STATE OF CALIFORNIA

DEPARTMENT OF TRANSPORTATION

Caltrans BMP Pilot Study Guidance Manual

April 2021

Document No. CTSW-RT-21-395.04.01

2021 BMP PILOT STUDY GUIDANCE MANUAL

This manual is not intended for use as a textbook or a substitute for engineering knowledge, experience, or judgment. It establishes uniform standards and procedures to use when planning and implementing pilot studies for evaluating stormwater best management practices.

CALTRANS BMP PILOT STUDY GUIDANCE MANUAL

ISSUED BY:

STATE OF CALIFORNIA
DEPARTMENT OF TRANSPORTATION
DIVISION OF ENVIRONMENTAL ANALYSIS
OFFICE OF STORMWATER PROGRAM DEVELOPMENT

Information concerning the contents of this manual may be obtained by writing:

STATE OF CALIFORNIA

DEPARTMENT OF TRANSPORTATION

DIVISION OF ENVIRONMENTAL ANALYSIS

OFFICE OF STORMWATER PROGRAM DEVELOPMENT

1120 N STREET

SACRAMENTO, CALIFORNIA 95814

Website address: https://dot.ca.gov/programs/environmental-analysis

2021 BMP PILOT STUDY GUIDANCE MANUAL

1. Report No.: CTSW-RT-21-395.04.01

2. Type of Report: Guidance Manual

3. Report Phase and Edition: Final

4. Title and Subtitle: Caltrans BMP Pilot Study Guidance Manual

5. Report Date: April 2021

6. Copyright Owner(s): California Department of Transportation

7. Caltrans Project Coordinator(s): Bhaskar Joshi

8. Performing Organization Names and Addresses:

Office of Water Programs
California State University, Sacramento
6000 J Street, 1001 Modoc Hall
Sacramento, CA 95819
www.owp.csus.edu

9. Contract No.: 43A0395

10. Task Order No.: 4

11. Sponsoring Agency Name and Address:

State of California Department of Transportation
Division of Environmental Analysis
Office of Stormwater Program Development
1120 N Street
Sacramento, CA 95814
https://dot.ca.gov/programs/environmental-analysis

12. Caltrans Functional Reviewer(s): Bhaskar Joshi

13. Supplementary Notes:

14. External Reviewer(s):

15. Key Words: stormwater, monitoring, BMP, pilot study

16. Distribution Statement: No restrictions. This document is available to the public upon request

17. No. of pages: 126 (including cover, TOC, etc.)

18. Abstract:

This manual contains guidelines for planning and conducting pilot studies that assess the feasibility, performance, costs, and maintenance requirements for stormwater best management practices (BMPs).

For individuals with sensory disabilities, this document is available in alternate formats upon request.

Please call or write to:

Stormwater Liaison,

Caltrans Division of Environmental Analysis, MS 27

P.O. Box 942874

Sacramento, CA 94274-0001

(916) 653-8896 Voice, or dial 711 to use a relay service.

Table of Contents

LIST OF APPENDICES

List of Tables

LIST OF FIGURES

Changes to this document from previous (2009) version

Change
General: • Changes made to text, tables, figures, captions, and headers to make the document ADA compliant • Throughout the document, hyperlinks were updated as necessary so that they continue to lead to the correct resources • Clarified that instructions in this manual follow Architectural-Engineering (A-E) Services Contracts and practitioners should follow practices for that type of delivery method. If other delivery methods are to be used, document now refers reader to Caltrans Project Development Procedures Manual (PDPM, Caltrans 2019b) and Workplan Standards Guide Release 12.0 (Caltrans 2018c). Subsequently removed most references to Work Breakdown Structures that were included in the previous document version and apply to non- A-E Services Contracts. • Updated relevant references • Updated formatting and ordering of Appendices for consistency
Chapter 1: • Revised order and content of each subsection to streamline chapter and referred to BMP Evaluation process described in updated Stormwater Management Plan (Caltrans 2016) • Replaced BMP Evaluation Process figure with a revised one planned for a future SWMP update
Chapter 2: • Streamlined chapter to simplify • Moved BMP study types and variable types discussions to new appendix • Revised step titles and order to better align with EPA 2009 Urban Stormwater BMP Performance Monitoring (EPA et al. 2009) and updated Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a) and provided consistency in terminology • Retitled and reordered subsections to make them task oriented (instructional as opposed to topical) • Revised examples of each step to include those from the Roadside Vegetated Treatment Sites (RVTS) study so user may understand entire project step by step • Incorporated applicability to Caltrans Product Evaluation Program
Chapter 3: • Retitled and reordered subsections to make them task oriented (instructional as opposed to topical)
Chapter 4 • No changes other than those listed in “General” above
Chapter 5 • Streamlined chapter to simplify
Chapter 6 • No changes other than those listed in “General” above
Chapter 7 • No changes other than those listed in “General” above
Chapter 8 • Streamlined chapter to simplify
Chapter 9 • Created new Chapter 9 that provides guidelines for developing an approval package for Caltrans Stormwater Advisory Team (SWAT). Included relevant appendices

Chapter 1 Introduction

This BMP Pilot Study Guidance Manual (PSGM) presents procedures to be used by California Department of Transportation (Caltrans) staff and consultants for planning, performing, evaluating, and documenting stormwater Best Management Practice (BMP) pilot studies. The guidance helps develop and test new or improved BMPs that can be used in projects to comply with the Caltrans National Pollutant Discharged Elimination System (NPDES) permit (State Water Board 2017).

Chapter Purpose and Desired Outcome

Chapter Purpose: Describe the intent and content of the PSGM

Desired Outcome: Understand why this manual exists, what it contains, and how it should be used

1.1 Expected Outcomes

Adherence to this PSGM will provide consistency in the monitoring, scoping, development, deployment, and reporting methods implemented among Caltrans BMP pilot projects. The reasons for using consistent procedures for all BMP pilot studies include:

- More efficient execution of BMP studies (i.e., less reinventing of the wheel)
- Fewer mistakes that would lead to inconclusive or erroneous results
- Improved clarity on how to interpret of results
- Proper gathering of information necessary for BMP approval
- Increased confidence in approving appropriate BMPs for implementation

1.2 Contents

This manual is organized according to the primary tasks used to conduct pilot studies:

- Chapter 1, Introduction
- Chapter 2, Project Planning
- Chapter 3, Project Site Selection
- Chapter 4, Permits and Environmental Clearance
- Chapter 5, Project Design
- Chapter 6, Project Construction
- Chapter 7, Operation, Maintenance, and Monitoring
- Chapter 8, Interim and Final Reports
- Chapter 9, Stormwater Advisory Team Package

Several appendices provide supplemental information. Note that there are circumstances when Caltrans staff may deviate from this guidance, but doing so must be approved by the Office Chief for Caltrans Stormwater Program Development.

1.3 Background

1.3.1 Regulatory Basis for Conducting Pilot Studies

The Caltrans NPDES permit (which falls under Section 402 of the Federal Clean Water Act) requires implementation of stormwater BMPs to prevent, capture, and treat runoff from Caltrans properties, ultimately reducing the runoff volumes and pollutants that are discharged to water bodies. The permit, issued by the California State Water Resources Control Board (State Water Board) stipulates that these BMPs must be implemented to the Maximum Extent Practicable (MEP).

Maximum Extent Practical (MEP) as Described in the Caltrans Permit

Compliance with the MEP standard involves applying BMPs that are effective in reducing or eliminating the discharge of pollutants to the waters of the United States. MEP emphasizes pollutant reduction and source control BMPs to prevent pollutants from entering stormwater runoff. MEP may require treatment of the stormwater runoff if it contains pollutants. BMP development is a dynamic process, and the menu of BMPs contained in a SWMP may require changes over time as experience is gained and/or the state of the science and art progresses. MEP is the cumulative effect of implementing, evaluating, and making corresponding changes to a variety of technically appropriate and economically feasible BMPs, ensuring that the most appropriate controls are implemented in the most effective manner. The State Water Board has held that "MEP requires permittees to choose effective BMPs, and to reject applicable BMPs only where other effective BMPs will serve the same purpose, the BMPs would not be technically feasible, or the costs would be prohibitive." (State Water Board 2017).

The permit also requires Caltrans to develop a Stormwater Management Plan (SWMP), and states that, "The Department shall continue to evaluate and investigate new BMPs through pilot studies." (State Water Board 2017).

As further required by the permit, the Caltrans SWMP (Caltrans 2016) lists BMPs that have been evaluated and selected for use on Caltrans properties. Only BMPs that have been approved as described in the SWMP are incorporated into projects. If project conditions prohibit the use of an approved BMP, Caltrans has the option of proposing a non-approved

BMP as a pilot project. In addition, Caltrans may conduct pilot projects to evaluate new or modified BMPs for inclusion in their approved list.

1.3.2 Operational Basis for Conducting Pilot Studies

Before approving BMPs for general deployment in the Caltrans operational environment, information is required on the performance capabilities, technical feasibility, maintenance requirements, and life-cycle costs of the BMPs. To obtain this information, well-designed and carefully monitored pilot studies are conducted to test components of the BMP or the BMP itself before full-scale deployment.

This PSGM augments the Caltrans BMP evaluation process (Figure 1-1 and Figure 1-2) documented in the 2016 SWMP (Caltrans 2016). As stated in the 2016 SWMP, treatment BMPs must be evaluated with respect to the following criteria:

- Feasibility: right-of-way, design, siting, construction, safety, environmental compliance
- Operations and maintenance (O&M)
- Treatment performance
- Life-cycle costs

Guidance on evaluating feasibility, O&M, treatment performance, and life-cycle cost criteria, as listed in the 2016 SWMP, is presented in Chapter 2 of this PSGM. Caltrans Director's Policy 33 (DP-33, Sustainability Policy) and Deputy Directive 107 (DD-107, Use of Life-Cycle Cost Analyses in Project Decision Making) must also be considered and followed. DP-33 directs Caltrans staff to apply sustainability principles in the planning, design, construction, maintenance, and operation of California's multimodal transportation system. DD-107 directs Caltrans staff to use life-cycle cost analyses (LCCAs) to ensure that the costs over the life of a facility are considered when making project decisions.

Note that, as indicated in Figure 1-2, a needs assessment determines that adequate studies are available to demonstrate that the BMP meets Caltrans feasibility, O&M, treatment performance, and life-cycle cost criteria.

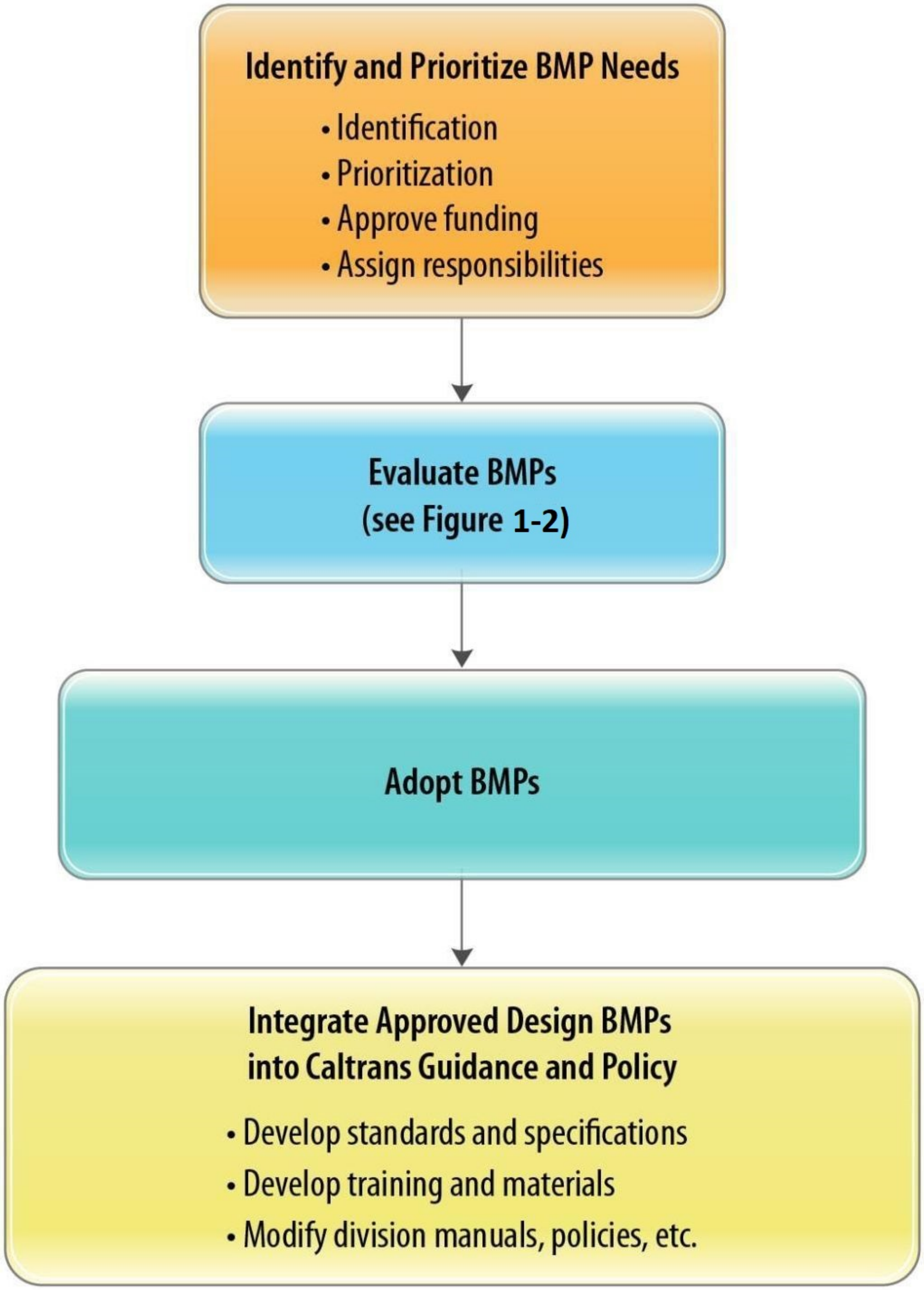


Figure 1-1 BMP Identification, Evaluation, and Integration Process

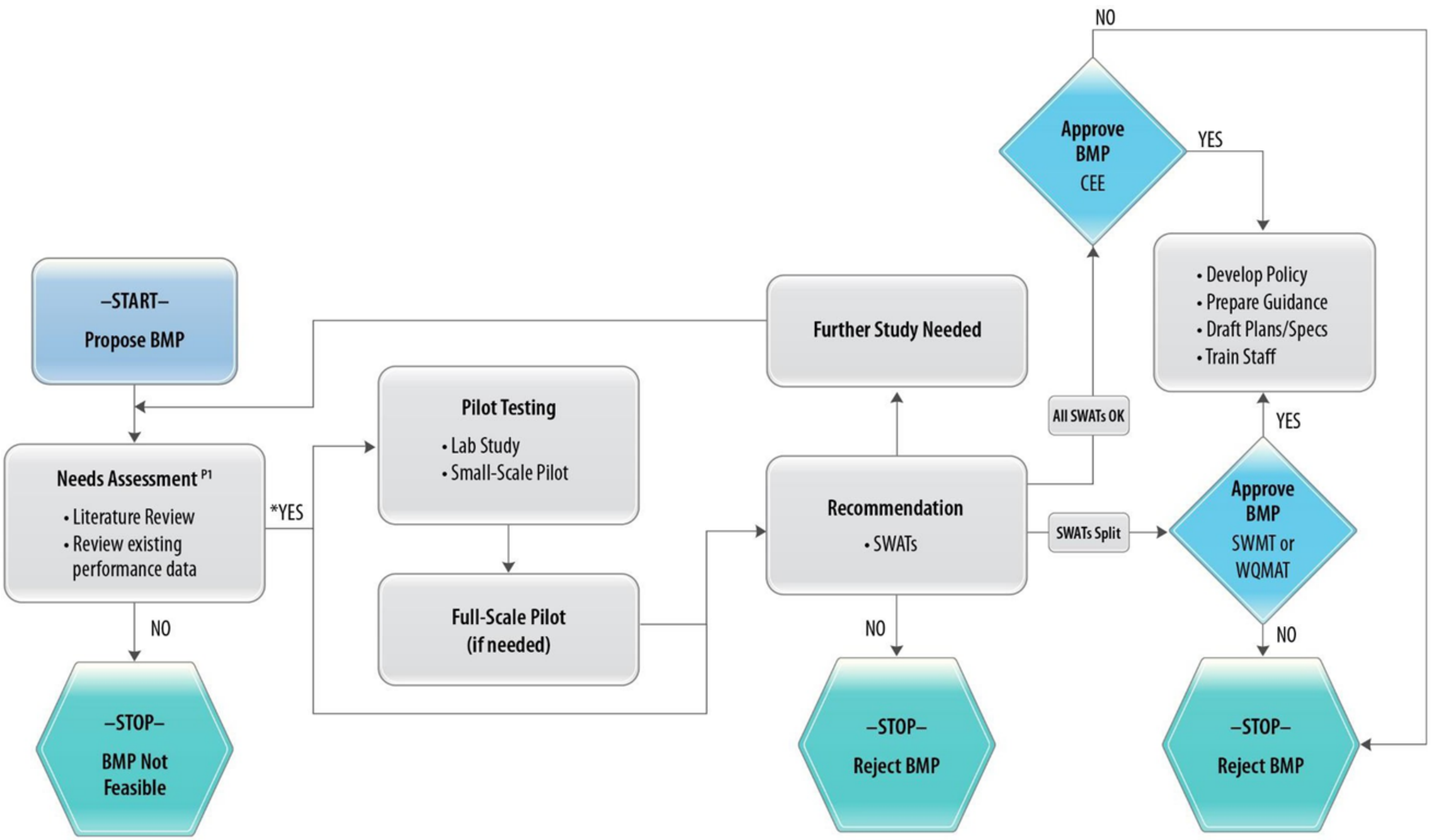


Figure 1-2 Treatment BMP Evaluation Process

1.4 Before Using This Manual

There are a few things that must be done before planning or executing a pilot study:

- Develop the BMP Pilot Project Team (PPT)
- Identify the Relevant Delivery Method

1.4.1 Develop the BMP Pilot Project Team

Before beginning a study, the PPT, which is comprised of subject matter experts, is convened. The experts will include Headquarter and District staff and possibly consultants and academics. The PPT is responsible for ensuring that (1) the study is conducted, documented, and assessed using valid, scientific protocols, and (2) the study results address Caltrans needs. Staff from Caltrans headquarters Division of Environmental Analysis (DEA), Division of Design (DOD), Division of Maintenance, Landscape Architecture Program, and Division of Construction must be part of the PPT to inform and ensure consistency with Caltrans contract management, BMP evaluation, design, O&M, and construction practices. District staff must also be part of the PPT as directed by the District NPDES Coordinators. The District NPDES Coordinators must be involved during the resolution of local project issues and informed throughout the life of the project. Caltrans Chief Environmental Engineer (CEE) may also be involved, as might members of Caltrans Stormwater Management Team (SWMT) and Stormwater Advisory Teams (SWATs). The PPT may also include consultants that can design projects and conduct monitoring and data evaluation as well as academia who can serve as peer reviewers or specific subject matter experts to ensure the study follows scientific protocols and incorporates the latest research available. One Caltrans staff member will serve as the Task Order Manager to coordinate between Caltrans staff, consultants, and academia.

1.4.2 Identify the Relevant Delivery Method

The guidelines presented herein assume that activities will be accomplished using Architectural-Engineering Services Contracts (A-E Contracts). This includes scenarios in which design and construction are performed by the same A-E Consultant or design and construction are performed by different A-E Consultants. For other delivery methods, follow the formal methods used by each district, namely the Caltrans Work Breakdown Structure codes and activities identified in the Project Development Procedures Manual (PDPM, Caltrans 2019b) and the Workplan Standards Guide for the Delivery of Capital Projects, Release 12.0 (Caltrans 2018c).

