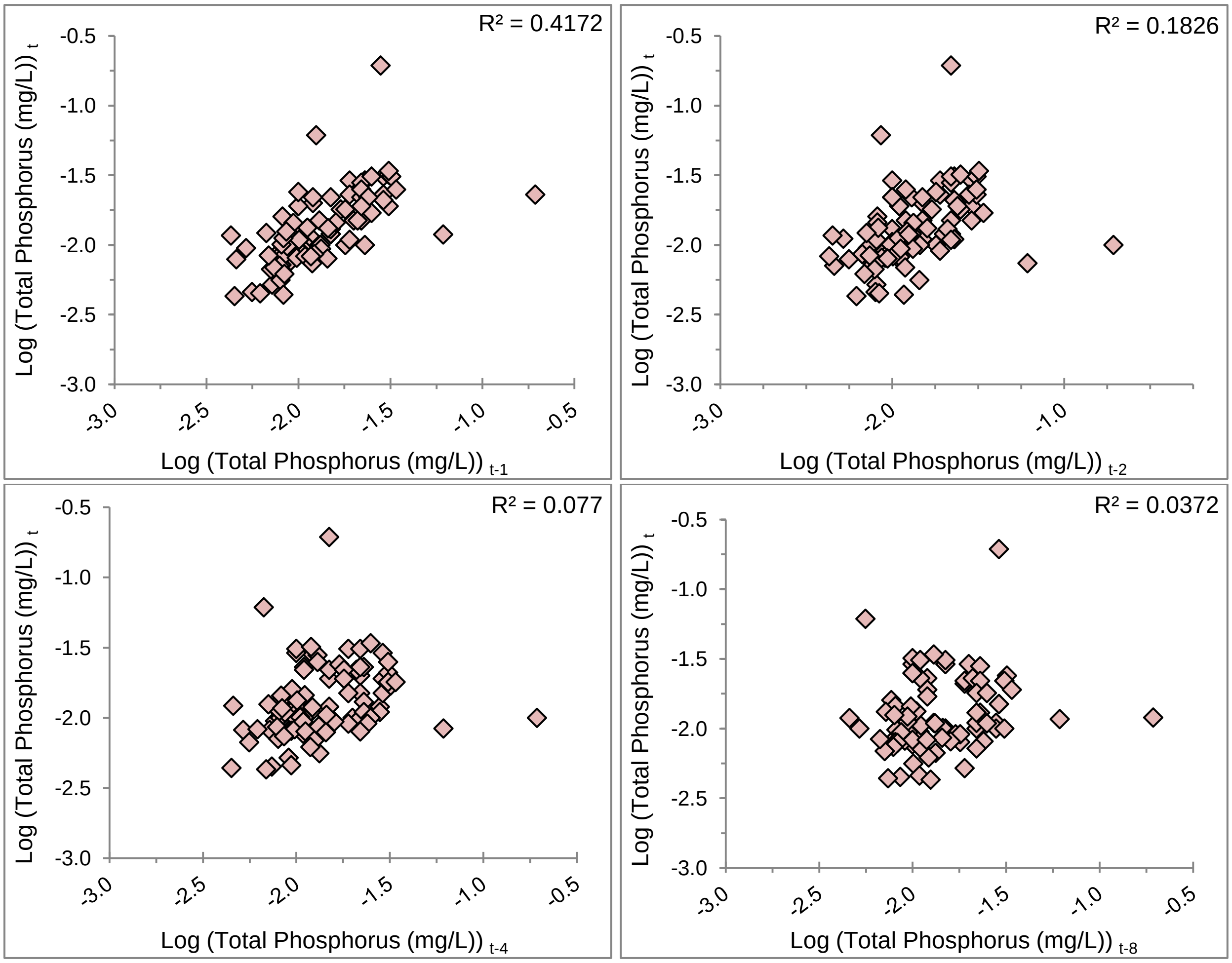
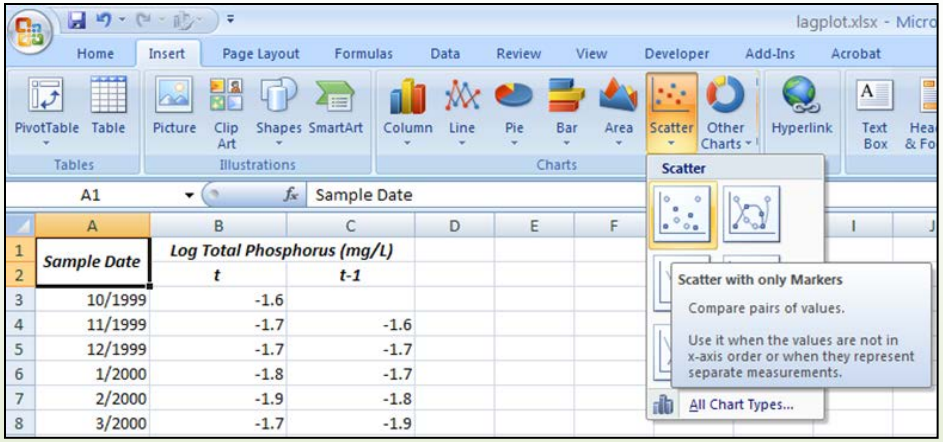


Figure 6. Lag plots of Cedar River, WA total phosphorus concentration at lag times of one month (top left), two months (top right), four months (bottom left), and eight months (bottom right). Note that the strong correlation evident at the one month time lag fades at higher lag times. This pattern is typical of data with short-term persistence.

EDA in Excel. Creating a Lag Plot

In Excel 2007, use the ***Scatter Chart*** tool to create a lag plot. In the example below, a 1-month lag plot is created for Burnt Bridge Creek fecal coliform concentration using monthly sample data collected over the period from October 1999 through September 2007.

Step 1. Arrange the data so that one column contains observed values and one column contains values offset by the selected time lag. Click the ***Insert*** tab. In the ***Charts*** section, click ***Scatter***. Select ***Scatter with only Markers***.



A1 | Sample Date

	A	B	C	D	E	F
1	Sample Date	Log Total Phosphorus (mg/L)				
2		t	t-1			
3	10/1999	-1.6				
4	11/1999	-1.7	-1.6			
5	12/1999	-1.7	-1.7			
6	1/2000	-1.8	-1.7			
7	2/2000	-1.9	-1.8			
8	3/2000	-1.7	-1.9			

Scatter with only Markers

Compare pairs of values.

Use it when the values are not in x-axis order or when they represent separate measurements.

Step 2. A blank chart is created. Right-click on the chart area and choose ***Select Data***.

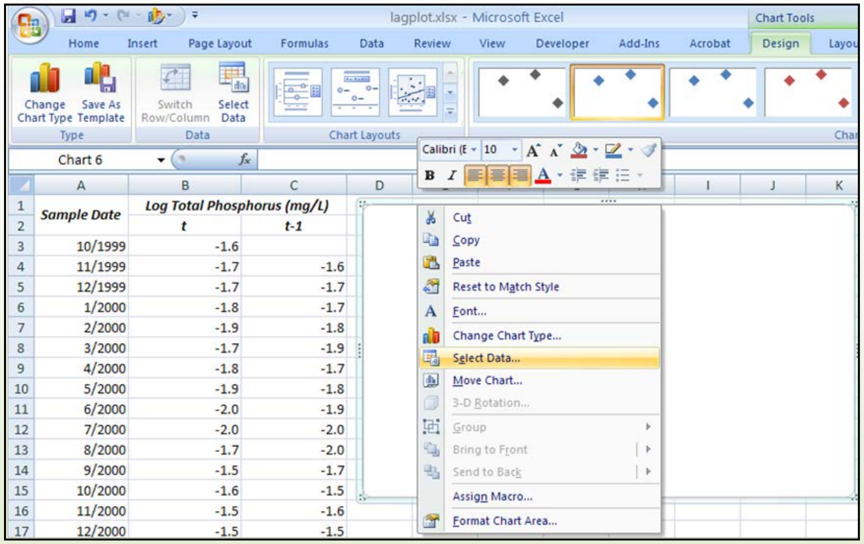
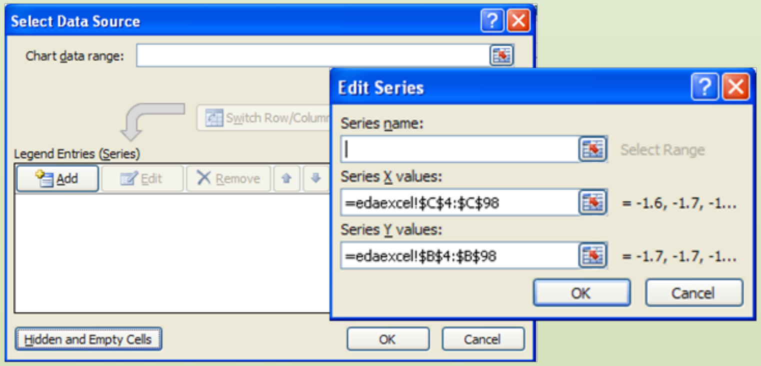