1.5 How to Use This Manual

This manual is organized according to the various phases of planning and implementing a pilot study. The relevant phases and corresponding chapters are shown in Figure 1-3.

Refer to Appendix A for a list of references mentioned throughout this manual. Refer to Appendix B for lists of acronyms and abbreviations used in this manual.

At the time of issuing this PSGM, Caltrans was updating their website for compliance with the American's with Disabilities Act (ADA). Therefore, many website or webpage links provided in this manual may not direct the reader to the correct file or website. The links will be reconnected once Caltrans has completed the update.

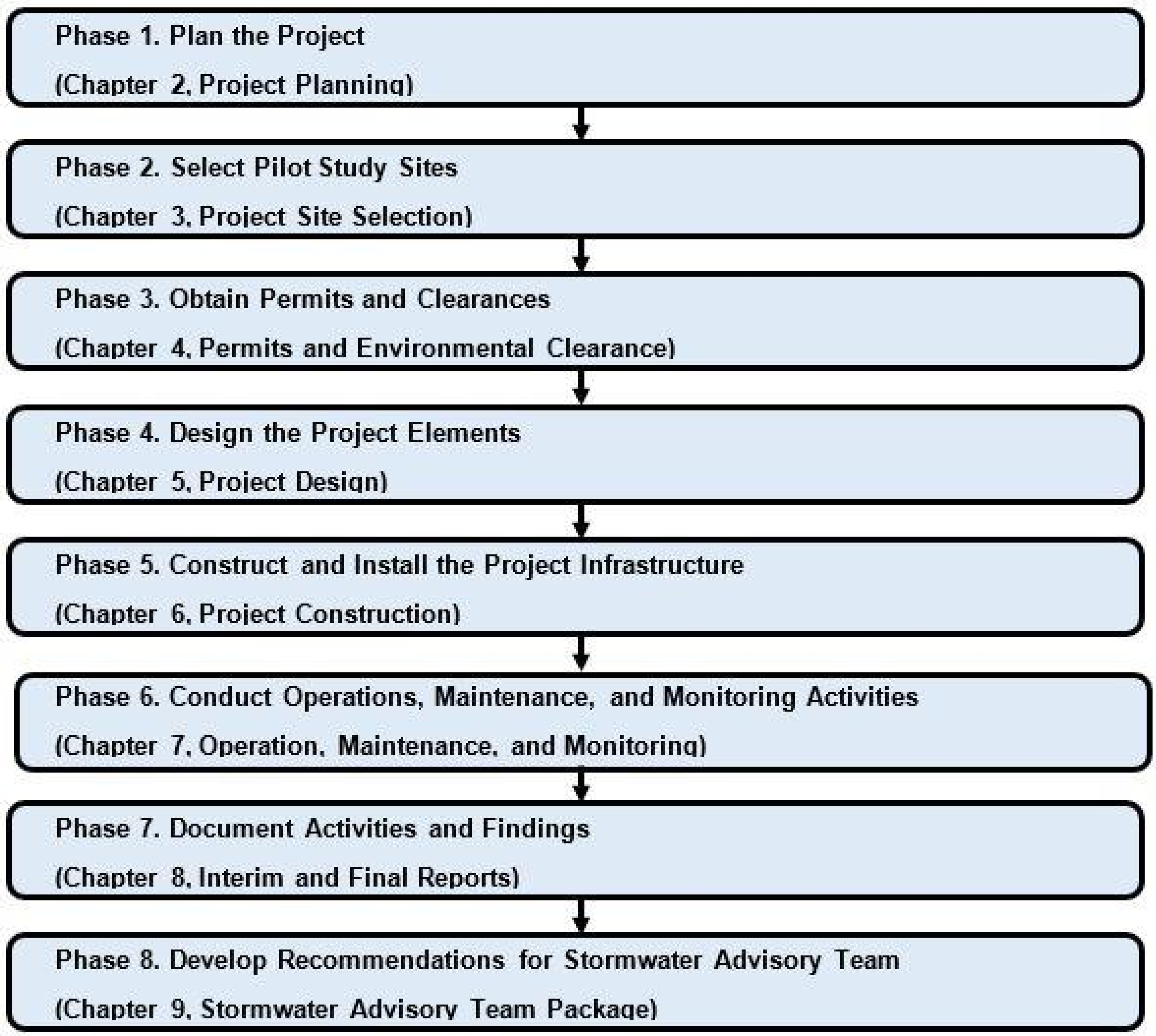


Figure 1-3 Various Phases of Planning and Implementing a Pilot Study

1.6 Examples of Previous Pilot Studies

The guidance provided in this manual has been successfully used to plan, design, and construct pilot BMPs in the Linear Filtration Alternatives Pilot Study and the State Route (SR) 267 Media Filter Pilot Study.

In the Linear Filtration Alternatives Pilot Study, the study plan considered siting, design, installation, and monitoring of seven different BMP concepts. Based on the study plan, Caltrans selected four BMPs for testing. The study plan, together with the site selection technical memorandum and basis of design report that were developed using the study plan, emphasized the original objective of selecting and testing BMPs that could be implemented in relatively poor soils (type C or D). This ensured that the pilot sites selected either were lined or had relatively poor type C soils, and not type A or B soils that are considered suitable for implementation of infiltration trenches. The study plan also identified the need for pretreatment to minimize clogging of the media surfaces and reduce frequency of maintenance. As a result, the final designs incorporated pretreatment at one of the two locations selected, and subsequent monitoring confirmed the importance of pretreatment in reducing maintenance needs.

In the SR 267 Media Filter Pilot Study, the final design report included a section titled “Items Requiring Special Attention” that discussed proper installation of geomembrane liners. Unfortunately, this information was not conveyed to the Resident Engineer and the liners were incorrectly installed below the spillway elevation instead of at the spillway elevation. This increased erosion of fill material beneath the liners, promoted short-circuiting of flows, and increased the risk of baseflow intrusion that had been observed at other sites in the Tahoe Basin. This lesson learned was highlighted in the post-construction report, which discussed the need to emphasize key construction items in the final design report and ensure that these are communicated to the contractor and the Resident Engineer.

Chapter 2 Project Planning

This chapter focuses on planning critical elements of the study, and documenting them in a Study Plan Technical Memorandum (Study Plan TM). Developing the Study Plan is the most critical task in any investigation. A poorly planned study can lead to erroneous conclusions and poor management decisions, resulting in misdirected or wasted time and resources.

Chapter Purpose and Desired Outcome

Chapter Purpose: Describe the steps required to plan a pilot study

Desired Outcome: A Study Plan TM that documents the problem to be addressed, the relevant Caltrans criteria, and how the study will be conducted to address the problem and meet the relevant criteria.

This chapter focuses on studies that assess the feasibility, treatment performance, O&M, and/or life-cycle costs of BMPs in field-scale applications. Other types of studies such as monitoring efforts to characterize discharges, and laboratory or small-scale experiments, also benefit from the creation of study plans. The planning steps described in this chapter should be considered for these projects.

Figure 2-1 shows the study planning process. The steps include:

- Reviewing information to describe the problem and state the study goal (Step 1)
- Identifying the study variables and relevant evaluation criteria (Step 2)
- Developing questions and objectives that state the variables to be studied (Step 3)
- Defining how the study will be conducted to address the study objectives (Step 4)
- Revisiting decisions to ensure the study objectives can be addressed (Step 5)
- Documenting all the details determined in the previous steps (Step 6)

Various members of the PPT will be involved in various steps, with specific roles and involvement assigned by Caltrans staff. Each step is described in the subsequent sections.

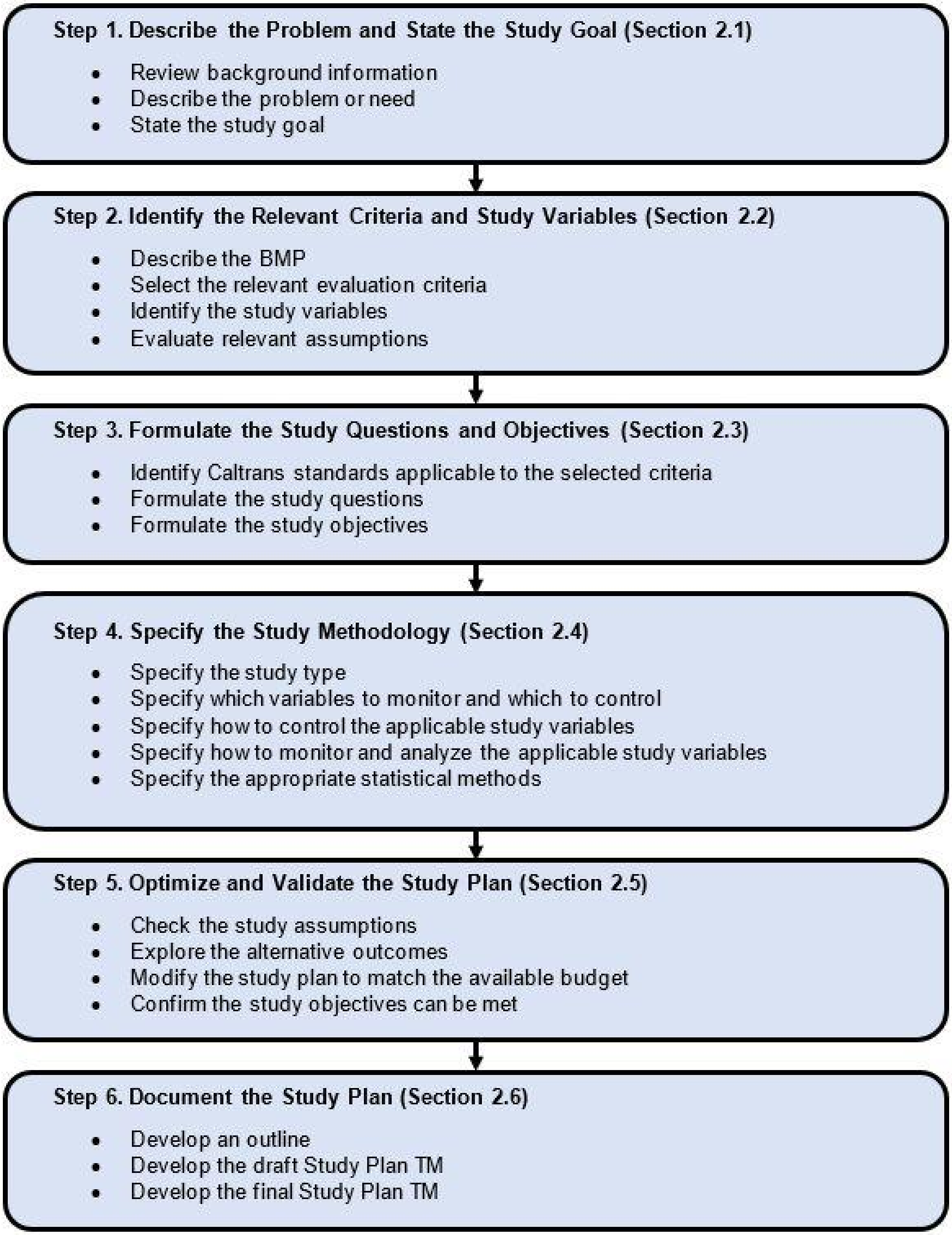


Figure 2-1 Flowchart for Developing a Study Plan

2.1 Step 1: Describe the Problem and State the Study Goal

Step 1 involves reviewing background information to describe the problem and state the study goal. These are used to document and communicate the PPT's understanding of why the new or modified BMP is needed and what the study is trying to achieve. These statements set the foundation for the study and are used throughout all study documents to ensure the study need and goal are addressed. The sub steps are:

- Review background information
- Describe the problem or need
- State the study goal

2.1.1 Review Background Information

As stated in the 2016 SWMP (Caltrans 2016), a needs assessment is conducted by various Caltrans staff before deciding to pursue a BMP pilot study. The assessment will have evaluated relevant performance, O&M, and safety information to confirm the need for the study. The PPT will examine the needs assessment and other relevant literature. Based on its review, the PPT will reevaluate the need for the study. If the PPT encounters serious questions as to whether the BMP will be effective, feasible, and legal, it should inform the Caltrans Task Order Manager who will inform the SWMT that will decide whether to continue, modify, or stop the proposed study. The PPT will also inform the Task Order Manager if the study objectives are already addressed to a sufficient degree in the literature.

2.1.2 Describe the Problem or Need

Develop a series of statements that explain the problem, need, or desire for the study and how it is tied to Caltrans' needs. As cited in the 2016 SWMP (Caltrans 2016), Caltrans develops and evaluates BMPs to address:

- Targeted pollutants that have been identified in Caltrans' stormwater runoff
- Total Maximum Daily Load (TMDL) compliance
- BMP feasibility
- BMP performance trends that identify weaknesses.

Typical problems or needs that may warrant a pilot study include:

- Determining BMP performance, costs, and limitations to approve it for general use;
- Measuring BMP performance to determine its ability to meet specific water quality standards such as those in TMDLs;
- Optimizing design parameters or O&M practices; and/or
- Determining BMP benefits for receiving waters.

As stated in the SWMP, BMP needs may be identified from literature reviews, past or ongoing monitoring, internal brainstorming sessions, feedback from approved BMP implementation, or conditions that arise during development of a project.

Example of Describing the Problem or Need

The Roadside Vegetated Treatment Sites (RVTS) Study

Caltrans needs to reduce the discharge of runoff and associated pollutants from its roadways and facilities. Caltrans has historically included vegetated landscapes in highway design and construction, but these were not specifically engineered to reduce pollutants from highway runoff. Caltrans wants to know if these existing roadside vegetated treatment sites (RVTS) provide water quality treatment similar to Caltrans-approved biostrip BMPs. If similar or better water quality treatment is observed, Caltrans could claim treatment credit for RVTS (which exist in hundreds of locations throughout the state) as it relates to compliance with its permit.

2.1.3 State the Study Goal

Write a statement that reflects the needs assessment and explains overall what the study is trying to achieve. The goal will ultimately be to address the problem or need identified above. Note that the study goal, unlike the study questions and objectives that will be developed in a future step, is general in nature, stating the study aspirations, but without details on criteria, applicable features, or study methods. The goal should be explicitly stated as a sentence that begins with "The goal of the study is to…" This clarifies the overall study intent and helps in developing study questions and objectives and defining how the study will be conducted (later steps). This statement can be used in documents developed throughout the study to ensure that the study goal remains consistent through all phases of the study.

Example of Stating the Study Goal

The RVTS Study

The goal of the study is to determine whether RVTS provide water quality treatment that is similar or better than that provided by Caltrans- approved biostrip BMPs.

2.2 Step 2: Identify the Relevant Criteria and Study Variables

Step 2 involves specifying relevant evaluation criteria and identifying study variables (BMP, site, and precipitation characteristics and O&M practices) that may impact the goal of the pilot study. The sub steps are:

- Describe the BMP
- Select the relevant evaluation criteria
- Identify the study variables
- Evaluate relevant assumptions

2.2.1 Describe the BMP

Identify the applicable treatment mechanisms and key factors influencing the BMP's performance. Section 1.0 of Appendix C of this PSGM tabulates common mechanisms applicable to Caltrans BMP implementation, as well as key factors that affect each mechanism. Then, write a narrative description of the BMP, including the applicable treatment mechanisms and key factors. Also develop a general schematic of the BMP that shows influent, effluent, and bypass/overflow locations and the dominant capture and treatment mechanisms. An example schematic is shown in Figure 2-2. Note that the specific dimensions and locations of monitoring components (although shown in Figure 2-2) are not needed at this point, but will be defined in subsequent steps.

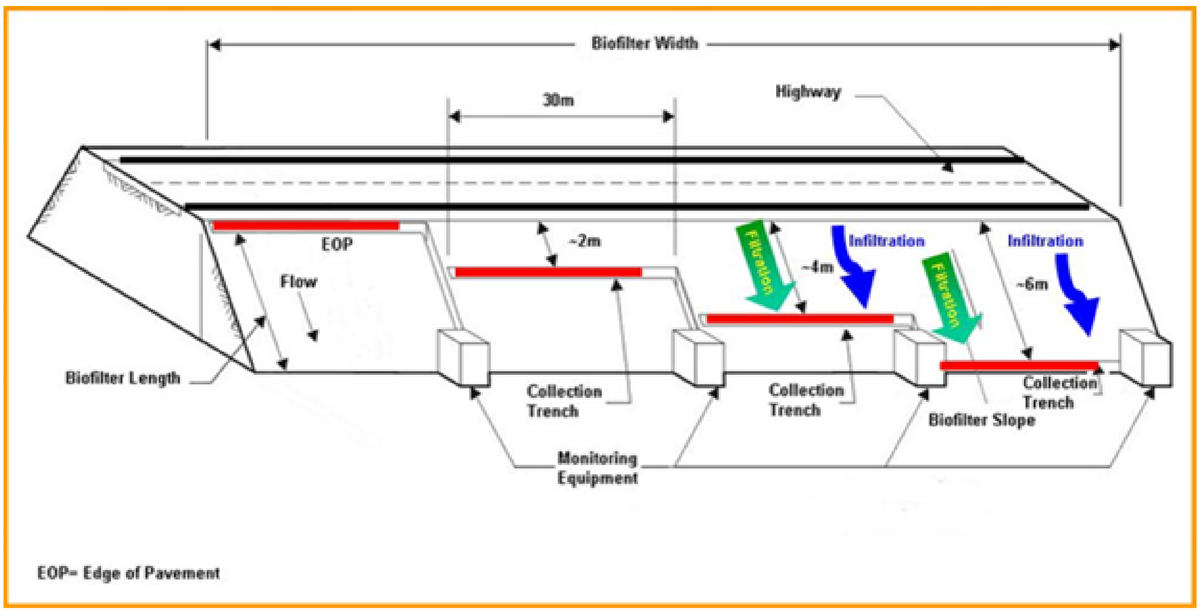


Figure 2-2. Example Schematic: A Roadside Vegetated Treatment Site

Example of Describing the BMP

The RVTS Study

Biostrips (an approved BMP) and RVTS (the proposed BMP) are broad, vegetated surfaces over which runoff flows. For both BMPs, the presence of plants slows the flow, which causes pollutant particles to be removed by settling and filtration. In biostrips, the underlying soils are often replaced with a permeable soil/compost mix, which allows runoff to infiltrate into the ground, thus reducing runoff volumes and pollutant discharges. Soils underlying RVTS may or may not have had an amendment added to improve infiltration. Biostrips are designed to be as long and flat as possible. RVTS have varying lengths and slopes, sometimes steep. In biostrips, pollutants are also reduced through filtration, adsorption, evapotranspiration, plant uptake, and biochemical transformation; it is assumed that the same treatment mechanisms occur within RVTS (although it is unknown if the extent of treatment is similar). RVTS characteristics that may influence the amount of treatment likely include length, width, slope, vegetation density and type, depth of amended soils, and type of soils underlying the amendment. A schematic of a typical RVTS is shown in Figure 2-2.

2.2.2 Select the Relevant BMP Evaluation Criteria

Select the feasibility, O&M, treatment performance, and/or life-cycle cost criteria from Caltrans SWMP (Section 4.2.1 of Caltrans 2016; copied in Section 2.0 of Appendix C of this PSGM for reference) that apply to the study goal and BMP characteristics.

Example of Selecting the Relevant BMP Evaluation Criteria

The RVTS Study

RVTS already exist and are maintained under Caltrans' O&M program. Therefore, only performance criteria were selected; feasibility, O&M, and life-cycle cost criteria were not applicable. The performance criteria, as worded in the SWMP (Caltrans 2016), are:

- Study provides results from full-scale field-testing of a stabilized (erosion-free) post-construction transportation-related impervious drainage area.
- Sampling and analysis were conducted in accordance to the Caltrans Stormwater Monitoring Guidance Manual, or other recognized protocol, such as the Urban Stormwater BMP Performance Monitoring (EPA et al. 2009).
- Testing was conducted at flow rates and volumes typical of Caltrans drainage areas. (Areas vary, usually between 0.1 and 15 acres. Flow and volumes can be found by using Caltrans Basin Sizer).
- Mean (influent) concentrations from study were below the 90th percentile of statewide characterization data.
- Data were collected from at least eight storm events over a minimum period of two years, and demonstrate a statistically significant removal ($p \leq 0.1$), which may require monitoring additional storm events.
- Particle size distribution (PSD) during they study was similar to the proposed field conditions (e.g., state whether traction sand was applied).
- The study's mean removal estimate corroborates the performance claim.
- The rainfall record for the study area or its vicinity during the evaluation period is documented and reported with the study results.

Example of Selecting the Relevant BMP Evaluation Criteria

The RVTS Study

(continued from previous page)

- Not all of the above criteria are met, but at least one of the following apply:
- Statistically significant (p-value ≤ 0.1) constituent removal was established from independent stormwater field monitoring for at least one year
- Removal efficiency based on best professional evaluation of unit operations and processes that are well established for treatment of other waters
- Load reduction of nutrients or BOD due to partial infiltration
- Statistically significant (p-value ≤ 0.1) constituent removal was established from independent laboratory testing that follows the Technology Assessment Protocol – Ecology (TAPE) from Washington State (ECY 2008), and testing used a volume of water equivalent to one year of runoff for a typical installation. Alternatively, a laboratory loading using actual stormwater could be used as with the Tahoe Small Scale

2.2.3 Identify the Study Variables

Use the BMP description, the site and precipitation characteristics, and the O&M considerations to identify the variables that may influence the study goal and applicable criteria. Variables include site and BMP characteristics, such as characteristics and practices that affect the volume and quality of the runoff and BMP effluent. Table 2-1 provides examples. Section 3.0 of Appendix C of this PSGM provides detailed descriptions of these variables.

Table 2-1. Example Study Variables

**Not all variables apply to all studies; select only those variables relevant to the study goal and study criteria*

Runoff Characteristics	Site Characteristics	BMP Characteristics	Precipitation Characteristics	O&M Characteristics
• Influent Volume • Influent Flow rate • Influent quality (constituent types & concentrations)	• Ecoregion • Drainage area • Slope • Traffic volume • Vegetation • Percent imperviousness • Soil type • Soil compaction • Soil infiltration rate • Representative-ness of inflows[1]	• Footprint • Slope • Media characteristics[2] • Ponding depth • Vegetation • Hydraulic loading rate • Effluent flow rate[3] • Effluent volume • Effluent quality (constituent types & concentrations)	• Precipitation intensity & duration • Inter-event timing[4] • Antecedent dry period[5] • Seasonal timing	• Activity type • Activity frequency • Adequacy of practice[6]

[1] Whether the majority of runoff entering the BMP is from Caltrans roadways or facilities, as opposed to commingled flow from other properties or sources

[2] Applies to various media types including soil amendments. Example characteristics include type, grain size, uniformity coefficient, depth, etc.

[3] Orifice sizing or other components controlling effluent flow

[4] Time between monitoring events

[5] Time since the last precipitation, regardless of whether the event was monitored

[6] Whether maintenance was done according to protocols and sufficiently met the established criteria, (e.g., sufficient to prevent an impact to performance, such as short circuiting or internal erosion)

Examples of Identifying Study Variables

The RVTS Study

All study variables listed in Table 2-1 were relevant to the RVTS Study, with the exception of effluent orifice sizing - RVTS do not have orifices. Refer to Section 4.0 of Appendix C of this PSGM as to how these variables can affect performance criteria.

2.2.4 Evaluate Relevant Assumptions

Identify assumptions relevant to the variables identified in the previous step. Examples include assuming that the drainage area estimated from plans is accurate, assuming that influent quality is representative of Caltrans sites, and assuming that infiltration rates can be estimated from soil type. Other common assumptions are listed in Section 4.0 of Appendix C of this PSGM. Some assumptions may not be accurate, and some can result in important variables being overlooked.

Example of Evaluating Relevant Assumptions

The Modified Infiltration Trench Study

Assumption Type	Assumption	Variable(s) Impacted	Measure Taken to Check Assumption
Site	Runoff represents Caltrans runoff & is similar among sites	Influent volume/quality	Select sites with minimal non-Caltrans drainage & with site characteristics typical of Caltrans highways
Site	There is no base flow intrusion	Influent volume/quality	Visually confirm that base flow is not present. (This can be done during site selection phase).
Site	Soil characteristics are as expected	Soil type/ compaction/ infiltration rate	Conduct soil infiltration tests before final site selection and after construction
BMP Design	Influent enters the BMP	Influent flow rate/volume	Review elevations & cross sections to assure flow path to BMP will be unimpeded by vegetated growth, soil uplift, or debris accumulation
BMP Design	Effluent exits the BMP	Effluent flow rate/volume	BMP is designed to drain within 96 hours & construction oversight is conducted to ensure it is built as designed
BMP Operation	Infiltration rates are maintained during entire study	Soil infiltration rate	Inspect for potential clogging at the soil interface throughout the study
BMP Operation	Appropriate O&M practices are done for all BMPs	Infiltration rate, ponding depth, effluent flow rate/volume	Implement an O&M & reporting plan to ensure practices are consistent for all BMPs
Monitoring	Monitoring accounts for most BMP inflows & outflows	Influent volume, effluent volume	Assess by mass balance analysis of monitored events
Monitoring	Construction & other site disturbances do not affect treatment	Influent quality, effluent quality	Do not begin monitoring until site has been stabilized post-construction

Example of Evaluating Relevant Assumptions

The RVTS Study

Originally, the RVTS Study assumed the contributing drainage area could be calculated as the length of the collection troughs times the width of the highway. Further review caught this error, and the highway and right-of-way were surveyed to identify grade breaks to determine the actual drainage area contributing runoff to the collection systems.

2.3 Step 3: Formulate the Study Questions and Objectives

Step 3 involves developing questions and objectives that determine which variables will be studied. The sub steps are:

- Identify Caltrans standards applicable to the selected criteria (from Step 2)
- Formulate the study questions
- Formulate the study objectives

2.3.1 Identify Caltrans Standards Applicable to the Selected Criteria

Identify Caltrans standards that apply to the criteria selected in Step 2. These standards come from Caltrans policies, practices, and requirements. Examples include:

- Caltrans vector requirements, drainage standards, and other safety requirements
- Caltrans O&M practices and maintenance staff expectations, and limitations
- Caltrans practices and expectations for "typical operating conditions"
- Caltrans monitoring requirements (e.g., storm sizes and return frequencies)
- Performance thresholds for specific pollutants of concern
    - Whether performance is to be evaluated based on an annual average, seasonal average, year-by-year, or other time unit
    - Whether performance is to be evaluated for volume, concentration, and/or load, as a percent reduction, nominal reduction, or effluent quality
    - Whether performance should equal or exceed that of existing approved BMPs (these applicable approved BMPs will need to be identified)
- Discharge limitations, downstream beneficial uses, water quality standards, and other considerations or legal restrictions for the pollutants of concern
- Caltrans practices and policies regarding life-cycle cost limitations (e.g., less expensive than a particular approved BMP)

One particular criterion to note is a 96-hour drain time to comply with vector control requirements stated in the Caltrans permit.

Example of Caltrans Standards Applicable to the Selected Criteria

The RVTS Study

The following Caltrans standards will apply to the RVTS Study:

- Storm sizes and return frequencies outlined in Caltrans monitoring protocols
- Number of years and storms per year to be monitored based on desired statistical confidence (Appendix D of this PSGM) and Caltrans monitoring protocols
- Concentration and load reductions observed for biostrips (the approved BMP to which RVTS performance will be compared)

2.3.2 Formulate the Study Questions

Using the BMP description and knowledge of the treatment mechanisms from Step 2 (during sub step 2.2.1, Describe the BMP), as well as the applicable standards, formulate questions that need to be answered to address the study goal and Caltrans criteria. The study questions will be general in nature and merely refer to Caltrans criteria; Caltrans standards, which are used to evaluate the criteria, will be incorporated during development of the study objectives (the next sub step). Questions for the technical feasibility, performance, O&M, and life-cycle cost criteria are listed below. Not all questions may be necessary to achieve the study goal, but each should be considered.

Example of Formulating the Study Questions

The RVTS Study

- What degree of treatment do RVTS provide?
- How does RVTS treatment compare to that for biostrips?
- How do variables affect the treatment performance of existing roadside vegetated slopes?
- Is there a maximum slope at which treatment is no longer comparable to that from biostrips?
- Is there a minimum strip width at which treatment is no longer comparable?
- Is there a minimum vegetation density for which treatment is no longer comparable?

2.3.3 Formulate the Study Objectives

For each study question, create statements that describe (1) what the study will assess, and (2) the Caltrans standards against which the study results will be evaluated. The objectives should be specific, measurable, achievable, relevant, and time bound so that the focus of the study and how its results will address the overall problem, study goal, and study questions are clear and unambiguous and directly tied to Departmental needs.

Example Study Objectives

The RVTS Study

- Determine if RVTS provide 50% or greater volume, concentration, and load reductions (the performance thresholds) for standard constituents cited in the Caltrans permit.
- Determine if the volume, concentration, and load percent reductions and effluent quality for standard constituents cited in the Caltrans permit are statistically similar, at the 90% confidence level (see Appendix D of this PSGM), for RVTS and biostrips specifically engineered for water quality performance.
- Determine how the volume, concentration, and load percent reductions vary as a function of the physical and hydrologic factors slope, width, drainage area, infiltration rate, influent concentration, and degree of vegetative cover.
- Determine if there is a maximum slope at which volume, concentration, and load percent reductions are not comparable to that of biofilters.
- Determine if there is a minimum strip width at which volume, concentration, and load percent reductions are not comparable to that of biofilters.
- Determine if there is a minimum degree of vegetated cover at which volume, concentration, and load reductions are not comparable to that of biofilters.

2.4 Step 4: Specify the Study Methodology

Step 4 involves defining how the study will be conducted to address the study objectives. This includes deciding which variables will be monitored or controlled and how. The sub-steps are:

- Specify the study type
- Specify which variables to monitor and which to control
- Specify how to control the applicable study variables
- Specify how to monitor and analyze the applicable study variables
- Specify the appropriate statistical methods

2.4.1 Specify the Study Type

Select a study type from the list in Table 2-2. Section 5.0 of Appendix C of this PSGM provides specific applications.

Table 2-2. Study Types

Study Types	Description
Influent-Effluent	• Compares runoff entering and exiting a BMP or a series of BMPs (i.e., a treatment train)
Upstream-Downstream	• Compares data from in-stream locations upstream and downstream of a BMP
Paired Watersheds	• Compares data from two or more similar watersheds during concurrent periods
Before-After	• Compares data from a location before and after a change is made (e.g., a BMP is implemented/modified)

Examples of Study Types

The Influent-Effluent Approach

This approach was used in the Retrofit Pilot Study. Influent and effluent sampling was performed on sand filters, biostrips, swales, dry and wet detention basins, and oil/water separators. This approach works best when there are discrete inflows to and outflows from a treatment system.

The Upstream-Downstream Approach

Caltrans used this approach for the Small Streams Crossing Study. Samples from upstream and downstream of bridge crossings were collected, analyzed, and compared.

The Before-After Approach

This approach was used in the Fresno Public Education Study. Litter was collected before and after an extensive public education program (" Don't Trash California") to determine if the program influenced public behavior and to what degree public behavior affected water quality.

The Paired Watersheds Approach

The Drain Inlet Cleaning Efficacy (DICE) Study used this approach. Water quality was measured downstream of six drain inlets that were cleaned once per year and six other drain inlets that were not cleaned. Midway through the study, the uncleaned inlets were cleaned and the cleaned inlets were not. The effectiveness of the BMP (drain inlet cleaning) was determined by comparing the water quality from cleaned inlets with that from the uncleaned inlets.

Example of Specifying the Study Type

The RVTS Study

The RVTS study will use an influent-effluent approach. One station that collects runoff from the edge of pavement will be located at each site (multiple sites throughout the state) to represent influent water quality and volumes. Other stations will be placed at varying widths within the RVTS to collect runoff, representing effluent water quality and volumes for that width.

2.4.2 Specify Which Variables to Monitor and Which to Control

Determine which study variables identified in Step 2 will be monitored and which will need to be fixed or controlled in order to minimize their potential impacts on data variability.

Typical variables to be monitored or recorded are listed in Section 6.0 of Appendix C of this PSGM.

Example of Specifying How the Study Variables Will be Treated

The RVTS Study

Monitored Variables	Controlled Variables
• Influent quality (constituent concentrations) • Influent & effluent runoff volumes & flow rates • Precipitation type, intensity, & duration • Antecedent dry period • Vegetation density	• Drainage area & slope • Precipitation probability of occurrence & estimated depth • Width • Traffic volume • Soil type & compaction • Representativeness of inflows • Seasonal & inter-event timing • O&M practices • Ecoregion and climate

2.4.3 Specify How to Control the Applicable Study Variables

For the selected variables to be controlled, determine how this will be done. This will inform site selection and design, as well as monitoring protocols. For example, “representativeness of inflows” can be controlled by establishing a threshold for percentage of non-Caltrans runoff as no more than 5%. Variables such as traffic and vegetation density can be controlled by establishing values or ranges for average annual daily traffic (AADT) and percent vegetation density, respectively. Alternatively, they can be controlled by selecting a number of sites that have a range of traffic volumes and vegetative densities. Key variables that cannot be fixed or constrained should be monitored and recorded so that their impact on the study results can be discerned at the end of the study. Section 7.0 of Appendix C of this PSGM describes how to control study variables.

Example of Specifying How to Control the Relevant Study Variables

The RVTS Study

- *Drainage area:* select sites with a 1-acre drainage area or less
- *Width:* select multiple sites throughout the state with varying widths, from 1 to ~10 m
- *Slope:* select multiple sites with varying slopes, from relatively flat to steep (2:1, H:V)
- *Vegetation density:* select multiple sites with varying density, from poor to full coverage
- *Traffic volume:* select multiple sites throughout the state
- *Soil type and compaction:* select multiple sites throughout the state
- *Inflow representativeness:* select sites where 100% of runoff is from Caltrans properties
- Probability of precipitation and estimated depth: use Caltrans monitoring protocols
- *Seasonal timing:* conduct monitoring between October 15 and April 15
- *Inter-event timing*: use standards identified in Caltrans monitoring protocols
- *O&M practices:* select multiple sites throughout the state
- *Ecoregion and climate:* select multiple sites throughout the state

2.4.4 Specify How to Monitor and Analyze the Applicable Study Variables

For the study variables to be monitored or recorded, specify how this will be done. Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a) describes various Caltrans sampling techniques, as well as analytical methods. Specific procedures for sample collection and data quality assurance will be developed and specified in the OM&M Plan (see **Chapter 7**).

2.4.5 Specify the Appropriate Statistical Methods

Specify the statistical methods needed for study planning and data analysis. Appendix D describes common statistical methods used for BMP studies, including selection of an appropriate method, interpretation of results, and method limitations. Table 2-3 lists the various topics covered in Appendix D. Appendix D describes the statistical analysis procedures for the interim and final reports.

Table 2-3. Statistical Topics Covered in Appendix D

Section	Title	Typical Study Questions Addressed
D1	How to Estimate an Adequate Number of Samples	• How many samples would I need to achieve desired confidence in the conclusions? • After one or two years of sampling, how do I decide whether I need more samples?
D2	How to Examine Data Quality and Detect Possible Outliers in the Data	• How do I prepare graphical and numerical data summaries to understand salient data features and identify potential outliers?
D3	How to Examine Data Quality in the Presence of Non-detect Values	• How do I account for non-detect results?
D4	How to Verify Common Assumptions for the Selection of an Appropriate Statistical Test	• How do I verify whether data are normally distributed? • How do I verify that the data variability of two or more groups is similar?
D5	How to Estimate Probabilities Using Data for a Single Variable	• How do I estimate how often the average BMP effluent concentration would meet a legal limit? • How do I estimate the BMP percentage removal of a pollutant with a specified confidence level?
D6	How to Compare Two Independent Data Sets	• In an upstream-downstream watershed approach or paired watersheds approach, how do I decide whether a given BMP is effective in removing a pollutant? • How do I compare the effectiveness of two pilot BMPs at a given geographic location?
D7	How to Compare Two Paired Data Sets	• In an influent-effluent approach or before-after approach, how do I decide whether a given BMP is effective in removing a pollutant?
D8	How to Compare Three or More Independent Data Sets	• How do I compare the effectiveness of three or more pilot BMPs at a given geographic location?
D9	How to Develop a Linear Regression Equation	• How does BMP effectiveness vary as a function of such other factors as storm characteristics, BMP design variables, and O&M practices?
D10	How to Evaluate Time Trends in BMP Monitoring Data	• How can I tell if the effectiveness of my pilot BMP is changing over time?
D11	How to Compare and Group BMPs	• How do I evaluate whether the treatment performance of BMPs is significantly different when each BMP type has been monitored at multiple locations?

2.5 Step 5: Optimize and Validate the Study Plan

Step 5 involves revisiting the study components to ensure the study objectives can be addressed within the allotted budget and in line with Caltrans feasibility, performance, O&M, and life-cycle cost criteria. The sub steps are:

- Check the study assumptions
- Explore the alternative outcomes
- Modify the study to the match the available budget
- Confirm the study objectives can be met

2.5.1 Check the Study Assumptions

Review the study details (e.g., the variables to be monitored and controlled, how variables will be controlled, the study type, applicable criteria), specifically looking for incorrect or inaccurate assumptions. This may include looking for assumed relationships between variables that are not true, questioning whether variables assumed to be negligible actually are, and rechecking for assumptions that can hide variables (from Step 2). If any issues are identified, revise the study as appropriate.

Example of Checking the Study Assumptions

The SR-73 Pilot Detention Basin Study

In this study, the Department needed guidance for designing and deploying detention basins in locations with insufficient space to capture the required water quality volume.

One study question was: “How does treatment performance vary with basin volume?” It was initially assumed that larger basins would hold water longer. Because rainfall intensity patterns and the size of the outlet orifice affect detention time, the relationship between basin size and detention time is not direct. For example, to drain a large detention basin in 48 hours (to prevent mosquito breeding) requires an outlet orifice of a certain size to be provided. However, this orifice may be so large that runoff from small storms (which make up the majority of annual rainfall events) will pass very quickly through even a large basin. In this example, the study question cannot be fully answered unless comparisons are made between sites that have similar outlet orifice sizes relative to their drainage areas, and storms with similar rainfall intensity patterns are monitored.

2.5.2 Validate the Study by Exploring the Alternative Outcomes

Check whether the study results will effectively address the study objectives, particularly as they relate to a Department need (i.e., the problem and study goal). To do this, the

PPT should explore a range of potential alternative outcomes to each study question. For each alternative outcome, an action item should be identified. If the PPT identifies situations where (1) alternative outcomes to a study have the same or no impact on the Department need, or (2) alternative outcomes are inconclusive, the study elements should be revisited and revised or augmented. One possible outcome of this process is a decision to cancel the study because it does not meet Department needs.

Example of Alternative Outcomes Analysis:

The RVTS Study

Below are examples of alternative outcomes for one RVTS study objective:

Objective: Determine if existing RVTS provide 50% or greater volume, concentration, and load reductions (the performance thresholds) for standard constituents cited in Caltrans permit.