Chart 6

	A	B	C	D
1	Sample Date	Log Total Phosphorus (mg/L)		
2		t	t-1	
3	10/1999	-1.6		
4	11/1999	-1.7	-1.6	
5	12/1999	-1.7	-1.7	
6	1/2000	-1.8	-1.7	
7	2/2000	-1.9	-1.8	
8	3/2000	-1.7	-1.9	
9	4/2000	-1.8	-1.7	
10	5/2000	-1.9	-1.8	
11	6/2000	-2.0	-1.9	
12	7/2000	-2.0	-2.0	
13	8/2000	-1.7	-2.0	
14	9/2000	-1.5	-1.7	
15	10/2000	-1.6	-1.5	
16	11/2000	-1.5	-1.6	
17	12/2000	-1.5	-1.5	

Step 3. In the ***Select Data Source*** dialog box, click ***Add***. In the Edit Series dialog box, enter the range of cells containing lagged values under ***Series X values***. Enter the range of cells containing unlagged values under ***Series Y values***. Click *OK*.

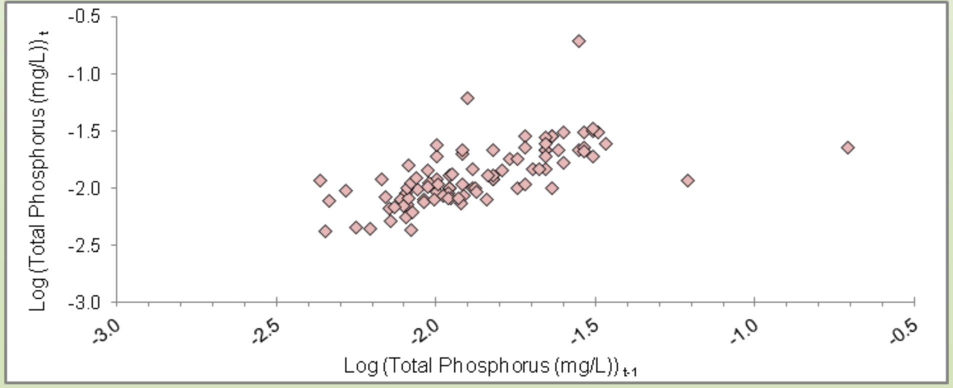


Series name:

Series X values: =edaexcel!$C$4:$C$98 = -1.6, -1.7, -1...

Series Y values: =edaexcel!$B$4:$B$98 = -1.7, -1.7, -1...

Step 4. Edit axis limits, titles, and data markers as needed.



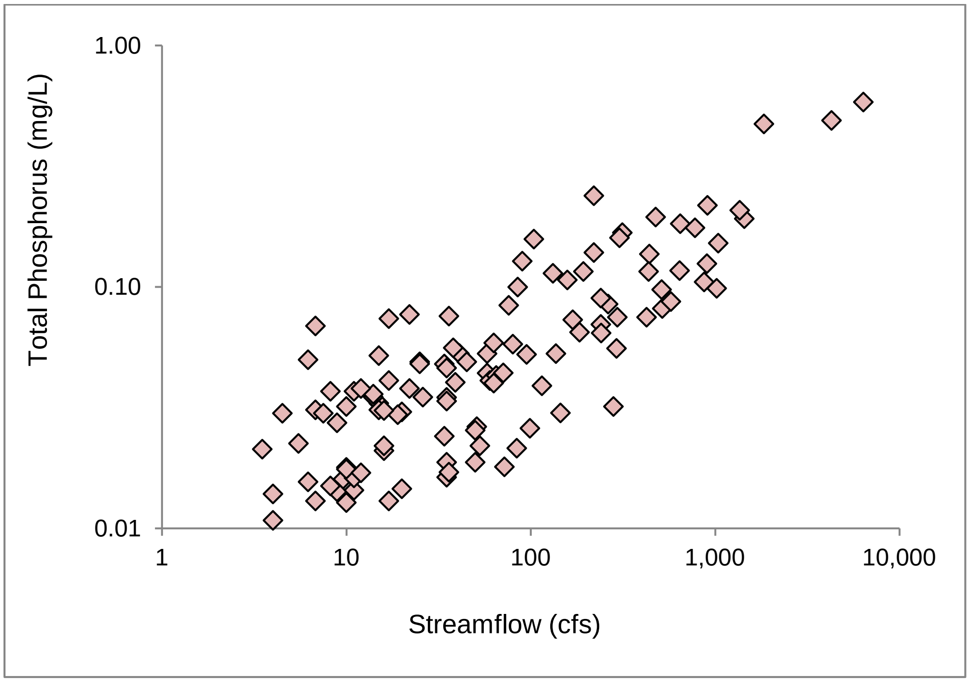
Covariates and Scatterplots

Under a TMDL effectiveness monitoring program, average water quality conditions, and variability in these conditions, are quantified at each monitoring site for the purpose of comparing sites or evaluating change over time. Oftentimes, variability in the parameter(s) of interest is tied to variability in other measurable parameters that directly or indirectly reflect the ultimate drivers of water quality change. Observations of these ***covariates*** increase/decrease with those of the parameter of interest (the data vary together). Covariate data can be used to "sort out" natural variability versus variability attributed to TMDL implementation; therefore, the inclusion of covariates in TMDL effectiveness monitoring can greatly improve the power of the statistical methods applied to detect differences between sites or change over time. For the case of common parameters addressed in TMDLs (e.g., nutrient concentrations, dissolved oxygen concentration, water temperature) common covariates include precipitation, air temperature, and streamflow.

Covariates can be identified by constructing a ***scatterplot***. Scatterplots contain the independent variable on the x-axis and the dependent variable on the y-axis. Note that time series and lag plots are specific types of scatterplots. A scatterplot can be created in Microsoft Excel 2007 by following the instructions provided on page 10 for creating a time series plot (with covariate data replacing sampling date).

A scatterplot allows for a visual review of the correlation between two parameters (do the two parameters vary together?).

Figure 7 contains scatterplots of streamflow vs. total phosphorous concentration and streamflow vs. fecal coliform concentration for Hangman Creek, WA. A strong relationship between flow and total phosphorous is apparent, with higher flows corresponding to higher concentrations. Conversely, no relationship is evident between flow magnitude and fecal coliform concentration. If a scatterplot reveals correlation between two parameters, parametric (multiple linear regression) and nonparametric (Kendall trend test with LOWESS) trend analysis techniques are available to account for covariates to better identify trends. A common method for including covariates when comparing two groups (e.g., before and after TMDL implementation) is Analysis of Covariance (ANCOVA). Refer to Helsel & Hirsch (2002) for a detailed discussion of such methods.



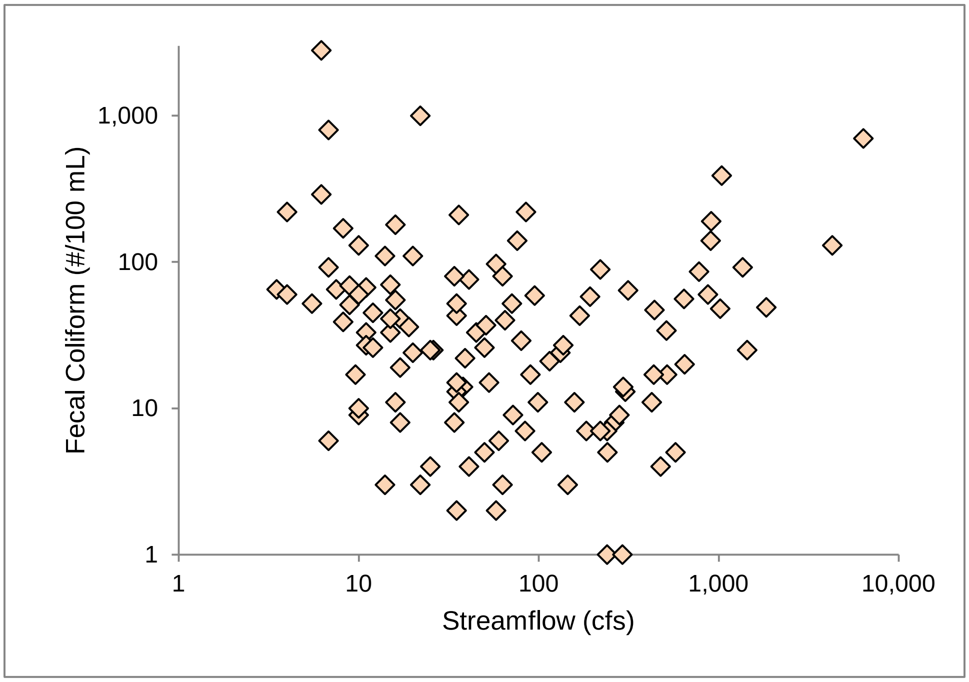


Figure 7. Scatterplots of total phosphorous (top) and fecal coliform (bottom) concentration vs. streamflow for Hangman Creek, WA from 10/1999 through 9/2009. Note the difference in the degree of correlation between streamflow and each water quality parameter.