- Outcome 1: For every constituent measured, RVTS provides 50% or greater volume, concentration, and load reductions.
- Outcome 2: For some of the constituents measured, RVTS provides 50% or greater volume, concentration, and load reductions.
- Outcome 3: For some of the constituents measured, RVTS provides less than 50% volume, concentration, or load reductions.
- Outcome 4: For every constituent measured, RVTS provides less than 50% volume, concentration, or load reductions.

2.5.3 Modify the Study to Match Available Budget

Develop a cost estimate and modify various elements of the study to fit the available budget. Be aware that modifying the study scope will usually affect the ability of the study to meet its original objectives. In considering various modifications, revisit the study objectives and revise them as needed.

2.5.4 Validate the Study Plan by Confirming the Study Objectives Can be Met

Describe how each study objective will be met in terms of the study plan elements. This is done to confirm that the study can provide the information needed to address the problem and study goal and meet Caltrans BMP evaluation criteria. If the study plan does not adequately address an objective, adjust the study elements as needed.

For example, consider the study objective "Determine if the frequency for maintaining the filter surface of a proposed BMP is annual or less." To confirm that it will be possible to address this objective, explain how the Study Plan requires observing and recording water level data (head) to indicate clogging of the filter media and appropriate O&M.

2.6 Step 6: Document the Study Plan

Step 6 involves documenting all the details determined in the previous steps to inform and direct the next phases of the BMP pilot study (permitting, design, construction, monitoring, assessment, and reporting). The sub steps are:

- Create an outline
- Develop a draft Study Plan TM
- Finalize the Study Plan TM

Well-documented Study Plans help ensure that the overall validity of the study is maintained during its execution. Additionally, the Study Plan will serve as a resource for making a final determination about whether the results achieved the study goal and objectives.

2.6.1 Create an Outline

Develop an outline that will guide you in organizing and documenting the details regarding the study problem/need, goal, questions, objectives, variables, and methodology. Appendix E provides an example, annotated outline that includes:

- Introduction (Section 1.0 of Appendix E)
- Problem Description and Study Goals (Section 2.0 of Appendix E)
- Study Questions and Objectives (Section 3.0 of Appendix E)
- Study Methodology and Analytical Approach (Section 4.0 of Appendix E)
- Schedule and Cost (Section 5.0 of Appendix E)
- Constraints and Optimization (Section 6.0 of Appendix E)
- Reporting Requirements (Section 7.0 of Appendix E)
- References (Section 8.0 of Appendix E)

2.6.2 Develop a Draft Study Plan TM

Using the outline, document the details, reasoning, and decisions relevant to each step. The Study Plan should describe how the study will be conducted and why various decisions were made. The draft should be routed to the PPT as well as Caltrans staff for review and comment.

2.6.3 Finalize the Study Plan TM

Upon receiving comments, address them as appropriate. If needed, follow up with reviewers to clarify or coordinate how comments can be addressed. Upon addressing all comments, finalize and distribute the Study Plan TM. A good example of a Study Plan TM that follows the steps described in this PSGM is Caltrans Final Linear Filtration Alternatives Pilot Study Study Plan Technical Memorandum (Caltrans 2007b).

Chapter 3 Project Site Selection

This chapter presents the procedures and guidelines for selecting pilot study sites. The steps are shown in Figure 3-1.

Chapter Purpose and Desired Outcome

Chapter Purpose: Describe the steps and considerations for selecting sites for the pilot study

Desired Outcome: Sites where BMPs can be installed/modified and evaluated according to the objectives documented in the Study Plan TM (Chapter 2)

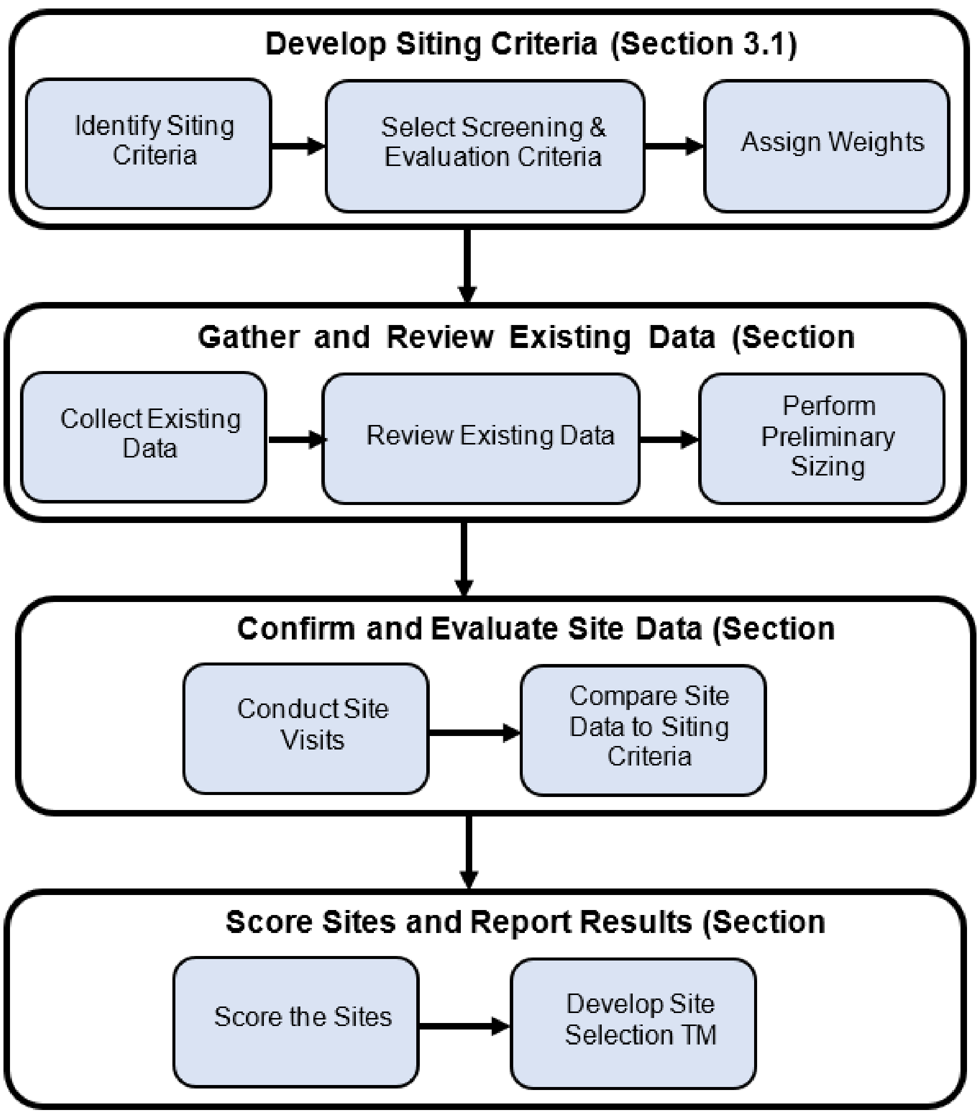


Figure 3-1. Site Selection Process Flowchart

3.1 Develop Siting Criteria

Using the pilot study objectives and key parameters documented in the Study Plan TM, along with any pilot BMP design criteria established after completion of the Study Plan TM, a list of ideal site characteristics (i.e., siting criteria) is prepared to evaluate and compare candidate sites. This activity encompasses three tasks: identification of appropriate siting criteria; identification of screening versus evaluation criteria; and assignment of weighting factors that are used to prioritize key siting criteria.

3.1.1 Identify Siting Criteria

Depending on the specific objectives and pilot BMP under evaluation, the list of desired site characteristics may include as many as 20 criteria. When developing the list, it is important to use selection criteria that are quantifiable in nature to avoid subjective evaluation and facilitate comparison of candidate sites. For example, the criterion "Available Hydraulic Head" would be better than "Minimum Hydraulic Head Met?" because the former would allow differentiation between sites that met the minimum hydraulic head requirement. Functionally dependent selection criteria should also be avoided to preclude double-counting bias in the results. For example, both "Hydrologic Soil Group (A, B, C, D)" and "Soil Infiltration Rate" would not be used because the two criteria are not mutually exclusive (hydrologic soil groups A and B imply a higher infiltration rate than hydrologic soil groups C and D).

For BMP pilot studies, site-selection criteria may be divided into four categories: BMP design, monitoring, safety and implementation. Each of these categories is discussed below. Additional information on site selection considerations is provided in the Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a).

BMP Design Criteria

BMP design criteria are BMP and site characteristics that are important for the successful operation of the pilot facility. Actual criteria used will depend on the technology being tested, but the more common ones include:

- *Drainage area:* Some pilot devices may need to be designed for a specific range of tributary drainage areas, such as multi-chamber treatment trains, which are typically designed for drainage areas no more than 2.5 acres, or 1 hectare, in size.
- *Hydraulic head:* Some devices may require a minimum amount of head to operate by gravity, such as the inclined screen gross solids removal device, which requires at least 5.5 feet (1.68 m) of hydraulic head. In addition to the head needed for the BMP itself, certain flow measuring devices (e.g., H-Flumes) are not designed to operate under submerged conditions and require additional hydraulic head to ensure no backwater conditions within the device.

- *Design storm:* Pilot study objectives might include evaluating a volume-based BMP designed to capture a specific water quality volume (WQV). In most cases, the WQV will be calculated from the 85th percentile design storm as specified in the Caltrans permit. Sometimes smaller WQVs are evaluated.
- *Water Quality Flow (WQF):* Pilot study objectives might include evaluating a flow-based BMP that requires a specific WQF.
- *Power availability:* Some devices may require electrical power to operate specific components, such as some chemically enhanced detention basins that require electrical power to operate the chemical dosing systems.
- *Soil classification and infiltration rates*: Pilot studies that incorporate infiltration as a removal mechanism will require certain soil types or measured infiltration rates. For example, infiltration trenches require hydrologic soil group type “A” or “B” soils. Some BMPs may require a minimum (or maximum) infiltration rate, such as infiltration basins, which require an infiltration rate between 0.5 inches (12.5 mm) per hour and 2.5 inches (60 mm) per hour.
- *Site geometry*: Some devices may require specific shapes (e.g., length to width ratios) for optimal performance, such as detention basin devices for which rectangular shaped basins are preferred over square shaped basins.
- *BMP Footprint:* Studies will require a certain amount of space based on preliminary hydrologic/hydraulic and sizing calculations (such as concrete vault Austin sand filters, which require at least 36 feet by 84 feet of space for 10,000 cubic feet of WQV). This is different from site geometry, in that a site may have adequate overall space but not be in the desired shape. Apart from the space required for the BMP itself, the site must have adequate room for all monitoring equipment (e.g., flumes, samplers, enclosures). More importantly, sufficient distance must be available upstream from certain flow measuring devices (e.g., some flumes require the influent flow in-line with the flume for a length equivalent to 12 times the pipe diameter) to establish uniform flow conditions if these conditions do not already exist.
- *Depth to groundwater:* A minimum depth to seasonally high groundwater is typically required for some BMPs (such as infiltration devices), and preferred for all BMPs to avoid groundwater contamination as well as construction cost escalations (dewatering, structure anti-buoyancy, etc.).
- *Traffic*: Some pilot studies may require sites to be located in an area with specific AADT requirements. AADT is one of the most important factors affecting runoff quality, so sites with unrepresentative AADT should be avoided to ensure influent quality is representative of Caltrans sites.

- *Stormwater sources:* Typically, sites are preferred where the majority of the runoff is wet weather runoff from impervious (paved) Caltrans rights-of-way. Sites with significant dry weather flows, significant non-Caltrans runoff, or significant runoff from pervious areas should be avoided.
- *Existing BMPs*: Some sites may have existing structural or non-structural BMPs that can impact monitoring results. Examples of existing BMPs include roadways that undergo street sweeping or have open graded friction course (OGFC).

Monitoring Criteria

Monitoring criteria are characteristics necessary for the successful monitoring of the pilot BMP under evaluation. These typically include:

- *Number of inlet pipes/streams*: Monitoring pilot performance requires accurate characterization of the runoff discharging into the device (i.e., before-treatment), and a single inlet stream is preferred. Sites with multiple inlet streams either require costly drainage modifications to combine multiple streams or costly monitoring programs to accurately characterize the influent runoff.
- *Proximity to next closest or paired site*: As part of an “Influent and Effluent” study, if one of the pilot study objectives is to compare the performance of a device at multiple sites with similar runoff characteristics, then it is typically preferred to have the sites close together such that the runoff at each site may be considered similar. As part of a paired watersheds study, the distance between the paired sites is important to facilitate monitoring activities.

Example Monitoring Criteria - Site Proximity

The State Route 73 Pilot Project

Although the existing compost stormwater filter BMP at SR 73 and Newport Coast Drive had multiple inlet pipes, it was selected as a retrofit site for one of the SR 73 BMP pilots due to its other characteristics and proximity to other pilot sites.

Safety Criteria

Safety criteria are characteristics necessary for the safety of the PPT (particularly the monitoring staff) and the public. When evaluating these characteristics, coordinate with the District traffic operations unit either directly or by the District NDPES Coordinator. Such communications must be documented and shared with the Department Task Order Manager. Safety criteria typically include:

- *Distance to Edge of Traveled Way*: Caltrans policy requires that roadways include a traversable clear recovery area to provide a clear recovery zone (CRZ) for vehicles that have left the traveled way. Obstacles located in the recovery zone should be removed, relocated, or shielded (e.g., with guardrail or crash cushions). Pilot project components that would be considered an obstacle include flumes, sampler enclosures, and any above-grade concrete structure. Selecting sites with sufficient space to locate the pilot BMP outside the recovery zone will reduce construction costs associated with shielding the BMP and coordinating with Traffic Safety. Specific information on this issue may be found on the Caltrans traffic operations webpage (https://dot.ca.gov/programs/traffic-operations).

- *Access*: The pilot site must be accessible to construction and monitoring personnel without risk of injury. Sites that have access from a non-Caltrans roadway are ideal, but are not usually possible. Sites at which activities occur in roadway medians (i.e., between opposing lanes of traffic) are not safe and should be avoided.

Implementation Criteria

Implementation criteria are characteristics that can have a direct effect on the project's ability to meet its objectives (primarily related to schedule and cost). These criteria are sometimes referred to as implementation issues or site constraints, and may include:

- *Sufficient right-of-way:* Acquisition of rights-of-way is time consuming and costly, and should be avoided. Selection of sites with adequate rights-of-way is critical to meeting project schedules and budgets.
- *Base flow:* Candidate sites in which permanent sources of runoff (base flow) exist may or may not be desirable depending on the pilot device being tested. For example, base flow would be desirable for wet basins, but not for a device that is designed to be dry between storm events. Furthermore, base flow may have an undesirable impact on monitoring as the flow being recorded reflects more than just the storm event.
- *Conflicts with high-risk utilities*: Sites with high-risk utilities present should be avoided, as they require utility coordination and relocation/adjustment/protection, affecting both schedule and budget. Refer to the Caltrans PDPM (Caltrans 2019b) for what constitutes high-risk utilities and how to address conflicts.
- *Conflicts with low-risk utilities:* Although not as critical as high-risk utility conflicts, low-risk utility conflicts will result in some utility coordination, relocation, adjustment, and/or protection and should be avoided if possible. Refer to the PDPM (Caltrans 2019b) for what constitutes low-risk utilities and how to address conflicts.
- *Environmental impacts:* Pilot study projects with environmental impacts to waterways, biological, cultural, and other protected resources will require agency

coordination and may not qualify for a Categorical Exemption (CE). If this is the case, environmental documents (EDs) (e.g., environmental impact reports [EIRs], environmental assessments [EAs], environmental impact statements [EISs]) would be required before project approval, significantly delaying the project schedule (by one year or more). Sites with environmental issues should be eliminated from consideration.

- *Line-of-Sight Visibility:* Depending on the area in question and the BMP, visibility of the BMP from nearby residential areas or recreation areas may be a concern. The distance from the nearest visual receptacle to the site may be used as a method for assessing this criterion. Review the Highway Design Manual (HDM, Caltrans 2018e) for additional line-of-sight criteria.
- *Conflicts with other construction:* Unless it is desired to construct the pilot project under the contract change order (CCO) delivery method, sites with either ongoing or proposed (within the pilot project schedule) construction should be avoided. Consult Caltrans website for a list of ongoing contracts (available on their construction website at www.dot.ca.gov/hq/construc/statement.html) and the local District NPDES Coordinator to identify future projects in the area.
- *Completeness of Data:* Some sites may have sufficient data to proceed directly to design, while others may require additional information to be collected.

Example of Implementation Criteria - Utilities

Site Selection for the District 3 Chemically Enhanced Detention Basin BMP Pilot Study included the evaluation of 26 candidate sites. Although one of the sites had the highest overall score, it was rejected from future consideration because it had a potential conflict with a high-risk utility.

$$$ Cost Reduction Strategy $$$

Perform site selection and assessment activities to avoid hidden costs associated with obstructions such as utility conflicts and buried objects.

3.1.2 Select Screening and Evaluation Criteria

Once the list of siting criteria has been established, the next step is to identify which criteria will be used for site screening and which will be used for site evaluation. Screening criteria are used as an initial screening tool to determine if a site has a specific characteristic that automatically preclude it from serving as a pilot site (i.e., fatal flaw). Evaluation criteria are then used to compare candidate sites that passed the screening criteria. This approach is valuable when there are a large number of candidate sites and the schedule/budget does not allow a detailed evaluation of all sites.

Screening criteria may be any of the three types of criteria described above. For example, if the pilot BMP under evaluation requires 3 feet (900 mm) of hydraulic head to function by gravity, and one of the candidate sites has a maximum available hydraulic head of 1 foot (300 mm), the site would be considered unsuitable. Alternatively, if one of the candidate sites does not have enough space for the BMP and requires acquisition of additional right-of-way, it might be rejected, or, if one of the candidate sites is in the middle of protected Coastal Sage Scrub habitat, the site might be eliminated from further consideration due to the environmental impact and necessary agency coordination. Table 3.1 presents an example list of siting criteria that may be used for screening purposes. For each criterion, the desired value that would result in the candidate site passing the screening phase is given.

All screening criteria should be worded such that the same response (either Yes or No) represents the same meaning throughout. With this approach, interpretation of the results is more straightforward.

Table 3-1 Example Screening Criteria for Site Selection

Screening Criterion	Desired Value	Passing Value
Sufficient Hydraulic Head	4 ft	Yes
Sufficient Water Quality Volume	5000 cu-ft	Yes
Sufficient Right-of-Way	20 ft	Yes
No Conflicts with Other Construction	Yes	Yes
No Environmental Issues	Yes	Yes
No High-Risk Utility Conflicts	Yes	Yes
Sufficient Soil Infiltration Rates	1 in/hr	Yes
Sufficient Depth to Groundwater	10 ft	Yes

Note that some siting criteria may be used for both screening and evaluation. For example, hydraulic head may be used as a screening criterion if the BMP requires a certain amount of head to operate by gravity, and it may also be used as an evaluation criterion to differentiate sites that passed the screening criteria (e.g., sites with more hydraulic head are scored higher than sites with less hydraulic head).

3.1.3 <u>Assign Weights</u>

When reviewing the final list of evaluation siting criteria, some criteria might be considered more important to the study objectives than others. For example, it is more important to know how much hydraulic head there than knowing if the site will encroach upon the recovery zone. Alternatively, hit is more important to know if a site has high-risk utilities than if it has low-risk utilities.

To account for and control the relative importance of evaluation siting criteria, a weighting factor is assigned to each criterion. The higher the weighting factor the more important the criterion. Weighting factors assigned to an evaluation criterion that was also used as a screening criterion may be lower, because once a site has passed the screening phase, such criteria may not be as important as other evaluation criteria. For example, available head is often used as a screening criterion because certain BMPs have a minimum head requirement. If a specific site meets that minimum requirement, the actual amount of available head may not significantly influence the ranking of the site. In that case, the hydraulic head evaluation criterion would be given a lower weighting factor. Weighting factors for pilot study siting activities are presented in Table 3.2. The actual assignment of weighting factors is somewhat subjective. As a result, it is important to obtain Caltrans' review and approval before scoring the evaluation siting criteria (see Section 3.4.1).

Table 3-2 Evaluation Siting Criteria Weighting Factors

Value	Relative Importance
1	Not Very Important
2	Somewhat Important
3	Important
4	Very Important

Example of Developing Siting Criteria

A pilot study being conducted in the Lake Tahoe area will evaluate the performance of detention basins enhanced by chemical coagulant additions. Example screening and evaluation criteria and weighting factors are tabulated below. Some criteria (e.g., hydraulic head) are used for both screening and evaluation purposes. See Section 3.4.1 for assessing and scoring the criteria.

Criterion	Screening Criterion	Evaluation Criterion	Weighting Factor
Right-of-way	Yes	No	NA
Conflicts with other construction	Yes	No	NA
Agency coordination required	Yes	No	NA
Impacts to high-risk utilities	Yes	No	NA
Soil classification	Yes	No	NA
Hydraulic head	Yes	Yes	2
Vegetation type	Yes	Yes	2
Space for dosing/mixing/monitoring	No	Yes	4
Basin capacity (percent WQV)	No	Yes	4
Percent Caltrans runoff	No	Yes	3
Sufficient space for design WQV	No	Yes	3
No. of inlet pipes	No	Yes	3
Distance to paired site	No	Yes	3
Depth to groundwater	No	Yes	2
Impacts to low-risk utilities	No	Yes	2
Soil percolation rate	No	Yes	2
Distance to edge-of-traveled Way	No	Yes	1
Space for 4:1 side slopes	No	Yes	1

3.2 Gather and Review Existing Data

Once the siting criteria have been established, the information needed to evaluate the candidate sites may be gathered or computed based on existing data. This activity encompasses three individual tasks: data collection, data review, and preliminary sizing calculations.

3.2.1 Collect Existing Data

Much of the information needed to evaluate the candidate sites already exists in one form or another. Performing data collection after the siting criteria have been established ensures that resources focus their efforts on gathering only the information that is actually needed. Both Caltrans and non-Caltrans sources should be included in the data gathering activity as they serve different purposes and complement each other. Internal Caltrans documents may include:

- Roadway as-built drawings
- Design drawings (for project areas under design)
- Drainage reports
- Aerial photography
- Storm drain outfall inventory database
- Project initiation documents
- Project approval documents
- Environmental documents
- Previous water quality/pilot study reports
- Geotechnical report
- Ongoing contract reports

During data gathering, consult with District staff to determine if there are any future planned projects in the area or if there is any additional information available. Non-Caltrans sources may include:

- Natural Resources Conservation Service (NRCS) soils data
- US Geological Survey topographic maps
- National Weather Service precipitation data
- US Geological Survey stream gauge records
- County land-use maps
- Previous water quality or pilot study reports
- Literature on pilot BMP under evaluation
- Public domain or proprietary aerial photography
- Public domain or proprietary GIS data
- City Plan, General Plan, or County Plan

A checklist that lists all candidate sites and possible sources may be used to document and record the data-gathering efforts. Check marks would be placed next to each site under the corresponding source.

3.2.2 Review Existing Data

As the information is collected, perform a data review to make sure they are appropriate for the intended purpose. For example, the following questions might be asked:

- Are the data relevant?
- Is any data missing?
- Are the data recent enough to be considered valid?
- Do the data reference another document or source?
- Are the data consistent with other documents?

As a result of the review, it might be determined that the data are insufficient or incomplete, and additional data gathering is necessary.

3.2.3 Perform Preliminary Hydrology and Sizing Calculations

Some siting criteria may not be explicitly reported in the gathered documents and must be computed (such as the WQV or WQF). For example, as-built drawings will show drainage layouts and profiles, roadway vertical alignments, roadway lengths, and embankment slopes, but will not show runoff time of concentrations, drainage areas, available hydraulic head, or right-of-way needs. In addition, initial area (footprint) requirements may be estimated to assist with right-of-way assessments. One source for drainage area information is the Caltrans storm drain outfall inventory database. Note that the information in this database is not necessarily accurate (areas provided in the database were found to be off by as much as 100 percent in one of the Lake Tahoe pilot studies) and should be considered preliminary. If there is a concern regarding the information in the existing documentation, a basic topographic survey may be considered. In addition, the Caltrans Basin Sizer program may be used to estimate detention basin footprint requirements.

3.3 Confirm and Evaluate Site Data

The purpose of site evaluations and site visits is to confirm information gathered from the existing data collection task, gather additional site information, confirm space availability outside CRZ for BMP and monitoring equipment, and document the site photographically. This section presents some basic guidelines on conducting site visits and completing the site evaluations.

3.3.1 Conduct Site Visits

Visits to candidate pilot sites should be planned so that the time spent in the field is efficient and cost-effective. Some basic considerations are:

- Safety first. The safety of the inspection team and traveling motorists is the number one priority when working within the right-of-way. All necessary precautions should be taken to ensure safety.
- Get an Encroachment Permit (EP). An EP is required if siting activities are conducted by either non-Caltrans staff or an A-E consultant that is unable to use their A-E Contract as an EP (see Section 4.1.1). Although this is not necessary if a Caltrans employee is a member of the siting team, a permit should be obtained in case the Caltrans employee is unable to attend and siting cannot be delayed. The permit will specify any requirements the siting team must comply with, such as parking restrictions and dress codes (hard hats, safety vests, and safety goggles must be worn at all times).

- Travel in pairs. Siting should never be conducted alone and at least two people should perform site visits. This provides additional support for visual observations and data recording and ensures the safety of the individuals.
- Coordinate with District. Site visits be coordinated with District and/or Headquarters (HQ) staff. The site evaluation team should include a representative from the District Environmental Branch to facilitate the identification of any potential environmental issues.
- Bring gathered information. Information gathered during the data collection and analysis phase should be brought to the site visits (especially maps, construction plans, and aerial photographs) to help guide the field evaluation and confirm documented information (in the event that it may not be recent).
- Fill out field form. Field forms should be filled out while conducting the site visit. Taking notes on a separate piece of paper and transcribing them onto the field forms at a later time is not cost effective and leads to mistakes. Examples of site evaluation forms (i.e., field forms) are presented in Appendix F.
- Take photographs. Digital photographs should be taken of each site to record the conditions during the site visit. A photograph log should be maintained as each photograph is taken, recording the date, site location, and direction the camera is facing (use a compass if necessary).
- Review recorded data. Before departing from the site, review the data collected and photographs taken to avoid having to return to the site.
- Take Soil Samples. If soil analyses are required for design purposes (such as identifying concentrations of aerially deposited lead [ADL] or structural loading), consider taking the samples now but holding on to them until the analyses are actually needed.

3.3.2 Compare Site Data to Siting Criteria

Although the Screening Criteria Matrix is filled out for all candidate sites, the Siting Criteria Evaluation Matrix may not be. If resources are limited, it may be more beneficial to only fill out the evaluation matrix for sites that passed the screening criteria. Then, sites that did not pass screening may be revisited if the number of potential sites is insufficient. The Caltrans Task Order Manager should be consulted regarding which approach is taken and if there is a need to fill out the Siting Criteria Evaluation Matrix for all candidate sites.

An example Screening Criteria Matrix for the example siting criteria is presented in Table 3-3, and a corresponding example Siting Criteria Evaluation Matrix is presented in Table 3-4..

Table 3-3 Example Screening Criteria Matrix

Site ID	County	State Route	PM	Need Right-of-way?	Conflict with Other Construction?	Agency Coordination Required?	Impact to High-Risk Utilities?	Insufficient Hydraulic Head?
7	Placer	267	7.35	No	No	No	No	No
8	Placer	267	7.50	No	No	No	No	No
9	Placer	267	7.61	No	No	No	No	No
10	Placer	267	7.73	No	No	No	No	No
11	Placer	267	7.84	No	No	No	No	No
12	Placer	267	8.03	No	No	No	No	No
13	Placer	267	8.14	No	No	No	Yes	Yes
14	Placer	267	8.36	No	No	No	Yes	No
15	Placer	28	3.23	No	No	No	No	No
16	Placer	28	3.47	No	No	No	No	No
17	Placer	89	0.91	Yes	No	Yes	No	Yes
18	El Dorado	89	1.84	Yes	No	Yes	No	No
19	El Dorado	89	2.59	No	No	Yes	No	No
20	El Dorado	89	3.00	No	No	Yes	No	No
21	El Dorado	89	3.04	No	No	Yes	No	No
22	El Dorado	89	4.49	No	No	Yes	No	No
23	El Dorado	89	18.40	Yes	No	Yes	No	No
24	El Dorado	89	20.37	Yes	No	Yes	No	No
25	El Dorado	89	25.13	Yes	No	Yes	No	No
26	El Dorado	89	67.91	Yes	No	Yes	No	No

Table 3-4 Example Siting Criteria Evaluation Matrix

Site ID	County	State Route	PM	Hydraulic Head (ft)	Space to Dose/Monitor/Mix (m)	Basin Capacity (% WQV)	Caltrans % of Runoff (%)	Design WQV (m^3)	No. of Inlet Pipes	Distance to Paired Site (m)	Depth to Seasonally High Groundwater (m)	Impacted Low-Risk Utilities	Distance to Edge of Traveled Way (m)	Space for 4:1 Side slopes
7	Placer	267	7.35	3.4	> 6	200	100	100	2	0.2	> 3	No	> 9	No
8	Placer	267	7.50	3.4	> 6	100	100	160	1	0.2	> 3	No	> 9	Maybe
9	Placer	267	7.61	3.4	> 6	190	100	140	1	0.2	> 3	No	> 9	Maybe
10	Placer	267	7.73	3.4	> 6	140	100	120	1	0.2	> 3	No	> 9	Maybe
11	Placer	267	7.84	3.4	> 6	290	100	50	2	0.3	> 3	No	> 9	Maybe
12	Placer	267	8.03	3.0	> 6	140	100	200	1	0.2	> 3	No	> 9	Maybe
13	Placer	267	8.14	1.8	> 6	50	100	110	1	0.4	> 3	No	> 9	Maybe
14	Placer	267	8.36	3.4	> 6	160	100	190	1	0.4	> 3	No	> 9	Maybe
15	Placer	28	3.23	3.4	> 6	380	100	80	1	0.4	> 3	No	> 9	Maybe
16	Placer	28	3.47	3.4	> 6	80	100	390	1	0.4	> 3	No	> 9	Maybe
17	Placer	89	0.91	1.8	> 6	180	100	170	1	5.3	> 3	Yes	> 9	Yes
18	El Dorado	89	1.84	3.4	> 6	150	100	170	1	1.2	> 3	Yes	> 9	Yes
19	El Dorado	89	2.59	3.4	> 6	100	100	150	1	0.7	> 3	Yes	> 9	Yes
20	El Dorado	89	3.00	3.4	> 6	210	> 60	140	1	0.1	> 3	Yes	> 9	Yes
21	El Dorado	89	3.04	3.4	> 6	210	> 60	120	1	0.1	> 3	Yes	> 9	Yes
22	El Dorado	89	4.49	3.4	> 6	310	100	80	1	2.3	> 3	Yes	> 9	Yes
23	El Dorado	89	18.40	2.4	> 6	500	> 60	30	1	3.2	> 3	Yes	> 9	Yes
24	El Dorado	89	20.37	3.4	> 6	1500	100	30	1	3.2	> 3	Yes	> 9	Yes
25	El Dorado	89	25.13	3.0	> 6	90	100	530	1	5.3	> 3	Yes	> 9	Yes
26	El Dorado	89	67.91	4.6	> 6	200	100	620	1	10.9	> 3	Yes	> 9	Yes

3.4 Score Sites and Report Results

The final phase of project site selection is the analysis of the data to determine the most appropriate site(s) and the preparation of a TM to document the work performed.

3.4.1 Score the Site

The score for each site is computed by adding up the individual scores for each selection criterion. The individual scores for each criterion are computed by multiplying the criterion weighting factor and the normalized siting value. Given that the raw values for the selection criteria are in different formats (numbers, percentages, text), they must first be converted into a normalized value to ensure consistency and equality in the analysis. Table 3-5 presents the normalized values to be used and the corresponding definition.

Table 3-5. Pilot Study Siting Normalized Values

Normalized Value	Description
0	Unacceptable
1	Poor
2	Fair
3	Good
4	Excellent

To convert raw values into normalized values, create a conversion chart for each selection criterion, taking into consideration the preferred (desired) value, the minimal acceptable value, and the range of possible values. A number of different types of conversion charts are possible. For example, for selection criteria with numeric value, like hydraulic head, each normalized value may correspond to a range of raw values. If three (3) feet is the minimum acceptable hydraulic head, with four (4) feet being desired and greater than four (4) feet increasingly more difficult to manage, the following conversion chart might be used:

Raw Value (e.g., Hydraulic Head, ft)	Normalized Value
< 3.0	0
3.0 – 3.5	1
3.5-4.0	3
4.0	4
4.0-4.5	3
4.5-5.0	1
>5.0	0

The mapping in the above example is somewhat arbitrary and may vary based on the specific needs of the project. One approach to standardizing conversion of numeric values is to use the absolute difference between the desired value and the raw value, divided by the desired value, as a weighting factor:

Weighting Factor = |Desired Value – Raw Value|/Desired Value

The weighting factor is then multiplied by the normalized score for the desired value. For example, using hydraulic head as an example, if 4 feet of hydraulic head is desired, a range of normalized values of 0 through 4 is selected, and the maximum head is 5 feet, the resulting values may be:

RAW VALUE (E.G. HYDRAULIC HEAD, FT)	WEIGHTING FACTOR	NORMALIZE VALUE	ROUNDED NORMALIZED VALUE
0.5	1-|4-0.5|/4 = 0.0625	0.125	0
1	1-|4-1|/4 = 0.25	1	1
2	1-|4-2|/4 = 0.50	2	2
2.5	1-|4-2.5| = 0.625	2.5	3
3	1-|4-3|/4 = 0.75	3	3
4	1-|4-4|/4 = 1.00	4	4
4.5	1-|4-4.5|/4 = 875	3.5	4
5	1-|4-5|/4 = 0.75	3	3
>5	-	0	0

Note that using weighting factors may result in non-nominal normalized values. If nominal values are desired, round the normalized value following conventional rounding practices. Also note that the weighting factor approach symmetrically weights raw values according to its difference from the desired value.

For criteria in which the possible raw values are Yes or No, and Yes represents the undesired value (as in the case of the criterion Impacts to Low-Risk Utilities), the following mapping is suggested:

VALUE	NORMALIZE VALUE
Yes	1
No	3

The actual method used to convert raw values into normalized values is not critical; it is somewhat subjective. However, it is important that a logical process is used and documented (in a Siting Criteria Normalization Matrix) so that it can be reviewed, revised, and approved. An example matrix for the previously developed example is presented in Table 3-6.

Table 3-6. Example Siting Criteria Matrix for Normalized Values (NV)

Evaluation Criteria	NV=1	NV=2	NV=3	NV=4	NV=5
Hydraulic Head (yd)	< 2.1	2.1 – 2.2	2.2 – 2.3	2.3 – 3.0	> 3.0
Space to Dose/Monitor/Mix (yd)	< 4.9	4.9 – 5.1	5.1 – 5.3	5.3 – 5.5	> 5.5
Caltrans % of Runoff	< 60	60 – 70	70 – 80	81 – 99	100
Design WQV (cy)	< 100 or > 400	100 – 110 or 350 – 400	110 – 123	132 – 180	180 – 350
Caltrans % of Runoff	< 60	60 – 70	70 – 80	81 – 99	100
Design WQV (cy)	< 100 or > 400	100 – 110 or 350 – 400	110 – 123	132 – 180	180 – 350
No. of Inlet Pipes	> 3	3	2	-	1
Distance to Paired Site (miles)	> 5.0	3.0 – 5.0	1.0 – 3.0	0.5 – 1.0	< 0.5
Depth to Seasonally High Groundwater (yd)	< 2.0	2.0 – 2.5	2.5 – 3.0	3.0 – 4.0	> 4.0
Impacted Low-Risk Utilities?	Yes	-	Maybe	-	No
Distance to Edge of Traveled Way (yd)	< 3.0	3.0 – 6.0	6.0 – 9.0	9.0 – 10.0	> 10.0
Space for 4:1 Side Slopes	No	-	Maybe	-	Yes

After the raw values are converted into normalized values, compute the scores for each candidate site and document them in the Siting Criteria Scoring Matrix. Table 3-7 presents an example. Note that the sites that did not pass the screening criteria are highlighted.

Table 3-7 Example Siting Criteria Scoring Matrix

Site ID	County	State Route	PM	Hydraulic Head Score [Max of 4]	Space for Dosing/Monitoring/Mixing Score [Max of 4]	Basin Capacity (%WQV) Score [Max of 4]	Caltrans % of Runoff Score [Max of 4]	Design WQV Score [Max of 3]	No. of Inlet Pipes Score [Max of 3]	Distance to Paired Site Score [Max of 3]	Depth to Seasonally High Groundwater Score [Max of 2]	Impacted Low-Risk Utilities Score [Max of 2]	Distance to Edge of Traveled Way Score [Max of 1]	Space for 4:1 Side Slopes Score [Max of 1]	Total Score (Max Possible of 120)
7	Placer	267	7.35	4	4	4	4	1	2	4	3	4	3	0	98
8	Placer	267	7.50	4	4	3	4	3	4	4	3	0	3	2	100
9	Placer	267	7.61	4	4	3	4	3	4	4	3	4	3	2	108
10	Placer	267	7.73	4	4	3	4	2	4	4	3	4	3	2	105
11	Placer	267	7.84	4	4	4	4	0	2	4	3	4	3	2	97
12	Placer	267	8.03	3	4	3	4	4	4	4	3	4	3	2	107
13	Placer	267	8.14	0	4	0	4	2	4	4	3	4	3	2	77
14	Placer	267	8.36	4	4	3	4	4	4	4	3	4	3	2	111
15	Placer	28	3.23	4	4	4	4	0	4	4	3	4	3	2	103
16	Placer	28	3.47	4	4	2	4	1	4	4	3	4	3	2	98
17	Placer	89	0.91	0	4	3	4	3	4	0	3	0	3	4	74
18	El Dorado	89	1.84	4	4	3	4	3	4	2	3	0	3	4	96
19	El Dorado	89	2.59	4	4	3	4	3	4	3	3	0	3	4	99
20	El Dorado	89	3.00	4	4	4	1	3	4	4	3	0	3	4	97
21	El Dorado	89	3.04	4	4	4	1	2	4	4	3	0	3	4	94
22	El Dorado	89	4.49	4	4	4	4	0	4	2	3	0	3	4	91
23	El Dorado	89	18.40	3	4	4	1	0	4	1	3	0	3	4	75
24	El Dorado	89	20.37	4	4	4	4	0	4	1	3	0	3	4	88
25	El Dorado	89	25.13	3	4	2	4	0	4	0	3	0	3	4	73
26	El Dorado	89	67.91	4	4	4	4	0	4	0	3	0	3	4	85

3.4.2 Develop the Site Selection Technical Memorandum

The Site Selection TM documents the activities performed during site selection and the results obtained. The Site Selection TM should contain the following:

- Study goal, objectives, and BMP descriptions (from Study Plan TM)
- Caltrans and non-Caltrans data collected
- District coordination
- Candidate sites
- Screening selection criteria
- Evaluation selection criteria
- Evaluation selection criteria weighting factors
- Screening criteria matrix
- Siting criteria evaluation matrix
- Siting criteria normalization matrix
- Siting criteria scoring matrix
- Conclusions
- Siting deviations that impact study plan
- Selected site limitations (for future monitoring)
- Field forms
- Site photographs

In addition to the above, the following supplementary information should be included for the selected site(s) to ensure consistency with the International Stormwater BMP Database:

- City
- Zip code
- Altitude to nearest 100 feet
- Watershed name
- Total watershed area
- Total percent impervious area in watershed
- Most relevant regional climate station
- Land uses (for nonstructural pilots)

The Site Selection TM should follow the outline provided in Appendix G.