Censored Data

Laboratory analysis of water quality samples often returns values that are less than the lower detection limit (non-detects) or greater than the upper detection limit of the analytical method applied. In such cases, the actual sample value is only partially known. Data with measured values that are only partially known are referred to as ***censored data***. A researcher may be tempted to discard censored samples to ease analysis. Though these data contain information that is limited relative to other data points, they contain information nonetheless, and therefore should not be removed for subsequent analysis. Doing so will bias measures of the central tendency (e.g., mean) of the dataset and its variability, and can lead to inaccurate interpretation of the data.

Several options are available to address censored data. A basic method is to replace the censored value with a real number value to conform to other measured data. For example, values reported as less than the detection limit can be estimated as zero, as the detection limit, or as a percentage of the detection limit. Replacement values are then used when calculating summary statistics or performing a statistical test. This approach can bias the sample mean and standard deviation and should only be used if a small number of data points are censored. More complex methods of estimating summary statistics should be applied when several data points are censored (see Helsel & Hirsh (2002); U.S. Environmental Protection Agency (2006)). A general rule of thumb is to apply simple substitution only if censored data make up less than 15% of the dataset (U.S. Environmental Protection Agency, 2006).

For hypothesis testing, substitution of censored data is not required if a nonparametric test is performed. For example, substitution is not needed for trend analysis if the nonparametric Mann-Kendall test is applied. However, results of nonparametric tests should be viewed with caution if the number of censored data points rises above 50%. If censored data are prominent, refer to Helsel & Hirsch (2002) and U.S. Environmental Protection Agency (2006) for a detailed review of estimation and analysis of censored data.

3. Conclusions

This guide presents several graphical methods for exploring and understanding the basic characteristics of TMDL effectiveness monitoring data. The information acquired through data exploration is highly valuable for multiple phases of a TMDL effectiveness monitoring project. During project planning, it is recommended that practitioners undertake exploration of pilot data. At minimum, pilot data should include observations of those water quality parameters that are the focus of TMDL effectiveness monitoring. Pilot data collected from study sites is preferred (though data from comparable sites can also be explored) and should include potential covariate data whenever possible. Exploration of pilot data allows planners to make data-driven decisions regarding which parameters to include in TMDL effectiveness monitoring, and sampling frequency, timing, and duration. Pilot data exploration also provides preliminary direction for post-monitoring confirmatory data analysis and informs selection of a study design (see *Technical Guidance for Designing a TMDL Effectiveness Monitoring Plan*). To facilitate the exploration of pilot data, an Excel-based TMDL effectiveness monitoring planning tool is provided as a complement to this document.

Following project implementation, exploration of TMDL effectiveness monitoring data is required to finalize the selection of methods used for confirmatory data analysis. While highly dependent on data characteristics, confirmatory data analysis decisions should always be made with the analysis objective in mind. Potential analysis objectives for TMDL effectiveness monitoring data are outlined in Table 1 and include:

1. Compare two independent data groups;
2. Compare two data groups with matched pairs;
3. Compare two data groups while adjusting for covariates;
4. Evaluate the relationship between one data group and time;
5. Evaluate the relationship between one data group, time, and covariates.

Analysis objectives 1 and 2 deal with two distinct groups of observations of a single water quality parameter (e.g., dissolved oxygen concentration before and after TMDL implementation). The difference between the two relates to the presence of paired observations between groups. For example, a monitoring program that includes concurrent sampling of upstream and downstream locations provides a group of “upstream” observations with a logical matched pair in the “downstream” group. Alternatively, a study of water quality before and after TMDL implementation provides “before” observations that have no matched pair in the “after” group (the groups are independent). Analysis objective 3 is common for a paired watersheds study, where observed values of a parameter during pre-treatment and post-treatment periods are compared while adjusting for natural variability using observations from the control watershed (the covariate). Analysis objectives 4 and 5 result from trend monitoring, with objective 5 applied if covariates are included in monitoring activities.

A number of confirmatory tests are available to address the analysis objectives discussed above. Alternative tests for a given analysis objective can generally be grouped as parametric or nonparametric tests. Remember that the fundamental difference between the two is that parametric tests assume that data are normally distributed, while nonparametric tests include no assumptions about the distribution of data. Other data characteristics identified through data exploration (e.g., the presence of outliers or censored data) will also influence the selection of a parametric vs. nonparametric test, and the assumptions of each alternative test should be evaluated before proceeding.

Table 1 contains several common parametric and nonparametric tests available to address the potential analysis objectives for TMDL effectiveness monitoring data discussed above. Detailed information on these and other statistical tests is provided in Helsel & Hirsch (2002) and U.S.

Environmental Protection Agency (2006). A number of these tests can be carried out manually using spreadsheet software such as Microsoft Excel (or using Excel's built-in Data Analysis ToolPak). Others require advanced statistical software (such as the free R statistical package).

Table 1. Some useful statistical tests for analysis of TMDL effectiveness monitoring data.

Objective	**Study Design**	**Statistical Test**	
		Parametric	**Nonparametric**
Compare two independent data groups	Before/After; Upstream/Downstream	Two-Sample (Unpaired) t-Test	Rank Sum Test
Compare two data groups with matched pairs	Paired Watersheds; Upstream/Downstream	Paired t-Test	Signed Rank Test
Compare two data groups while adjusting for covariates	Paired Watersheds	Analysis of covariance (ANCOVA)	
Evaluate the relationship between one data group and time (without seasonality)	Trend Monitoring	Linear Regression	Mann-Kendall Test
Evaluate the relationship between one data group and time (with seasonality)	Trend Monitoring	Linear Regression with Seasonal Term	Seasonal Kendall Test
Evaluate the relationship between one data group, time, and other variables (without seasonality)	Trend Monitoring	Multiple Linear Regression	Mann-Kendall Trend Test with LOWESS
Evaluate the relationship between one data group, time, and other variables (with seasonality)	Trend Monitoring	Multiple Linear Regression with Seasonal Term	Seasonal Kendall Test with LOWESS

4. References

Helsel, D. R. (2006). *Computer program for the Kendall family of trend tests: U.S. Geological Survey Scientific Investigations Report 2005–5275.*

Helsel, D. R., & Hirsch, R. M. (2002). Statistical Methods in Water Resources. In *Hydrologic Analysis and Interpretation*. Avaiable online at: http://water.usgs.gov/pubs/twri/twri4a3/.

NIST/SEMATECH , www.itl.nist.gov/div898/. (2003). *Engineering Statistics Handbook*. Available online at: http://www.itl.nist.gov/div898/handbook/.