3.5 Task Order Development

Unless directed otherwise by the Department Task Order Managers, task orders with site selection activities should include, at a minimum, the following scope elements:

3.5.1 Kickoff Meeting

A kickoff meeting will be held at the Department's offices in Sacramento. The purpose of this meeting is to discuss the purpose, study questions, and approach for the project. In addition, potential locations or corridors of interest to the Department will be identified during this meeting. Meeting minutes will be taken and distributed to all meeting participants.

- Deliverables:
  - Meeting agenda
  - Meeting minutes

3.5.2 Develop Siting Criteria

Appropriate site screening and evaluation criteria for the pilot study will be developed, along with weighting factors. Siting and screening criteria will follow the guidance in the approved Study Plan TM and the Caltrans PSGM. A draft set of screening and evaluation criteria will be submitted to the Department Task Order Manager for review and approval before initiating field activities.

- Deliverables:
  - Draft siting criteria
  - Final siting criteria

3.5.3 Develop List of Candidate Pilot Sites

A list of potential sites will be identified through review of as-built plans and other information (for the locations or corridors of interest) and discussions with appropriate District staff. The PPT will review the information available and request additional relevant information if the available information is insufficient to evaluate the sites at a prescreening level.

- Deliverables:
  - List of candidate sites

3.5.4 Candidate Site Evaluations

Site visits of candidate study sites will be performed to complete the site evaluations and determine if specific sites need to be precluded because of site-specific characteristics. A site selection field form will be filled out for each site.

- Deliverables:
  - Site selection field forms

3.5.5 Site Analysis and Reporting

Sites will be scored in accordance with the approval Siting Criteria, and the siting matrices will be completed. A Draft Site Selection TM will be prepared and submitted to the Department Task Order Manager for distribution and review. The Draft Site Selection TM should include a “lessons learned” section. Following receipt of reviewer comments, a meeting will be held to review the draft submittal and comments received. A Response to Comment Form will be prepared with proposed responses to comments received, and submitted to the Department Task Order Manager for review. Following approval of the responses, a Final Site Selection TM will be prepared and submitted.

- Deliverables:
    - Draft Site Selection TM
    - Draft TM review meeting minutes
    - Draft TM responses to comments
    - Final Site Selection TM.

Chapter 4 Permits and Environmental Clearance

Following successful planning and selection of sites for the pilot study, appropriate permits and environmental approvals should be obtained. This chapter presents the approval and permitting requirements with which pilot studies need to comply.

Chapter Purpose and Desired Outcome

Chapter Purpose: Guide the PPT in needs and considerations related to required project permits and environmental clearances

Desired Outcome: Understand and obtain all required permits and environmental clearances

Note: Although final design documents (see Chapter 5) may be required to obtain final permits and approvals, the process should start before initiating design activities. Coordination with the appropriate District Environmental Branch at this stage will facilitate the permit/approval process and identify the level of design needed for various permits/approvals.

4.1 Permits

Permits required to conduct pilot studies fall within two major categories: those issued by Caltrans; and those issued by other agencies.

4.1.1 Encroachment Permits

Execution of a pilot study involves field activities within the Caltrans right-of-way in almost every phase, including site selection, reconnaissance surveys, topographic surveys, construction, construction site inspections, O&M, and monitoring. To perform these activities, individuals other than Caltrans employees are required to obtain an EP from the local District. As described in the Caltrans EPs Manual (Caltrans 2018a):

"An encroachment is defined in Section 660 of the California Streets and Highways Code as '[…] any tower, pole, pole line, pipe, pipeline, fence, billboard, stand or building, or any structure, object of any kind or character not particularly mentioned in the section, or special event, which is in, under, or over any portion of the [State] highway right-of-way. "Special event" means any street festival, sidewalk sale, community-sponsored activity, or community approved activity.'"

An EP issued by Caltrans, therefore, provides the permittee the authority to enter the state highway right-of-way to construct, alter, repair, improve facilities, or conduct specified activities (Caltrans 2018a). It is not, however, a property right. EPs can be revoked by any departmental representative or law enforcement office if the permitted activity is deemed detrimental to the integrity of the state highway or to the safety of the traveling public.

Typically, each activity requires a separate permit. However, at the discretion of the District Permit Engineer, a permit rider may be issued to amend an existing permit to cover additional subsequent activities (for example, an EP may be initially issued for site reconnaissance work, and then amended with a permit rider to include surveys and geotechnical investigations).

An A-E Consultant does not always need to obtain a Caltrans EP. As specified in the May 2002 Division of Procurement and Contracts Memorandum on Guidelines on the Issue of EPs for A-E Consultants (Caltrans 2002), the executed A-E Contract serves as the EP. The A-E Consultant needs to carry the contract and pertinent task orders at all times while working within the right-of-way. Subconsultants and subcontractors are also exempt from EP requirements as long as the work they are performing is within the approved scope of work. If the work is unforeseen and outside the approved scope of work, the subconsultant or subcontractor must obtain an EP from the local District. The memo also specifies that under special circumstances, the Contract Manager may request an EP. Besides these provisions specified in the May 2002 Memorandum, the District may still require an EP for some or all pilot study activities (such as traffic control). Therefore, coordination with the Caltrans Task Order Manager and District Encroachment Permit Office is required during the planning phase to determine if an EP is necessary for any pilot study activities. If it is determined that a formal EP is not necessary and the A-E Consultant will be using the A-E Contract as the EP, the respective District Encroachment Permit Office should be notified that the work will be performed in their District under an A-E Contract.

If it is determined that an EP is required, guidelines for preparing EP applications may be found at https://dot.ca.gov/programs/traffic-operations/ep.

It should be noted that the Environmental Division does not obtain EPs. However, Caltrans Environmental may review EP applications to determine whether environmental studies are necessary.

4.1.2 Other Permits and Approvals

Federal, state (other than Caltrans), and local resource agencies often have vested interests in projects, which they protect by requiring mitigation of project effects, or by requiring various approvals, permits, or agreements. When conducting pilot studies, these agencies typically include the US Army Corps of Engineers (USACE), the US Fish and Wildlife Service (USFWS), the California Department of Fish and Wildlife (CDFW), the State Water Board, and the local Regional Water Quality Control Board (Regional Water Board). Other agencies, such as the California Coastal Commission (CCC), the Tahoe Regional Planning Agency (TRPA), and the Bureau of Indian Affairs (BIA), may become involved, depending on the project's location or other circumstances. It is essential that all possible permits and approvals be evaluated and obtained for a specific pilot study, as the respective agency can shut down the pilot study, or impose fines, for failure to secure the necessary permit, certification, and/or agreement.

Additional information on non-Caltrans permits may be found in Caltrans Standard Environmental Reference (SER). Table 4-1 presents a summary of the non-Caltrans permits that may be required for a pilot study.

Table 4-1. Non-Caltrans Permit and Approval Requirements

Resource	Agency	Permit/Approval
Coastal Shoreline (except San Francisco Bay area)	CA Coastal Commission	Coastal Development Permit
Coastal Shoreline (San Francisco Bay area)	Bay Conservation and Development Commission	Coastal Development Permit
Lake Tahoe Watershed	TRPA	TRPA Project Permit
Stream Environment Zone	TRPA	TRPA Project Permit
Central Valley Floodways	Reclamation Board	Encroachment Permit
Water	USACE	Section 404 Permit
Water	Regional Water Board	Section 401 Water Quality Certification
Water	State Water Board	NPDES Permit (CWA Section 402)
Groundwater	Regional Water Board	NPDES Permit (CWA Section 402)
Fish and Wildlife Habitat, Threatened and/or Endangered Species	CDFW	Section 1602 Streambed Alteration Agreement
Fish and Wildlife Habitat, Threatened and/or Endangered Species	USFWS US Forest Service	Biological Opinion
Cultural Issues	State Historic Preservation Office	National Historic Preservation Act Section 106 Programmatic Agreement Approval
Cultural Issues	Advisory Council on Historic Preservation	National Historic Preservation Act Section 106 Programmatic Agreement Approval
Cultural Issues	Native American Tribes	Consultation

The District's Environmental Branch is responsible for the implementation of Caltrans policies, programs, and procedures concerning environmental considerations, analysis, and compliance with environmental laws and regulations under CEQA and NEPA as well as other state and federal regulations. Identification of applicable permits/approvals for pilot studies conducted under the A-E Contract Delivery Method are the responsibility of the PPT in coordination with the District NPDES Coordinator and District Environmental staff. As a result, execution of pilot studies requires close coordination with District Environmental staff to determine project schedules; identify potential project issues, criteria, constraints, and impact mitigation; and ensure that all laws and regulations are followed during the course of project development and system testing.

The PDPM (Caltrans 2019b) and SER (Caltrans 2019c) identify necessary federal, state, and local permits and approvals based on three project criteria: project location, affected resources, and construction activities. This section presents the more common environmental permits and approvals for each criterion that may be required for a pilot study. Table 4.2 presents a list of possible pilot study activities that may require an environmental permit/approval. Coordinate with the District NPDES Coordinator to identify any region-specific requirements, or exemptions that pertain to Regional Water Board processes. For example, in San Diego, a dewatering permit is only necessary if the dewatering quantity exceeds 100,000 gallons per day.

Table 4-2. Examples of Activities Requiring Environmental Permits/Approvals

Activity	Possible Permit/Approval
Constructing any part of a BMP pilot within Coastal Zone (except in the San Francisco area)	CCC Coastal Development Permit
Constructing any part of a BMP pilot in San Francisco within 3,000 feet of the coast	BCDC Coastal Development Permit
Constructing any part of a BMP pilot in the Lake Tahoe Watershed	TRPA Project Permit
Constructing any part of a BMP pilot in a Lake Tahoe Stream Environment Zone	TRPA Project Permit
Constructing any part of a BMP pilot within the floodway of a Central Valley stream regulated by the Reclamation Board	Reclamation Board Encroachment Permit
Constructing any part of a BMP pilot on a federal flood control levee or within the surrounding 10-foot Reclamation Board easement	Reclamation Board Encroachment Permit
Constructing an outlet pipe that directly connects to a waterbody or channel regulated as a water of the US	USACE Section 404 Permit Regional Water Board Section 401 Water Quality Certification
Discharging groundwater to a receiving water during geotechnical investigations	Regional Water Board NPDES Permit
Discharging groundwater to a receiving water during dewatering operations	Regional Water Board NPDES Permit
Removing (either temporarily or permanently) existing wildlife habitat within a USGS Blue Line Stream	CDFW Section 1602 Streambed Alteration Agreement, USACE Section 404
Removing (either temporarily or permanently) existing protected habitat (such as Coastal Sage Scrub)	CDFW Section 1602 Streambed Alteration Agreement
Removing (either temporarily or permanently) existing habitat used by federally endangered species (such as the California gnatcatcher)	USFWS Biological Opinion
Excavating in an area known to have archaeological significance and/or Native American concerns	National Historic Preservation Act (NHPA) Section 106 Programmatic Agreement; Native American Tribes Consultation
** This list is not all-inclusive.*	

Project Location

Work within certain geographic areas within the state may require a specific permit or approval, regardless of the pilot facility itself and any protected resources that may be affected. The two most common areas that require agency coordination are the Coastal Zone and the Lake Tahoe Basin.

Coastal Zone

"Coastal Zone" refers to the land and water area of California from the Oregon border to the border of the Republic of Mexico, extending seaward to the state's outer limit of jurisdiction (3 miles, offshore), including all offshore islands, and extending inland generally 3,000 feet from the mean high tide line of the sea. In significant coastal estuarine, habitat, and recreational areas it extends inland to the first major ridgeline paralleling the sea or five miles from the mean high tide line of the sea, whichever is less. In developed urban areas, the zone generally extends inland less than 3,000 feet. The CCC maintains detailed maps of the Coastal Zone for each coastal county, and should be consulted to determine if a specific site is actually within the Coastal Zone.

The CCC, in partnership with coastal cities and counties, plans and regulates the use of land and water in the Coastal Zone. In the San Francisco Bay and surrounding tributaries, development in the Coastal Zone is regulated by the Bay Conservation and Development Commission (BCDC), whose jurisdiction is defined as the open water, marshes and mudflats of greater San Francisco Bay, including Suisun, San Pablo, Honker, Richardson, San Rafael, San Leandro, and Grizzly Bays and Carquinez Strait; the first 100 feet inland from the shoreline around San Francisco Bay; and the portion of the Suisun Marsh below the ten-foot contour line, including levees, waterways, marshes, and grasslands.

If the pilot facility is located within the Coastal Zone, a coastal development permit will likely be required from the CCC, the BCDC, or the local government. More information on the Coastal Zone and associated permits may be found in Caltrans SER.

Example of Permit Requirements Affecting Study Design and Site Selection

During the design of the I-5/Palomar Road Biofiltration Swale Pilot BMP, it was discovered that the site was within the Coastal Zone, and the adjacent trees had to be protected in accordance with the Coastal Development Permit (CDP) in effect for the Cannon Road improvements. The CDP required that any existing trees removed by construction activities had to be replaced at a 5:1 ratio. To avoid tree disturbance and mitigation, the BMP was redesigned using short concrete channels to convey the runoff around the trees. As a result, the BMP final design consisted of three biofiltration swales and two intermediate concrete swales. Additional protection was provided by restricting excavation activities to the area beyond the tree dripline.

Lake Tahoe

Lake Tahoe is a unique national treasure known for its beauty, clarity, and many recreational opportunities. It is designated as an *Outstanding Natural Resource Water*, a special designation under the Clean Water Act. Since measurements began in the 1960s, Lake Tahoe has been losing an average of one foot of clarity per year and is currently listed as an Impaired Water Body under Section 303(d) of the Clean Water Act. In October 1998, Lake Tahoe was listed as a Category I impaired priority watershed under the California Watershed Assessment.

The TRPA is charged with protecting the water quality of Lake Tahoe and issues permits for activities within the Lake Tahoe Basin. If the pilot study is located within the Lake Tahoe Basin, a TRPA permit will most likely be necessary. Additional information on TRPA may be found at www.trpa.org/.

Stream Environment Zones (SEZs) and related hydrologic zones consist of the natural marsh and meadowlands, watercourses and drainage ways, and floodplains that provide surface water conveyance from upland areas into Lake Tahoe and its tributaries. SEZs are determined by the presence of riparian vegetation, alluvial soil, minimum buffer strips, water influence areas, and floodplains. The TRPA is responsible for the long-term preservation and restoration of SEZs.

TRPA policy requires the preservation of existing naturally functioning SEZ lands in their natural hydrologic condition; restoration of all disturbed SEZ lands in undeveloped, un-subdivided areas; and restoration of SEZ lands that have been identified as disturbed, developed, or subdivided to obtain a 5 percent total increase in the area of naturally functioning SEZ lands. Therefore, if the pilot facility must be located within or adjacent to a SEZ, coordination with the TRPA is required and mitigation may be necessary. In addition to water quality, TRPA has eight other thresholds to be met as part of the permitting process. Scenic resources could be affected by the proposed pilot projects. Additional information on SEZs is available at the TRPA website (www.trpa.org/) and the CA Tahoe Conservancy website (https://tahoe.ca.gov/).

Example of Requirements Affecting Environmental Clearances and Permits

The detention basin with outlet skimmer pilot BMP at the SR 73 and SR 133 intersect was created by retrofitting an existing compost stormwater filter BMP. However, to achieve the necessary basin capacity, the existing basin invert had to be lowered. As a result, the outlet channel's discharge point had to be moved 200 feet further downstream. Upon review by the Environmental Planning Unit, it was discovered that the outlet channel was a Water of the United States. Therefore, a Regional Water Board Section 401 Water Quality Certification and USACOE Section 404 Permit were necessary to construct the new discharge point.

Affected Resources

Federal, state, and local regulations are in place to ensure projects do not detrimentally affect protected resources. The primary resources that the pilot projects may affect are wetlands and waters of the US, cultural resources, and biological resources.

Wetlands and Waters of the US

The purpose of the Clean Water Act of 1977 and 1987 is to restore and maintain the chemical, physical, and biological integrity of the nation's waters through prevention and elimination of pollution. Section 404 of the Clean Water Act establishes a permit program administered by the USACE regulating activities affecting waters of the US. Waters of the US include surface waters such as navigable waters and their tributaries, all interstate waters and their tributaries, natural lakes, all wetlands adjacent to other waters, and all impoundments of these waters.

Section 401 of the Clean Water Act requires a water quality certification from the State Water Board or Regional Water Board when a project: (1) requires a federal license or permit (such as a USACE Section 404 permit); and (2) results in a discharge to waters of the United States.

Sections 1600 through 1616 of the California Fish and Wildlife Code were adopted to protect waters under state jurisdiction. Section 1602 of the code requires the CDFW to be notified before any project which would divert, obstruct, or change the natural flow or bed, channel, or bank of any river, stream, or lake.

Under the federal regulations, if a pilot study affects a water of the US, a Section 404 Permit and a Section 401 Water Quality Certification may be necessary. For example, installing a new discharge point from a BMP into a water of the US requires a permit. However, some of the nationwide permits do not require notifying the USACE to conduct specific activities. The permit itself will say whether notification is required. Nonetheless, most applicants inform the USACE when a 401 Water Quality Certification is required. Under state regulations, if the pilot study affects a state-jurisdictional water, a Section 1602 Streambed Alteration Agreement may be necessary.

Cultural Resources

Cultural resources encompass archaeological, traditional, and built environment resources, including but not necessarily limited to buildings, structures, objects, districts, and sites. Qualified cultural resources professionals, consulting with their peers, Native Americans, subject matter experts, or review authorities as necessary, conduct studies of those cultural resources that could have the potential to possess significance and that could be affected by transportation projects.

If it is determined by the PPT (in coordination with the District Environmental Branch) that the pilot study could affect a cultural resource, coordination and/or consultation with the appropriate agency (e.g., the State Historic Preservation Office or the BIA) may be required. Additional information on cultural resources and associated approvals may be found in the Caltrans SER.

Biological Resources

Caltrans must comply with federal and state environmental laws and regulations designed to protect biological resources in all phases of project planning and development, construction, permitting, and O&M. Biological resources include Habitats and Vegetative Communities, Migratory Corridors, Plants, Wildlife, Fisheries, and Special Status Species (regulated by a law, regulation, or policy, such as threatened and endangered species). In addition to the CDWF, federal agencies associated with the protection of biological resources include the USFWS and USACE. The PPT, in coordination with the District Environmental Branch, will determine if any biological resources are affected by the pilot study and what coordination is required with the appropriate resource agencies. Additional information on biological resources may be found in the SER.

4.1.3 Construction Activities

Federal, state, and local agency coordination may also be required depending on what activities will be performed during construction of the pilot facility. The PPT will determine the complete list of necessary permits and approvals during project development. In addition to any permit conditions already identified for the project, other activities during construction that may result in agency coordination are managing hazardous wastes and encroaching upon local streets and highways.

Hazardous Wastes

Hazardous wastes are different from other environmental issues in that sites contaminated with hazardous wastes will have an impact on the project rather than the project having an impact on the environment. Hazardous waste and hazardous materials include chemicals discharged to the environment that may adversely affect the environment or human health and safety. As presented in the SER, the word "contamination" is also used to indicate soil and groundwater impacted by chemicals.

The most common hazardous waste that might be encountered during the construction of a pilot study is ADL in the soil. In the past, Caltrans has applied for and received variances from the California Department of Toxic Substances Control (DTSC) for the reuse of some lead-contaminated soils in certain Caltrans Districts. The current variance allows Districts 4, 6, 7, 8, 10, and 11 to reuse lead-contaminated soil within Caltrans right-of-way in the roadway corridor boundaries under certain conditions if the soil was considered a non-Resource Conservation and Recovery Act (RCRA) waste. Districts not subject to the variance are required to haul all contaminated soil off to an appropriate disposal facility. The Caltrans permit requires written notification to the Regional Water Board at least 30 days before advertisement for bids of projects that involve soils subject to the variance. The Regional Water Board will then determine the need for development of Waste Discharge Requirements (WDRs) or written conditional approvals by Regional Water Board staff. It is recommended that the notification be submitted as early in the design phase as possible because the Regional Water Board may take as long as 180 days to issue WDRs. In addition, if the variance is to be invoked, public notification is required and the DTSC must be notified at least five days before construction. Coordination with the District NPDES Coordinator is required to determine if a variance is in effect at the time of the pilot study and the appropriate procedures are in place.

If lead is determined to be present in the soil at concentrations considered hazardous, the contract documents will have to clearly identify the contaminated soil and provide appropriate reuse/disposal procedures. Additional information on hazardous wastes may be found in the Caltrans SER. Additional information on the lead variances may be found on the Division of Environmental Analysis Hazardous Waste website: https://dot.ca.gov/programs/environmental-analysis/hazardous-waste.

Example of Hazardous Material Requirements Affecting a Pilot Study

The design of the SR 73 full sedimentation earthen berm Austin sand filter pilot BMP required excavation depths of up to five feet. During construction, unsuitable hazardous material from an apparent old abandoned dumpsite was discovered at lower depths. This hazardous material was not identified during the design-phase field investigations. The contractor was directed to over-excavate the area to remove all the material and replace it with imported borrow. The cost of this work (close to $200,000) used all the contract contingency funds and forced a suspension of the contract until additional funds were secured.

Encroachments

Although the pilot facility may be located within Caltrans right-of-way, access from a non-Caltrans roadway might be necessary. For example, if the pilot facility is constructed adjacent to an elevated roadway, the best access might be from a city street and not from the roadway itself. If this is the case, a driveway approach might be needed to get from the city street to the pilot facility. This approach is an encroachment upon the city's right-of-way and requires a permit from the city.

4.1.4 Environmental Commitments Record

An environmental commitment is a measure that Caltrans or a local agency commits to implement to avoid, minimize, and/or mitigate a real or potential environmental impact. It can be identified as early as the planning and scoping stages, during the ED or design processes, or as late as construction or O&M of a project. It can be something as simple as a requirement for seasonal work windows or as complex as a treatment plan for cultural resources.

An Environmental Commitment Record (ECR) tracks and documents the completion of Environmental Commitments through the Project Delivery Process. It brings all the relevant environmental compliance information together in a single place, making it easier to track progress and easier for project team members (Environmental staff, Project Engineer, Project Manager, Resident Engineer) to identify actions they need to take. The ECR also aids in preparing and updating the RE Pending File, in executing Environmental Certification at the Ready to List stage, and in preparing the Certificate of Environmental Compliance.

An ECR is required for all projects on the State Highway System. The form of the ECR is determined at the District level. For details, see the June 10, 2005 memorandum by the Caltrans Chief Engineer.

For a listing of typical commitments, see the guidance document attached to the memorandum and titled “Environmental Commitments Record Standards and Instructions.” This document guides you to ask the “who, what, when, and where” questions to help logically document the commitments.

A sample ECR form can be found on the SER forms and templates webpage: https://dot.ca.gov/programs/environmental-analysis/standard-environmental-reference-ser/forms-templates. Please note that regardless of the format chosen to record environmental commitments, mitigation measures required for significant impacts under CEQA should be listed separately from other avoidance, minimization, and mitigation measures.

4.1.5 Traffic Control Certification

A Traffic Control Contractor must be certified to perform lane or shoulder closures. Specifically, the Traffic Control Contractor must possess a valid, current C31 license. A Traffic Control Contractor prepares or removes lane closures, flagging, or traffic diversions, using portable devices, such as cones, delineators, barricades, sign stands, flashing beacons, flashing arrow trailers, and changeable message signs, on roadways, including, but not limited to, public streets, highways, or any public conveyance. This requirement applies to all phases of a project where lane closures are required. For additional information, see the California Business and Professions Code, Sections 7008, 7058, and 7059.

Lane closures must be registered with the respective Caltrans District Traffic Management Control (TMC) before the Traffic Control Contractor can perform the closure. The closure needs to be scheduled two weeks in advance. Once a closure is registered with the respective District TMC, an authorization number and approved work hours will be issued. The Traffic Control Contractor must call the respective District TMC with a valid authorization number each time a closure is set up and removed. If an EP was issued for the pilot study, the respective District encroachment permit office will register the closure with the District TMC and issue the authorization number and approved work hours. If the pilot study is being performed under an A-E Contract, and no encroachment permit was issued, the HQ PPT will register the closure with the respective District TMC and issue the authorization number and approved work hours.

All lanes closures must conform to the appropriate traffic control system. Requirements are available on Caltrans Traffic Operations website (https://dot.ca.gov/programs/traffic-operations).

4.2 Environmental Clearance

Environmental clearance involves preparation of the ED by the PPT unless otherwise instructed by the Caltrans Contract Manager or as directed by the local District Environmental Branch. Specific environmental clearance requirements may vary by District, so it is important to coordinate efforts with appropriate District environmental staff early in the permitting/approval process.

Depending on the funding source of the pilot study, it will fall under the jurisdiction of CEQA and may fall under NEPA. As all projects require environmental review, contact the designated District environmental staff for direction on any technical studies needed to be performed by professional archaeologists and/or biologists to complete the ED for water quality, biological, cultural, and other resources. For each site, procedures in the SER for CEQA (and NEPA if needed) compliance will be followed. However, because specific requirements may vary between districts, appropriate procedures should be confirmed with the District Environmental Branch before initiating any technical studies. Specific guidelines for CEQA and NEPA compliance are provided in the SER.

Effective July 1, 2007, the Department has been assigned environmental review and consultation responsibilities under NEPA pursuant to Section 6005 of the Safe, Accountable, Flexible, and Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU). The assignment applies to all projects on the State Highway System (SHS) and all Local Assistance projects off the SHS, with the exception of responsibilities assigned for certain CEs under the June 7, 2007 Memorandum of Understanding (MOU) with the Federal Highway Administration (FHWA), projects excluded by definition, and specific project exclusions. On projects for which Caltrans has assumed NEPA responsibilities, Caltrans has also assumed responsibility for environmental review and consultation under other federal environmental laws. Refer to Caltrans SER for detailed guidance on the policy and procedures for compliance with NEPA and other federal environmental laws, regulations, and executive orders for projects assigned to the Department.

Compliance with NEPA is required if the pilot project will also involve federal funding or approval by the FHWA, or Caltrans as assigned by the FHWA (23 U.S. Code §327 2010.), or a permit or approval from a federal agency. Under NEPA, the ED may be a CE, EA, or EIS. The preferred document under NEPA is the CE because it indicates minimal or no environmental impacts, which requires the following conditions:

- The project does not have a significant impact on the environment.
- The project does not involve substantial controversy on environmental grounds.
- The project does not involve significant impacts on publicly owned lands of a public park, recreation area, or wildlife and waterfowl refuge of national, state, or local significance, or land of an historic site of national, state, or local significance.
- The project does not involve significant impacts on properties protected by the National Historic Preservation Act.
- The project comes from either a currently conforming plan under a Transportation Improvement Program, or is exempt from regional conformity.
- The project is consistent with all federal, state, and local laws relating to the environmental aspects of the project.

CEQA compliance is required for projects for which Caltrans has a discretionary action unless the project is exempted by statute in an act of the state legislature. Under CEQA, the ED may be a CE, Initial Study (IS), or EIR. The preferred document for a pilot study is the CE. This means that there is no possibility that the pilot study may have a significant effect on the environment. To qualify for a CE, the following criteria must be met:

- The project does not affect an environmental resource of hazardous or critical concern.
- There will not be a significant cumulative impact by the project and successive projects of the same type in the same place, over time.

- There is not a reasonable possibility that the project will have a significant effect on the environment due to unusual circumstances.
- The project does not damage a scenic resource.
- The project is not located on a site that is listed on the DTSC Hazardous Waste and Substances Site List (i.e., the Cortese List).
- The project does not cause a substantial adverse change in the significance of a historical resource.

Additional information on preparation of EDs may be found in the SER and Section 4.2 of this PSGM.

Example of Environmental Clearance

The installation of stormwater monitoring stations in the Northern California region required the filing of a Categorical Exemption that included technical studies examining biological resources and cultural resources at the installation sites.

The biological documentation included a letter report documenting development of a species list for each monitoring station's 7.5-minute quadrangle by searching the California Natural Diversity Database (CNDDB) and the Sacramento USFWS website; and examination of photos of the site to determine the existing conditions at each location and the possibility for special-status species to occupy these sites. Where necessary a 9-quadrangle species list was obtained from the CNDDB and the California Native Plant Society Inventory.

The archeological documentation included historical resource compliance reports prepared by professional archeologists. The archeological survey was performed in accordance with the Caltrans Section 106 Programmatic Agreement (Section 106 PA) and CEQA Guidelines §15064.5(a).

Aerially deposited lead was the only hazardous waste concern at the monitoring stations. An email from the North Region Environmental Engineering Office stating that previous testing in the region of the monitoring stations showed aerially deposited lead levels below hazardous concentrations.

4.3 Task Order Development

Unless directed otherwise by the Department Task Order Manager, task orders with environmental clearance and permit/approval activities should include, at a minimum, the following elements (with language as proposed).

4.3.1 Draft Environmental Document

The Department Task Order Manager will provide project information and mapping to the designated PPT contact when a pilot study project is proposed. Based on direction from District staff, the necessary ED must be completed. Any environmental technical studies (such as for water quality, biological, or cultural resources) required by District Environmental staff must be performed following the approval of the Department Task Order Manager. Coordination and consultation with District staff must be performed as needed. The draft and final technical reports must be reviewed and approved by District Environmental staff. The procedures in the SER must be followed, including the preparation of environmental approval documents for CEQA and possibly NEPA compliance. Draft environmental documents must be submitted to the Department Task Order Manager for distribution and review by the PPT and District Environmental staff. If necessary, environmental documents must be circulated for public review.

4.3.2 Final Environmental Document

Following Department review and receipt of Department comments on the Draft Environmental Document, revised documents must be prepared and submitted to the Department Task Order Manager. The Department must sign the form and make the final determination.

4.3.3 Permits and Approvals

All EPs and environmental permits/approvals necessary to conduct the pilot study must be identified. Coordination with the appropriate District EP Office and resource agencies must be conducted and the necessary permit applications and supporting information must be prepared by the PPT. Application submittal and payment of permit fees must be performed by the Department DEA Stormwater Program.

Chapter 5 Project Design

This chapter presents guidance for designing a pilot study. Table 5-1 shows the steps involved.

Chapter Purpose and Desired Outcome

Chapter Purpose: Guide the PPT in designing and documenting the project elements

Desired Outcome: A Plans, Specifications, and Estimate (PS&E) package that will be used for construction contracts and activities

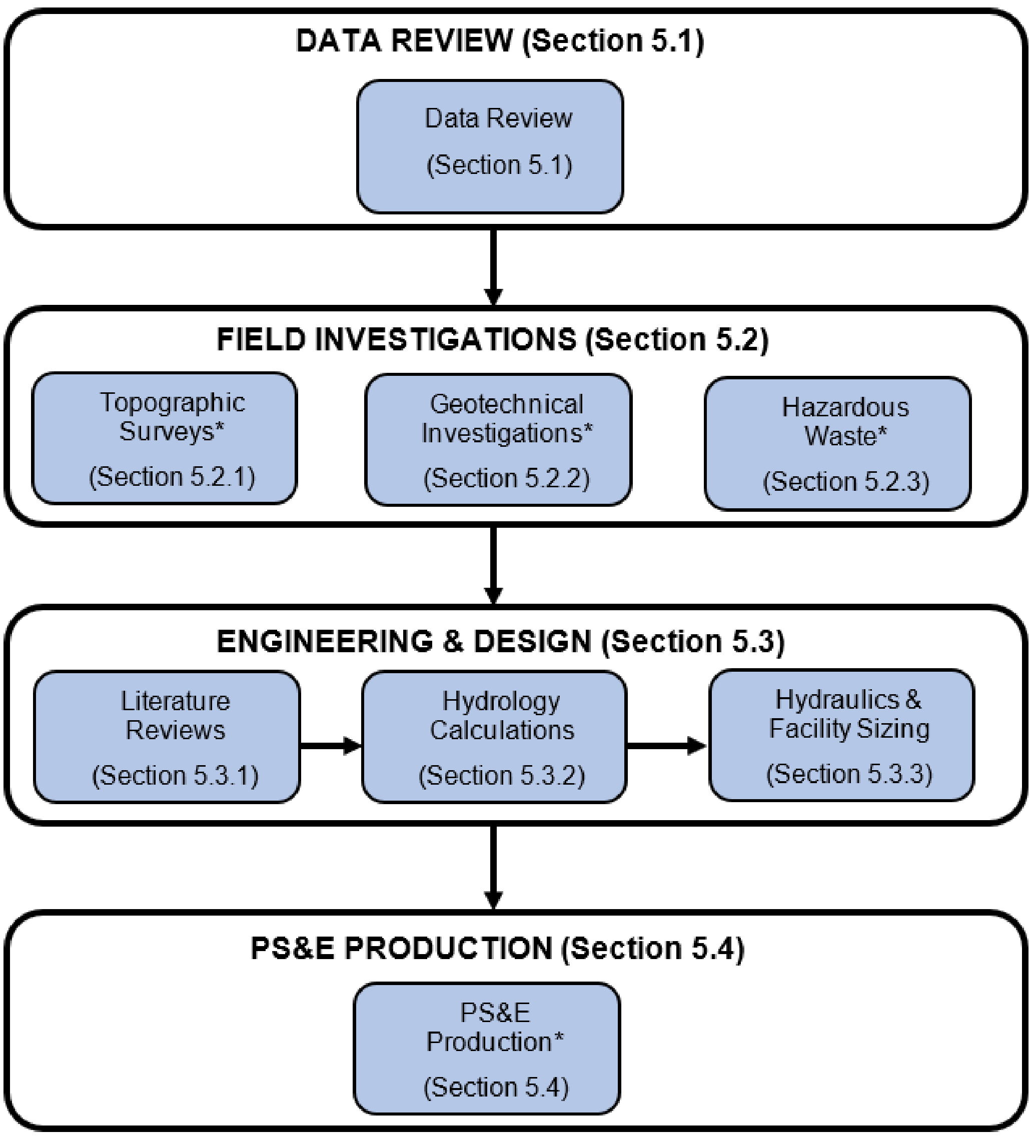


Figure 5-1 Design Phase Process Flowchart

5.1 Data Review

Review the Study Plan TM and the Site Selection TM for the study background, objectives, and decisions to-date. Additional/supplemental information requirements should be submitted to the Task Order Manager for approval so the information may be collected during subsequent field investigations.

5.2 Field Investigations

Conduct field investigations to collect information that was not collected during site selection and/or to determine current site characteristics and conditions. Such investigations can take 4 to 12 weeks to complete, depending on the number and size of the pilot sites.

5.2.1 Topographic Surveys

Evaluate the need for a survey. Surveys are performed to:

- Provide an accurate base map for the BMP layout and design.
- Verify information used during site selection, such as preliminary tributary drainage areas, WQVs and/or WQFs, and existing drainage facility connectivity (i.e., the drainage system as shown in as-builts).

To develop a base map, use aerial photography (with ground control) or conduct a ground survey. The result is a digital terrain model of the area with existing grade contours, existing visual roadway and drainage features (e.g., roadways, shoulders, curbs and dikes, barriers, light standards, drain inlets, manhole covers), and spot elevations of existing visual features (e.g., outfall pipe invert elevations, manhole/drain inlet rim elevations). The survey should include the immediate project site and enough of the surrounding area to support the design of a construction areas sign and the traffic handling elements. Refer to the PDPM (Caltrans 2019b) for relevant requirements.

Note that a topographic survey may not be cost effective, especially for large sites. A visual reconnaissance survey (using as-builts to field verify drainage area high and low points, grade breaks, flow paths, etc.) may be a more cost-effective alternative.

Examples of Surveys

For the BMP Retrofit Pilot Study, drainage areas were determined from topographic surveys. For the SR 73 Pilot Studies, drainage areas were determined using as-built drawings and visual reconnaissance surveys. Acceptable results were achieved in both studies.

$$$ Cost Reduction Strategies $$$

Perform adequate site and geotechnical surveys to avoid unexpected costs and ensure post-construction BMP effectiveness, especially for infiltration type BMPs.

5.2.2 Geotechnical Investigations

Evaluate the need for a geotechnical investigation. Such investigations can determine current soil and groundwater conditions at the project site and identify geotechnical constraints that must be incorporated into the design. This work is performed by licensed professionals who provide the project Design Engineer with recommendations for embankments, fill and cut slope design, expansive and soft soil treatment, foundation bearing pressures, groundwater control, seismic stability, potential liquefaction, retaining walls, earth/water retaining structures, sound walls, and culvert foundations. The work typically includes:

- A literature search
- Review of geologic maps
- Surface geologic investigation
- Subsurface geologic and geotechnical investigation
- Soil infiltration rates and other soil tests required by the Professional Engineer
- Laboratory analysis
- Submittal of a Geotechnical Design Report that includes a log of test borings

5.2.3 Hazardous Wastes

In the contract documents, specify any hazardous wastes within the project limits to ensure appropriate handling and disposal. Although the Task Order Manager may request a detailed investigation for any hazardous materials, the most common material encountered during pilot studies, and therefore tested for, is ADL. ADL concentrations within soil impacts the costs for proper reuse/disposal. Some sites may have localized areas of ADL-laden soils, where only portions of the site will be subject to ADL handling and disposal requirements. The most efficient strategy is to draw a grid over the excavated area and take one sample from each grid cell. For appropriate test methods, refer to Caltrans Division of Environmental Analysis Hazardous Waste website: https://dot.ca.gov/programs/environmental-analysis/hazardous-waste.

Example of Addressing Hazardous Wastes

The Initial Site Assessment (ISA) for the SR 73 pilot study concluded that hazardous wastes might be present at any of the 38 project sites. This was confirmed at the University Drive Phase II inclined screen GSRD BMP pilot site when ADL was discovered during the subsequent hazardous waste investigation. Although the excavation area covered more than 2 acres, enough samples were taken to determine that the ADL was limited to the northeast corner of the basin, thereby minimizing disposal costs.

5.3 Engineering and Design

During engineering and design, the information collected in previous activities is used to size and lay out the pilot study infrastructure.

The following Caltrans documents provide design guidance:

- Construction Contract Development Guide (CCD, Caltrans 2019d)
- Project Planning and Design Guide (PPDG, Caltrans 2019e)
- HDM (Caltrans 2018e)
- Design Memoranda (https://dot.ca.gov/programs/design/design-memoranda)
- PDPM (Caltrans 2019b)

Note that the HDM must be followed for drainage design methods and that water quality design must follow the process described in the PPDG and SWMP to comply with the Caltrans NPDES Permit.