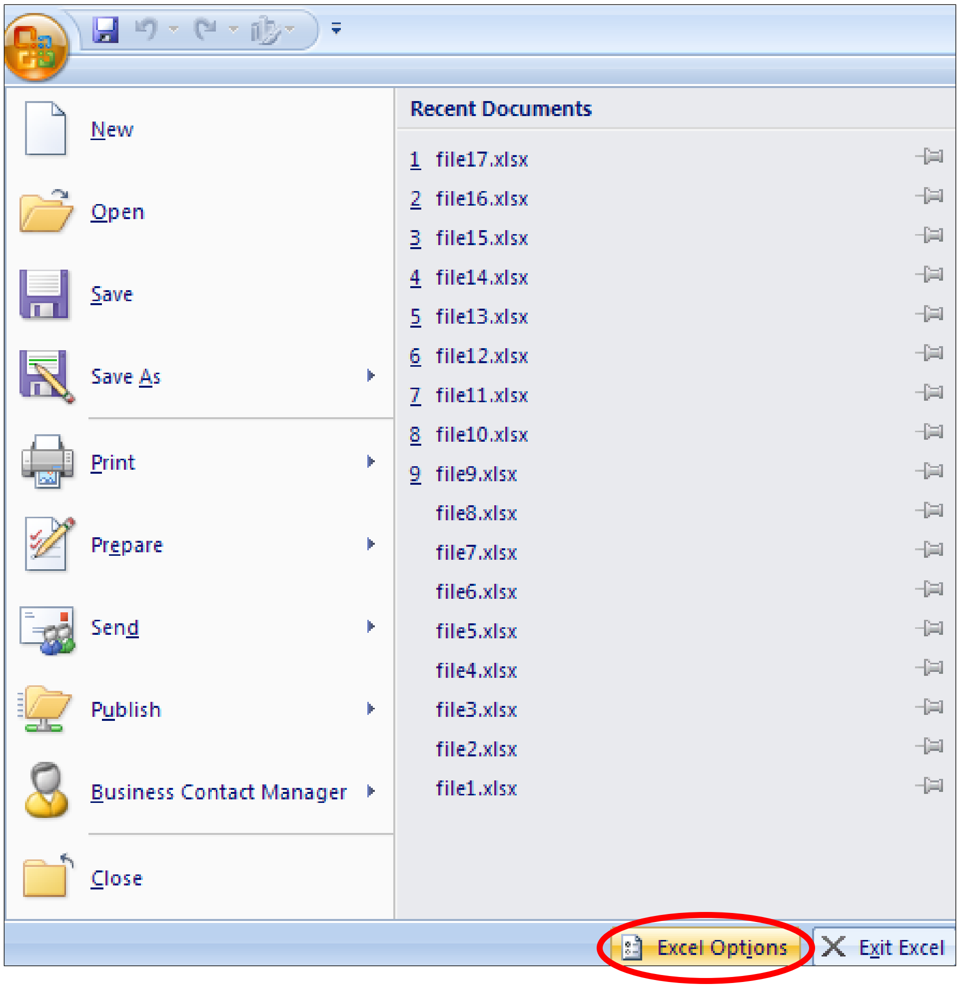
U.S. Environmental Protection Agency. (2006). *Data Quality Assessment: Statistical Methods for Practitioners EPA QA/G-9S*. Available online at: http://www.epa.gov/quality/qs-docs/g9s-final.pdf.

Velleman, P. F., & Hoaglin, D. C. (1981). *Applications, Basics, and Computing of Exploratory Data Analysis*. Duxbury Press. Available online at: http://dspacelocal.library.cornell.edu/DSpace/IFUP_Catalog.pdf.

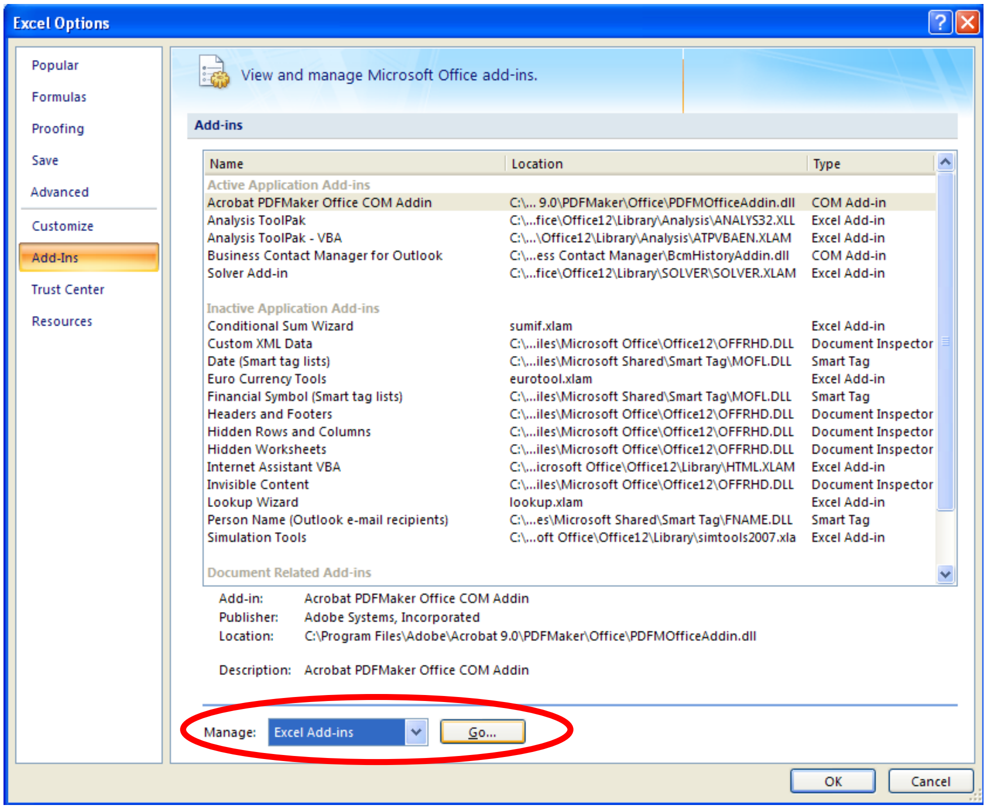
Appendix A. Loading the Analysis ToolPak in Microsoft Excel 2007

Microsoft Excel 2007 includes several built-in data analysis tools. These tools are not accessible to users until the Analysis ToolPak is installed. Below are step-by-step instructions for installing the Analysis ToolPak.

Step 1. Click the Office Button at the top-left of the screen. Click *Excel Options*.



Step 2. In the *Excel Options* box, click the *Add-Ins* tab on the left. In the *Manage* drop-down list, select *Excel Add-ins*. Click *Go*…



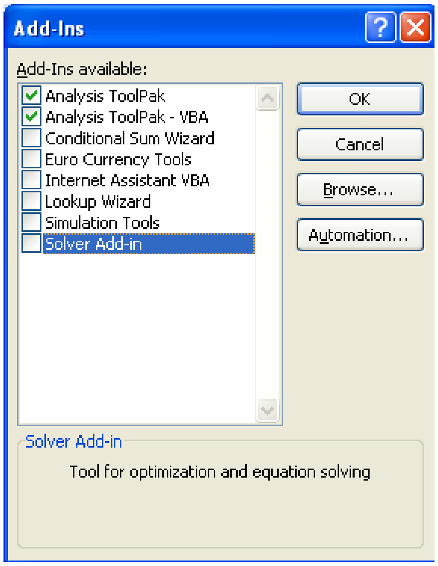
Name	Location	Type
Active Application Add-ins		
Acrobat PDFMaker Office COM Addin	C:\... 9.0\PDFMaker\Office\PDFMOfficeAddin.dll	COM Add-in
Analysis ToolPak	C:\...fice\Office12\Library\Analysis\ANALYS32.XLL	Excel Add-in
Analysis ToolPak - VBA	C:\...\Office12\Library\Analysis\ATPVBAEN.XLAM	Excel Add-in
Business Contact Manager for Outlook	C:\...ess Contact Manager\BcmHistoryAddin.dll	COM Add-in
Solver Add-in	C:\...fice\Office12\Library\SOLVER\SOLVER.XLAM	Excel Add-in
Inactive Application Add-ins		
Conditional Sum Wizard	sumif.xlam	Excel Add-in
Custom XML Data	C:\...iles\Microsoft Office\Office12\OFFRHD.DLL	Document Inspector
Date (Smart tag lists)	C:\...iles\Microsoft Shared\Smart Tag\MOFL.DLL	Smart Tag
Euro Currency Tools	eurotool.xlam	Excel Add-in
Financial Symbol (Smart tag lists)	C:\...iles\Microsoft Shared\Smart Tag\MOFL.DLL	Smart Tag
Headers and Footers	C:\...iles\Microsoft Office\Office12\OFFRHD.DLL	Document Inspector
Hidden Rows and Columns	C:\...iles\Microsoft Office\Office12\OFFRHD.DLL	Document Inspector
Hidden Worksheets	C:\...iles\Microsoft Office\Office12\OFFRHD.DLL	Document Inspector
Internet Assistant VBA	C:\...icrosoft Office\Office12\Library\HTML.XLAM	Excel Add-in
Invisible Content	C:\...iles\Microsoft Office\Office12\OFFRHD.DLL	Document Inspector
Lookup Wizard	lookup.xlam	Excel Add-in
Person Name (Outlook e-mail recipients)	C:\...es\Microsoft Shared\Smart Tag\FNAME.DLL	Smart Tag
Simulation Tools	C:\...oft Office\Office12\Library\simtools2007.xla	Excel Add-in
Document Related Add-ins		

Add-in: Acrobat PDFMaker Office COM Addin
Publisher: Adobe Systems, Incorporated
Location: C:\Program Files\Adobe\Acrobat 9.0\PDFMaker\Office\PDFMOfficeAddin.dll

Description: Acrobat PDFMaker Office COM Addin

Manage: Excel Add-ins Go...

Step 3. In the *Add-Ins* box, check the boxes next to *Analysis ToolPak* and *Analysis ToolPak – VBA*. Click *OK*.



Add-Ins available:

☑ Analysis ToolPak
☑ Analysis ToolPak - VBA
☐ Conditional Sum Wizard
☐ Euro Currency Tools
☐ Internet Assistant VBA
☐ Lookup Wizard
☐ Simulation Tools
☐ Solver Add-in

Solver Add-in

Tool for optimization and equation solving

This page is purposely left blank

Appendix F. DRAFT - Dungeness Bay and Dungeness River Watershed Fecal Coliform Bacteria TMDL: Addendum to Water Quality Effectiveness Monitoring Report: Detailed Methods and Results of Trend Analysis

This page is purposely left blank

Dungeness Bay and Dungeness River Watershed Fecal Coliform Bacteria Total Maximum Daily Load

Addendum to Water Quality Effectiveness Monitoring Report: Detailed Methods and Results of Trend Analysis

8/13/2010 Draft

Introduction

This addendum provides supplementary information on the fecal coliform trend analysis in the Dungeness Bay and Dungeness River Watershed Water Quality Effectiveness Monitoring Report. It contains an expanded description of the statistical methods and results and provides additional guidance on interpreting the results of the trend analysis. Some additional analyses were conducted for this addendum, but the results of those analyses did not substantially change the findings of the main report.

Methods

Dungeness Bay Trends

Multiple linear regression (MLR) was used to test for temporal trends in fecal coliform concentrations.

1. Fecal coliform (FC) concentrations follow a right-skewed, non-normal distribution. MLR requires normally distributed data, which can usually be achieved with a square root or log transformation of the raw data. The log transformation produced a more normal distribution with these data than the square root transformation. However, FC concentrations include some zero values, which cannot be log-transformed. To facilitate log-transformation of all values, one (1) was added to each raw FC value (i.e., LogFC = log(FC+1). This is a standard procedure for accommodating zero values in a log-transformation. However, because the choice of a constant that is added before transformation is subjective, the sensitivity of the trend analysis results to different constants (0.1 and 10) was evaluated. The distribution of ZLogFC (defined in step 3) was most normal with 1 as the constant. Coefficients on the model variables changed slightly with the alternative constants, but the overall conclusions about trends and variables that influence FC did not change.
2. MLR evaluates relationships between two or more predictor variables and a response variable; in this case, LogFC. The initial set of potential predictor variables included date, date^2 (to represent non-linear trends[1]), season, water temperature, tide status, salinity, and flow rate from the Dungeness River. Season and tide status were removed from this variable list because they are correlated with temperature and salinity, respectively. A stepwise variable selection procedure in Minitab version 13.30 was used to identify significant predictors from the remaining variables. The stepwise procedure enters significant variables into the model one by one (forward selection) and removes variables that become non-significant through the addition of another variable (backwards elimination), with the critical p-value (α) in this case set to 0.05. In all cases, the signs on

---

[1] Evidence for non-linear trends can be evaluated in regression analysis by including a squared version of the time variable. In this case, this approach treats log fecal coliform concentrations as quadratic function of time. The additional variable is warranted if a visual examination of the time series suggests a non-linear trend (see Appendix C of the EM report, Figure C-5 for an example).

the coefficients of each variable were evaluated to ensure that they were consistent with theoretical understanding of the influence of each variable.

3. Regressions were run for each station individually, for all stations grouped together, and for inner and outer bay station groups. In the grouped station regressions, the mean LogFC at each sample station was subtracted from all individual samples at that station to factor out differences among stations. This variable was called ZLogFC because of the similarity of this approach to a z-score transformation. This approach is equivalent to a mixed effects model, where station is a random effect.

Dungeness Bay Tributary Trends

MLR was also used to test for temporal trends in log-transformed FC concentrations at tributary sampling stations. Potential predictor variables included year, year^2, season (irrigation (April 15 to September 15) or non-irrigation), and rainfall (the cumulative rainfall depth at Sequim for a period spanning two days before the sample through the day of the sample).

The same stepwise variable selection procedure used in the marine analysis was used to identify significant predictors. Individual and grouped regressions were run for all sites where samples were collected throughout the 1999-2009 period (DR0.1, DR0.8, DR3.2, MAT0.1, MAT1.9, MAT3.2, and MC0.2). As with the marine analysis, in the grouped station regression, the mean LogFC at each sample station was subtracted from all individual samples at that station to factor out differences among stations. This normalized variable was called ZLogFC.

Results

Dungeness Bay Trends

Regression results on data from all marine stations indicate that FC concentrations significantly decreased by 24% from 1999 to 2009 (Figure A-1). This equates to a decrease in the geometric mean from 8.2 cfu/100 mL in 1999 to 6.1 cfu/100 mL in 2009. The statistical evidence for this decrease is moderately strong ($p=0.01$). The trend is non-linear, suggesting a decrease, followed by a leveling off, with perhaps a small increase in the last few years (Figure A-1). Because the shape of this trend was fit to data from 1999-2009, the trend should not be extrapolated before or after this period. The trend is in the geometric mean concentration among all marine stations. This trend was only detectable in the full dataset; inner and outer bay station groups and all but one of the individual stations did not have significant trends. There is weak evidence that FC concentrations decreased at station 182.

The finding of a trend in the full dataset coupled with the lack of trend at most individual stations suggests that unless sample size is high enough, the high variance of FC concentrations can obscure trends. In addition, it was only possible to detect the trend over time by factoring out the strong effects of temperature, salinity, and the flow rate of the Dungeness River on FC concentrations. Even after factoring out the influence of these environmental variables, the trend in FC concentrations over time is difficult to perceive visually (Figure A-2). It should therefore be emphasized that variance among individual sample values remains high and that careful partitioning of the sources of that variance of FC is required to detect trends over time. Detailed regression results are presented in Table A-1.

The influence of variable environmental conditions on FC concentrations has been documented elsewhere. Other studies have shown that FC bacteria survive longer in cold water than in warm water (Burkhardt et al., 2000; Wait and Sobsey, 2001; Boehm et al., 2004). FC bacteria also survive longer when their exposure to sunlight is limited (Burkhardt et al., 2000). Salinity is associated with the relative dominance of marine and fresh water in the bay, and therefore probably indicates the degree of dilution by low-FC seawater on high-FC river outflow. Notably, there was a significant increase in the salinity of samples between 1999 and 2009 ($p<0.001$) (Appendix C of EM report, Figure C-1). This trend could be due to changing circulation patterns in the bay or some aspect of the sampling protocol that inadvertently changed. However, because the trend over time remained significant even after accounting for the influence of salinity, the decreasing trend was probably caused by factors unrelated to changes in salinity.

There is no evidence that flows from the Dungeness River have changed over this time period (regression of log(monthly flow at river mile 0.8) on date, $p=0.29$). Because variations in salinity, temperature, and river flows have the potential to obscure trends in FC over time, future evaluations should also use MLR to factor out the effects of these variables.

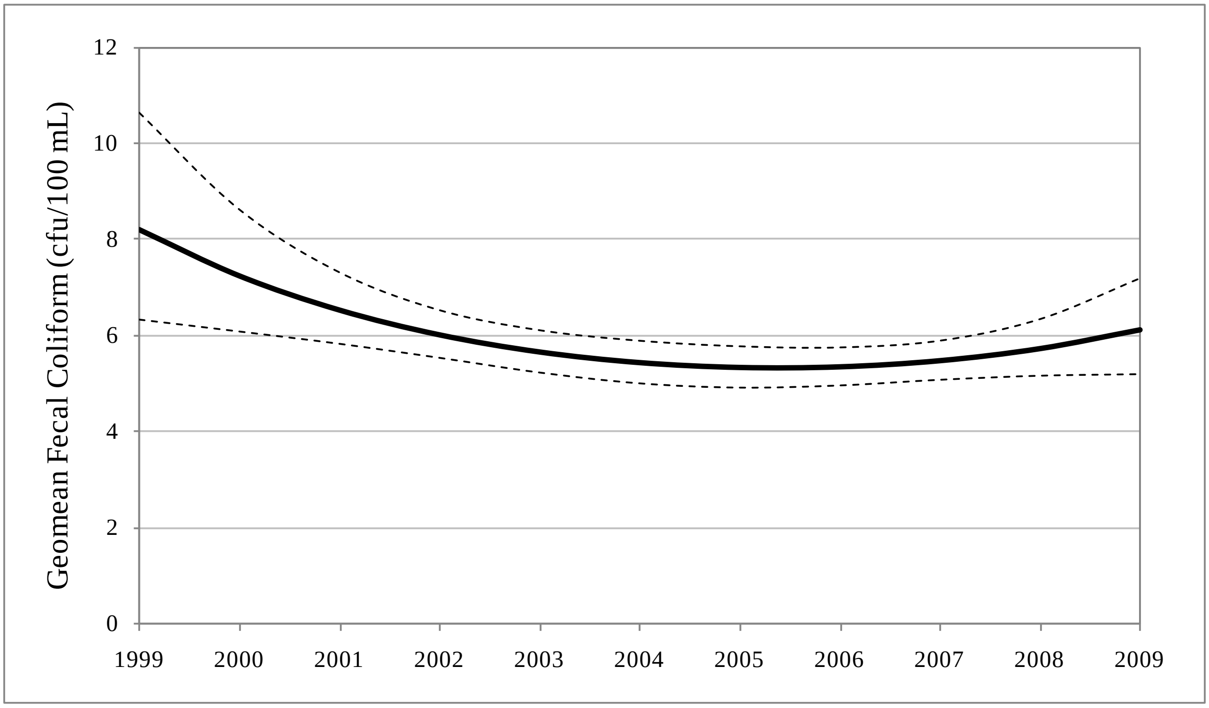


Figure A-1. Trend in geometric mean FC concentrations over time for pooled data from all marine sample stations from 1999-2009. The mean values of salinity, temperature and log(Dungeness River flow) were inserted into the ALL station regression equation in Table A-1 to isolate the trend over time from variance in these other variables. Geometric means were calculated by back-transforming log values predicted by the regression equation. Dashed lines delineate 95 percent confidence intervals.

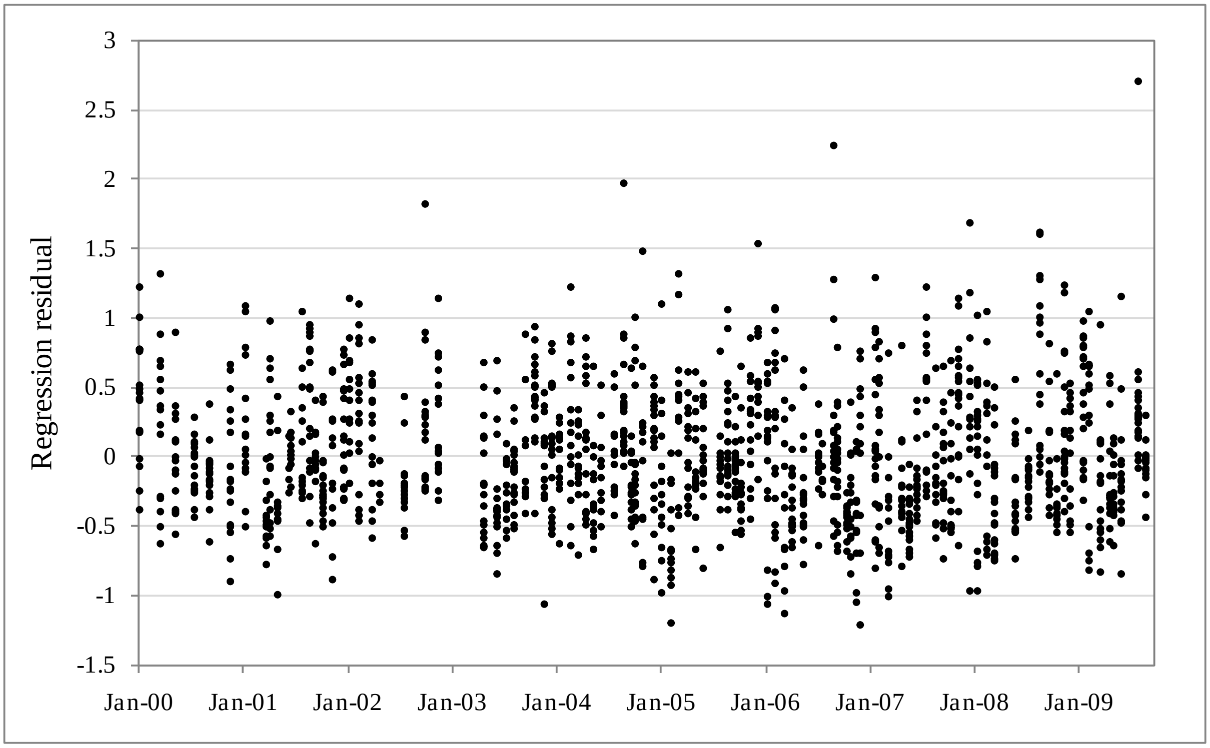


Figure A-2. Residuals from the regression equation ZLogFC ~ Salinity + Temperature + log(Dungeness River flow) on pooled data from all marine sample stations from 1999-2009. The slight, but statistically significant curved decreasing trend in these residuals is the basis for the trend in Figure A-1.

Table A-1. Regression model coefficients (± 1 standard error) for Dungeness Bay stations. The response variable is LogFC for individual stations and ZLogFC for the grouped models. Coefficients reported for significant ($p<0.05$) variables only. Bold text indicates $p<0.001$.

Station	Constant	Temperature	Salinity	River Flow	Date	Date^2
ALL	**2.050 ± 0.148**	**-0.104 ± 0.005**	**-0.016 ± 0.003**	0.109 ± 0.035	-0.151 ± 0.059	0.031 ± 0.013
Dungeness	**2.105 ± 0.159**	**-0.108 ± 0.006**	**-0.016 ± 0.003**	0.101 ± 0.038	-0.156 ± 0.065	0.030 ± 0.015
Inner Bay	**2.642 ± 0.161**	**-0.125 ± 0.007**	**-0.025 ± 0.006**	-	-	-
106	**2.717 ± 0.326**	**-0.095 ± 0.020**	-0.034 ± 0.011	-	-	-
107	**2.537 ± 0.389**	**-0.068 ± 0.018**	-0.041 ± 0.014	-	-	-
108	**2.110 ± 0.168**	**-0.145 ± 0.018**	-	-	-	-
109	**3.019 ± 0.417**	**-0.150 ± 0.016**	-0.030 ± 0.015	-	-	-
110	**2.047 ± 0.140**	**-0.140 ± 0.015**	-	-	-	-
111	**2.284 ± 0.155**	**-0.160 ± 0.016**	-	-	-	-
112	**2.009 ± 0.160**	**-0.138 ± 0.016**	-	-	-	-
Outer Bay	**1.898 ± 0.109**	**-0.069 ± 0.010**	**-0.020 ± 0.003**	-	-	-
103	0.515 ± 0.402	-	-0.022 ± 0.008	0.272 ± 0.111	-	-
104	**2.359 ± 0.252**	**-0.097 ± 0.023**	-0.020 ± 0.007	-	-	-
105	**1.905 ± 0.226**	**-0.114 ± 0.025**	-	-	-	-
113	**2.206 ± 0.246**	**-0.098 ± 0.026**	-0.018 ± 0.006	-	-	-
114	**1.521 ± 0.269**	-	-0.027 ± 0.010	-	-	-
115	**3.091 ± 0.415**	-0.074 ± 0.025	**-0.062 ± 0.014**	-	-	-
Jamestown	**2.452 ± 0.450**	**-0.064 ± 0.014**	-0.038 ± 0.016	-	-	-
99	-0.370 ± 0.255	-	-	0.329 ± 0.104	-	-
100	**3.755 ± 0.892**	-0.061 ± 0.025	-0.089 ± 0.031	-	-	-
101	**1.269 ± 0.266**	-0.072 ± 0.029	-	-	-	-
102	**1.565 ± 0.243**	-0.098 ± 0.027	-	-	-	-
182	**1.762 ± 0.467**	-	-	-	-0.359 ± 0.154	-

Dungeness Bay Tributary Trends

Time series plots of FC concentration data on the seven consistently sampled tributary stations are presented in Appendix C of the EM report (Figures C-2 through C-8). Regression results on pooled data from these stations indicate that FC concentrations significantly decreased from 1999 to 2009. As with the Dungeness Bay trend, the best model includes the squared term for year, which indicates a non-linear trend (Figure A-3, Table A-2).

The amount of rainfall just prior to the sample also significantly influenced FC concentrations. Specifically, the rainfall term was significant in the multiple regression model and higher FC concentrations were associated with higher amounts of rainfall. However, because the temporal trend remained significant even after accounting for the influence of rainfall events, it was probably caused by factors that are unrelated to patterns of rainfall over the study period.[2] (The total annual rainfall is presented in Appendix C of the EM report, Figure C-9.)

There was also a strong seasonal pattern, with higher FC concentrations occurring during the irrigation season. When data from the two seasons were evaluated separately, most of the overall trend appears to be in data collected during the irrigation season (Figure A-3). In addition, the irrigation season trend shows the most pronounced increase in recent years of all the trends.

Including season and rainfall information in the regression model increased the weight of evidence for a temporal trend by eliminating some of the variance in FC concentrations. However, even after factoring out the influence of these environmental variables, the trend in FC concentrations over time is difficult to perceive visually (Figure A-5). It should therefore be emphasized that variance among individual sample values remains high and that careful partitioning of the sources of that variance of FC is required to detect trends over time.

The pattern in Figure A-5 also suggests that high FC values in approximately the first year of the time series could be responsible for the significance of the time term in the regression model. Indeed, the p-value on year in a regression on data from 2001-2009 is 0.07, which indicates weak support for a trend over this period. Flow in the Dungeness River was low during 2000, which would increase FC concentrations with the same loading.

Based on regressions for individual stations, significant year-round decreases in FC concentrations only occurred at two sites on Matriotti Creek: MAT0.1 and MAT1.9 (Table A-2, Figure A-4). However, there is weaker evidence for decreasing trends during the irrigation season at one other site: DR3.2 (Table A-2). As with the all-station analysis, the shape of the trends in all the individual models was non-linear, suggesting a rapid decrease, followed by a leveling off, with perhaps a small increase in the last few years (Figure A-4).

---

[2] When several variables are sequentially added to a multiple regression model, the addition of one variable may negate the significance of another variable. This results from correlation among the predictor variables and can make it difficult to determine which variable is most directly influencing the response variable. However, when two or more variables are significant in the same model, it can generally be assumed that they have independent effects on the response variable.

The results of the individual station regressions suggest that the trends in the pooled data are solely the result of the strong trends at MAT0.1 and MAT1.9. Because Matriotti Creek contributes a small fraction of the flow of the Dungeness River, interpreting these trends as representing watershed-wide improvement would be misleading. To evaluate this possibility, additional regressions were run for subsets of the pooled data to evaluate whether the two stations with the most obvious trends (MAT0.1 and MAT1.9) were driving the overall results. The year and year^2 terms were significant (<0.05) in a model that included only the three Dungeness River stations and in a model that included all sites except MAT0.1 and MAT1.9 (Table 2). This result indicates that the decreasing trend in the pooled data is not solely the result of the strong trends at MAT0.1 and MAT1.9. The p-values were higher than for all the data and the coefficient on year was lower, which indicate a weaker trend. Pooling the data from multiple stations provides more statistical power to cut through the high variance in FC data.

Interpreting Trend Analysis Results

Because of high variance in FC concentrations, complex statistical models were required to detect trends over time and to attribute variance to the appropriate sources. To illustrate the results of these analyses in simple terms, trends were isolated from other sources of variation and plotted in Figures A-1, A-3, and A-4. The following points should be considered when interpreting these figures and the accompanying narrative results.

1. The trend lines in Figures A-1, A-3, and A-4 represent the best statistical estimate of the geometric mean FC concentration over time. The 95% confidence intervals around these lines indicate uncertainty about the true shape of trends in geometric means; they do not delimit the area that contains 95% of individual samples. The variance in individual samples is much greater and did not change over the period of analysis.
2. The simple curved shape of the trends in Figures A-1, A-3, and A-4 is a compromise between a linear trend and a more complex shape. The curved trends fit the data better than linear trends and the additional parameters required to fit more complex trends are not justifiable with high-variance data such as the FC concentrations in this study. When evaluating these trends, the reader should consider that the real trend may deviate substantially from the fitted line. These deviations are likely to be contained by the 95% confidence intervals, which are wider at the ends of the study period. In particular, a flat trend, rather than a small increase, fits into the confidence intervals at the end of most of the trends (except the pooled tributary stations in the irrigation season). In addition, because all the trends were fit to data from 1999-2009, they should not be extrapolated before or after this period.
3. Figures A-1, A-3, and A-4 show trends over time that are isolated from other sources of variation, including salinity, temperature, river flows, rainfall, season, and differences among sampling stations. By holding these natural variables constant, the trends show how FC concentrations have changed relative to a standard condition. Most of these variables vary at shorter time scales than the period of analysis. However, the average salinity of samples increased significantly during this period. Because the trend in FC concentrations shown in Figure A-1 was calculated by holding salinity constant, it does not reflect the possible effect of capping drainage ditches on freshwater flows into the bay. The available data do not allow for distinguishing this explanation from changes in

circulation in the bay or inadvertent changes in some aspect of the sampling protocol. Incorporating the trend in salinity into the trend in FC concentrations produces an estimated 36% decrease in FC concentrations (vs. 24% with a constant salinity), so the primary result should be considered a conservative estimate.

4. It may seem counterintuitive that the pooled data regressions included significant trends while most of the individual station regressions did not. The FC data in this study exhibit very high variance which can mask modest trends which may have occurred at individual stations. When data from multiple stations are combined, they are said to “borrow strength” from one another, making the overall picture more clear. A significant trend in a pooled model does not imply that all the individual stations also had trends, but that they belong to a common population of stations which, on average, had a trend.

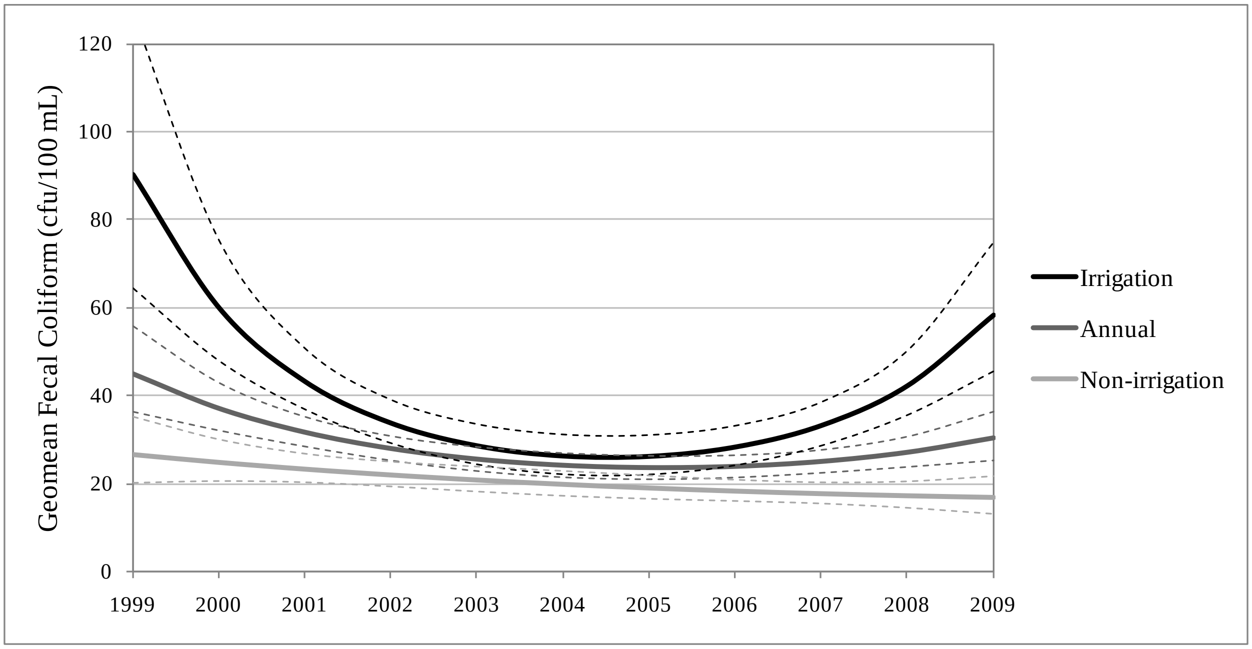


Figure A-3. Trends in geometric mean FC concentrations over time for pooled data from tributary stations from 1999-2009. Mean rainfall was inserted into the ALL station regression equation in Table A-2 to isolate the trend over time from variance in FC caused by rainfall. For the annual trend, season was set to 0.5 to show the average trend across seasons. Geometric means were calculated by back-transforming log values predicted by the regression equation. Dashed lines delineate 95 percent confidence intervals.

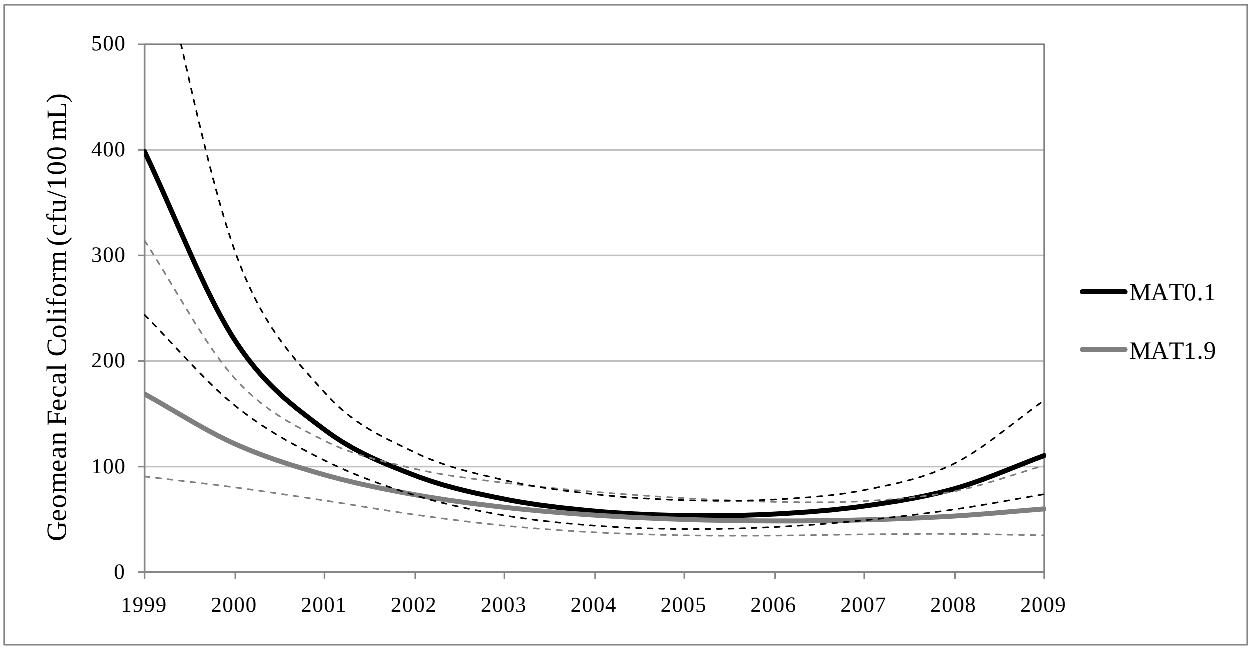


Figure A-4. Trends in geometric mean FC concentrations from 1999-2009 for two sites on Matriotti Creek. Mean values of rainfall and season were inserted into the annual regression equations for MAT0.1 and MAT1.9 in Table A-2 to isolate the trend over time from variance in FC caused by these environmental factors. Geometric means were calculated by back-transforming log values predicted by the regression equations. Dashed lines delineate 95 percent confidence intervals.

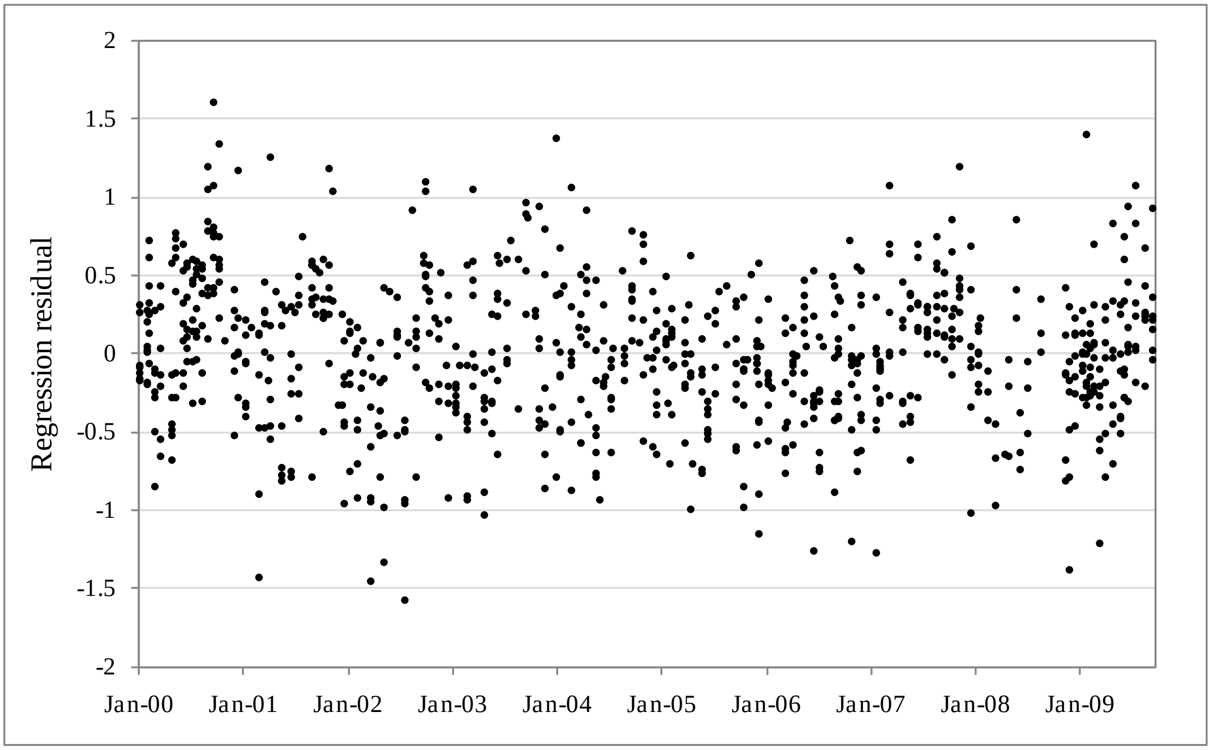


Figure A-5. Residuals from the regression equation ZLogFC ~ Season + Rain on pooled data from all tributary sample stations from 1999-2009. The slight, but statistically significant curved decreasing trend in these residuals is the basis for the trend in Figure A-3.

Table A-2. Regression model coefficients (± 1 standard error) for Dungeness Bay tributary stations. The response variable is LogFC for individual stations and ZLogFC for the grouped models. Coefficients reported for significant ($p<0.05$) variables only. Bold text indicates $p<0.001$.

Station	Constant	Season	Rain	Year	Year^2
ALL	**1.462 ± 0.050**	**0.309 ± 0.033**	**0.004 ± 0.001**	**-0.099 ± 0.021**	**0.008 ± 0.002**
	1.956 ± 0.075	Irrigation	-	**-0.195 ± 0.033**	**0.018 ± 0.003**
	1.241 ± 0.024	Non-irrigation	**0.004 ± 0.001**	-	-
DR0.1, 0.8, 3.2; MAT3.2; MC0.2	**1.360 ± 0.058**	**0.257 ± 0.038**	**0.004 ± 0.001**	-0.052 ± 0.024	0.005 ± 0.002
	1.823 ± 0.092	Irrigation	-	**-0.160 ± 0.041**	**0.015 ± 0.004**
	1.253 ± 0.069	Non-irrigation	**0.005 ± 0.001**	-	-
DR0.1, DR0.8, DR3.2	**1.386 ± 0.066**	**0.252 ± 0.043**	**0.004 ± 0.001**	-0.063 ± 0.028	0.006 ± 0.003
	1.801 ± 0.113	Irrigation	-	-0.152 ± 0.050	0.014 ± 0.005
	1.305 ± 0.075	Non-irrigation	**0.004 ± 0.001**	-	-
DR0.1	**0.895 ± 0.054**	**0.297 ± 0.077**	0.004 ± 0.002	-	-
	-	Irrigation	-	-	-
	-	Non-irrigation	-	-	-
DR0.8	**0.925 ± 0.046**	**0.295 ± 0.066**	0.003 ± 0.001	-	-
	-	Irrigation	-	-	-
	0.911 ± 0.041	Non-irrigation	0.004 ± 0.001	-	-
DR3.2	**0.573 ± 0.047**	-	0.005 ± 0.002	-	-
	1.108 ± 0.188	Irrigation	-	-0.299 ± 0.089	0.030 ± 0.008
	0.522 ± 0.058	Non-irrigation	0.006 ± 0.002	-	-
MAT0.1	**2.319 ± 0.113**	**0.492 ± 0.073**	0.004 ± 0.002	**-0.290 ± 0.048**	**0.024 ± 0.004**
	2.950 ± 0.169	Irrigation	-	**-0.371 ± 0.077**	**0.032 ± 0.007**
	2.310 ± 0.140	Non-irrigation	-	**-0.228 ± 0.062**	0.017 ± 0.006
MAT1.9	**1.852 ± 0.097**	**0.371 ± 0.098**	-	-0.039 ± 0.015	-
	2.496 ± 0.167	Irrigation	-	-0.198 ± 0.075	0.015 ± 0.007
	-	Non-irrigation	-	-	-
MAT3.2	**1.668 ± 0.064**	**0.441 ± 0.102**	-	-	-
	-	Irrigation	-	-	-
	1.526 ± 0.076	Non-irrigation	0.008 ± 0.003	-	-
MC0.2	-	-	-	-	-
	-	Irrigation	-	-	-
	-	Non-irrigation	-	-	-