5.3.1 Literature Review

Review the Study Plan TM to understand the intent of the project and previously identified BMP components, including variables that must be controlled. A literature review should have been conducted before or during the study planning stage, and relevant information and decisions should have been documented in the Study Plan TM. If necessary, consult the Caltrans Task Order Manager to obtain approval for an additional literature review to glean further knowledge or fill in gaps.

5.3.2 Detailed Hydrology Calculations

Review the Site Selection TM to see if preliminary hydrology calculations were performed to help identify sites with appropriate WQFs or WQVs. These calculations may have been sufficient for siting purposes, but may not be detailed or accurate enough for engineering and design purposes. Verify the hydrology assumptions and perform detailed calculations.

WQFs and WQVs are calculated using drainage areas, runoff coefficients, rainfall intensities, and rainfall depths, and are in turn used to design the BMP as well as associated monitoring equipment. Note that design criteria for treatment and monitoring differs from that for drainage and flood control, with smaller more frequent storms targeted for treatment and monitoring design, and larger, less frequent storms targeted for drainage and flood design. General guidelines for calculating roadway flows are provided in Caltrans HDM (Caltrans 2018e). Specific guidelines for calculating BMP WQFs and WQVs are provided in Caltrans PPDG (Caltrans 2019d). The WQF and WQV calculations are based on specific requirements stated in the Caltrans NPDES permit. Note that individual counties may have their own guidelines for flood control calculations.

The Caltrans tool Basin Sizer (http://svctenvims.dot.ca.gov/wqpt/basinsizer.aspx) computes rainfall depths and rainfall intensities for water quality events. The Caltrans Rainfall Intensity-Duration-Frequency (IDF) PC Program can also be used. The latter incorporates the California Department of Water Resources (DWR) short-duration precipitation data. Rainfall intensities are listed in the SWMP and have been previously negotiated with the State Water Board and Regional Water Boards. These rainfall intensities should be used as the basis for designing the approved flow-based Treatment BMPs.

Caltrans HDM (2018e) provides a list of hydrologic software that is approved for use by Caltrans. The use of software that is not included requires the approval of the Caltrans Task Order Manager and possibly the District Hydraulic Engineer.

Example of Hydrology Calculations

Two hydrologic software packages that are not listed in Caltrans HDM but have been approved for specific pilot studies are HydroSoft by Advanced Engineering Software (AES) and HydroCAD by HydroCAD Software Solutions, which have been used in pilot studies in Southern California and Lake Tahoe, respectively. If these or other methods or packages are needed for the study, check with Caltrans design staff to obtain approval.

5.3.3 Detailed Hydraulic Calculations and Facility Sizing

Use the hydrologic data (Section 5.3.2) and existing facility data to calculate velocities, water surface elevations, backwater effects, and scour depth, which will be used to size the BMP and associated components (e.g., culverts, outlet structures, flow measurement devices, spillways, energy dissipaters, and bypass flow splitters). Hydraulic calculations should include those for effluent discharge rates to confirm that they do not exceed established design criteria.

One particular criterion to note is vector control, as cited in the Caltrans permit: "All storm water BMPs that retain storm water shall be designed, operated and maintained to minimize mosquito production, and to drain within 96 hours of the end of a rain event, unless designed to control vectors. BMPs shall be maintained at the frequency specified by the manufacturer. This limitation does not apply in the Lake Tahoe Basin and in other high-elevation regions of the Sierra Nevada above 5000 feet elevation with similar alpine climates. The Department shall operate and maintain all BMPs to prevent the propagation of vectors, including complying with applicable provisions of the California Health and Safety Code relating to vector control." (State Water Board 2017).

The calculation and sizing methods should support the study objectives as well as adhere to applicable hydraulic design criteria and standards. Caltrans HDM (2018e) provides general Caltrans guidelines for roadway hydraulics. The guidelines presented previously on hydrologic software are also applicable to hydraulic software.

$$$ Cost Reduction Strategies $$$

- Use natural landscape features and materials instead of concrete and other structural components.
- Minimize support features such as fencing, access roads, and gates to those necessary for safety and O&M.
- Minimize access road surfaces to what is necessary for O&M and use permeable materials where feasible (although permeable materials may have a higher capital and O&M cost as compared to asphaltic concrete).
- Use prefabricated components as much as possible.

5.4 Plans, Specifications, and Estimate (PS&E) Production

The PS&E specifies what will be constructed (plans), how it will be constructed (specifications), and how much it is anticipated to cost (estimate).

5.4.1 Plans

Develop project plans that identify what work will be performed and where. The plans will consist of engineering drawings that contain information that is unique to the project and does not exist in the Caltrans Standard Plan or Revised Standard Plan. Prepare the plans following the guidelines in the Plans Preparation Manual (Caltrans 2018d) and CADD Users Manual (https://dot.ca.gov/programs/design/manual-cadd-users-manual). Table 5-1 presents a typical list of the drawings, although the actual list will depend on the study's purpose, complexity, required level of detail, available schedule, and available budget. Simple studies might only include a title sheet, drainage plans, and drainage detail sheets, while more complex projects might use the entire list. For simple projects, certain drawings may be combined, such as the layout, grading, and drainage plan.

Table 5-1. Typical Roadway Plan Set[1]

Sheet Name	Description/Contents
Title Sheet	Project vicinity map and limits of construction
Layouts	Layout and location of roadway items of work (access roads, driveways, fencing), existing utilities, permanent barriers (guard rail, curbs/dikes), right-of-way limits
Construction Details	Supplementary details (dimensions, materials, typical sections) of layout items of work that are specific to the project and are not in the Standard Plans
Temporary Water Pollution Control Quantities	Pay quantities of temporary water pollution control items
Erosion Control Plans	Layout and location of permanent erosion control materials
Erosion Control Quantities	Pay quantities of erosion control items
Contour Grading	Existing and proposed site grading, limits of disturbed soil, top of cut, toe of fill, limits of contaminated soils
Drainage Plans	Layout and location of existing and proposed drainage facilities (culverts, BMP structures, flow measurement devices, drainage inlets, manholes)
Drainage Profiles	Profiles of each drainage system, which permit the determination of excavation and backfill quantities
Drainage Details	Construction details for drainage items of work, which are specific to the project and are not in the Standard Plans
Drainage Quantities	Pay quantities of drainage items
Traffic Handling	Layout and location of long-term traffic control systems, to show how traffic is to be routed and maintained within the limits of the project
Construction Area Signs	Location and type of temporary signs required for the direction of public traffic through or around the project site
Summary of Quantities	Pay quantities for layout, grading, construction area signs, and traffic handling items of work
Planting Plans	Layout and location of permanent non-erosion control plants (e.g., shrubs, ground cover, trees)
Irrigation Plans	Layout and location of permanent irrigation facilities necessary to support the plants shown in the Planting Plans
Electrical Plans	Layout and location of electrical items of work (usually in support of BMP-related mechanical equipment)

[1] *Plan development must follow Caltrans Plans Preparation Manual (Caltrans 2018d)*

5.4.2 Specifications

Develop needed specifications for work items that are not addressed by Caltrans Standard Specifications, Standard Special Provisions (SSPs), or non-Standard Special Provisions (nSSPs). Specifications dictate how the work is to be performed and what materials are to be supplied. Refer to Caltrans Construction Contract Standards: http://ppmoe.dot.ca.gov/hq/esc/oe/old%20hq-esc-oe%20pages/construction_standards_01-15-16.html.

Caltrans Standard Specifications are general statewide procedures prepared and maintained by Caltrans for the management and execution of projects. They provide detailed information for items of work contained in the plans and the specific requirements for measurement. Caltrans Standard Specifications are developed in consultation with construction and maintenance. Modification of existing specifications must receive concurrence from construction and maintenance divisions.

Contract Special Provisions are procedures prepared by the Design Engineer that supplement or supersede the Standard Specifications for a particular project. Contract Special Provisions are made up of boilerplate Caltrans SSPs maintained by Caltrans. In addition, nSSPs are prepared by the Design Engineer for specific project needs. The nSSPs are maintained by functional owners and any modification need to be approved by the owner. New nSSPs may be submitted for approval to the concerned functions. The nSSPs developed for past pilot projects include those listed in Table 5-2. Contact the Office of Stormwater Management, Design for additional information.

Table 5-2. nSSPs Developed for Past Pilot Studies

nSSP	nSSP
Fiberglass Flume	Hold and Release Valve
Filter Media	Gate Valve
Floating Skimmer	Monitoring Well
Geomembrane Liner	Trench Rock
Gross Solids Removal Device	Tube Settlers
Roosting Bat Treatment	Environmentally Sensitive Areas
Bird Exclusion Netting (Swallows)	-

Note that if special, proprietary materials are to be used, the materials must meet product evaluation requirements for Caltrans Authorized Materials Lists: https://dot.ca.gov/programs/engineering-services/authorized-materials-lists.

Finally, incorporate important construction guidelines. For example, most BMPs infiltrate into underlying soils, so it is critical that heavy construction equipment is not used during basin excavation.

5.4.3 Estimate

Develop an estimate of the expected cost for construction. For each work item, include an item code, name, and description, the unit of measure (e.g., each, lump sum, cubic yard, linear foot), the quantity, the unit price (e.g., cost per liner foot or cost per cubic yard), and the amount (the unit price times the quantity).

The item code is a unique identifier that corresponds to the standard Basic Engineering Estimating System (BEES) codes maintained by Caltrans. However, these codes have been developed from roadway projects and do not always include nonstandard specialty items that may be included in a pilot study. Guidelines for assigning item codes to nonstandard items are provided in the CCD Guide (Caltrans 2019d).

Unit prices are developed based on experience, engineering judgment, vendor information, and values from previous work. The Caltrans annual Contract Item Cost Data Summary (https://d8data.dot.ca.gov/contractcost/) provides average unit rates for standard contract items from previously awarded contracts. Note that the rates represent weighted averages and are affected by many factors, including location, schedule, and quantity. The quantity of the work item is especially important, since pilot studies usually have much lower quantities than typical roadway projects, and therefore significantly higher unit prices. Caltrans publishes annual Contract Data Cost Data Books (https://sv08data.dot.ca.gov/contractcost/).

Due to the small quantity of asphalt material to be used for BMP pilots, the section for price cost adjustment for asphalt based on oil price index should be deleted. Quantity for all work of excavation and embankment should be based on 0% shrinkage.

The estimate is included with each PS&E submittal (draft, revised, and final). A sample format for a cost estimate is presented in Table 5-3.

Table 5-3. Sample Cost Estimate Format

Item	Code	Description	Unit	Qty	Unit Price	Item Total
1	74019	Prepare Stormwater Pollution Prevention Plan	LS	1	$7,000	$7,000
2	74020	Water Pollution Control	LS	1	$15,000	$15,000
3	74028	Temporary Fiber Rolls	FT	1,310	$4	$5,240
5	74032	Temporary Concrete Washout Facility	EA	3	$2,000	$6,000
6	74033	Temporary Construction Entrance	EA	3	$3,500	$10,500
7	74038A	Temporary Inlet/Outlet Protection	EA	4	$500	$2,000
8	120090	Construction Area Signs (S)	LS	1	$4,000	$4,000
9	120100	Traffic Control System	LS	1	$45,000	$45,000
11	150806	Remove Pipe	FT	76	$30	$2,280
12	150821	Remove Headwall	EA	2	$1,000	$2,000
15	153219	Remove Concrete Flared End Section	EA	2	$750	$1,500
16	160101	Clearing and Grubbing	LS	1	$6,000	$6,000

17	190101	Roadway Excavation	CY	4,660	$30	$139,800
18	190110	Lead Compliance Plan	LS	1	$4,500	$4,500
-	**-**	**-**	**-**	**-**	**Total**	**$125,000**

5.5 Basis of Design Report

After completing the hydrology calculations, develop a basis of design report to document the assumptions and criteria to be used in subsequent design tasks. Appendix H provides an example outline for a design report. Specific contents include:

- Introduction (Section 1.0 of Appendix H)
- Site Conditions (Section 2.0 of Appendix H)
- Design Criteria (Section 3.0 of Appendix H)
- Hydrologic and Hydraulic Analysis (Section 4.0 of Appendix H)
- Design Deviations that Impact Study Plan (Section 5.0 of Appendix H)
- Design Summary (Section 6.0 of Appendix H)
- Construction Support (Section 7.0 of Appendix H)
- OM&M Support (Section 8.0 of Appendix H)
- References (Section 9.0 of Appendix H)
- Appendices (Section 10.0 of Appendix H)

Prepare a draft basis of design report and submit it to the Caltrans Task Order Manager with the Draft PS&E. Following receipt of comments, prepare and submit the final basis of design report no later than one week before beginning construction.

5.6 Stormwater Data Report

The Stormwater Data Report (SWDR) summarizes the following pilot study information:

- Project description
- Stormwater quality design issues
- Regional Water Board agreements
- Proposed design pollution prevention, treatment, and construction BMPs

Pilot studies should typically use the “Short Form” SWDR depending on the water quality issues. Submit a Draft SWDR with the Draft PS&E. Submit the final SWDR with the Final PS&E. Detailed guidelines and instructions for preparing SWDRs can be found in the PPDG (Caltrans 2019d).

5.7 Typical Timelines

Table 5-4 lists typical durations for the PS&E process. Figure 5-2 depicts a timeline.

Table 5-4. Typical Pilot Study PS&E Timeline

Activity	Duration (weeks)
Field Investigations	2 – 8
Preparation and submittal of Draft PS&E	4 – 12
Review of Draft PS&E by Department Task Order Manager and Others	2 – 6
Preparation and submittal of Revised PS&E	2 – 12
Review of Revised PS&E by Department Task Order Manager and Others	2 – 6
Preparation and submittal of Final PS&E	1 – 4
Review/approval of Final PS&E by Task Order Manager	1 – 4
Total	**14 – 52**

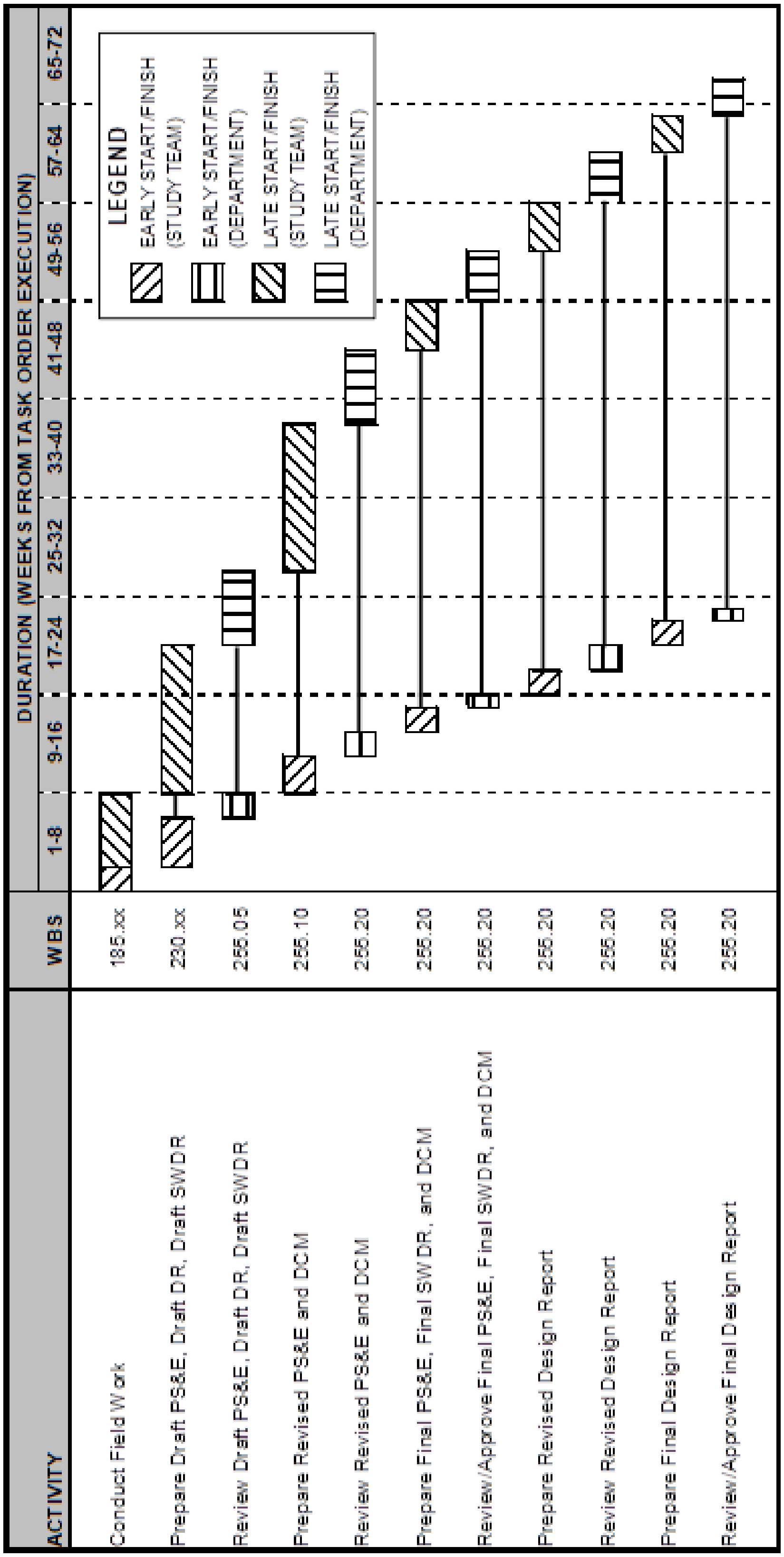


Figure 5-2. Typical Pilot Study PS&E Timeline

5.8 Task Order Development

Table 5-5 presents the scope of design activities to be used in task orders.

Table 5-5. Design Task Order Scope Elements

Scope Activity	Scope Element	Brief Description	Deliverable(s)
Data Review	Data Review	• Review the Study Plan Site Selection TMs and determine if additional data are needed	• Tech Memo recommending type of additional data
Data Review	ADL Site Investigation	• Determine levels of ADL within project limits	• ADL Report
Field Investigations	Surveys and Mapping	• Conduct field survey and create a topographic base map	• Base Map
Field Investigations	Geotechnical Investigation	• Determine soil and groundwater characteristics	• Geotechnical Design Report
Field Investigations	Hydrology & Hydraulics	• Conduct hydrology and hydraulic calculations to determine size and dimensions of pilot components	• Pilot Component Sizes
Engineering & Design & PS&E Package	Draft Roadway Plans	• Prepare draft roadway drawings	• Draft Plans
Engineering & Design & PS&E Package	Draft Specifications	• Prepare draft special provisions	• Draft Special Provisions
Engineering & Design & PS&E Package	Draft Estimate	• Prepare draft quantities and construction estimate	• Draft Estimate
Engineering & Design & PS&E Package	Draft SWDR	• Prepare draft SWDR	• Draft SWDR
Engineering & Design & PS&E Package	Draft Design Report	• Prepare draft design report	• Draft Design Report
Engineering & Design & PS&E Package	Revised PS&E	• Prepare Design Changes Memo (DCM) • Address comments on draft PS&E	• Response to Comments • Revised PS&E • DCM
Engineering & Design & PS&E Package	Final PS&E	• Address comments on Revised PS&E and draft SWDR • Prepare Final PS&E, Final SWDR, and DCM	• Response to Comments • Final PS&E • Final SWDR • DCM
Engineering & Design & PS&E Package	Final Design Report	• Address comments on DCM and draft Design Report • Incorporate DCM into draft Design Report • Prepare final Design Report	• Final Design Report

Chapter 6 Project Construction

This chapter presents guidelines for performing construction-related activities for a pilot study. Figure 6-1 presents the flowchart.

Chapter Purpose and Desired Outcome

Chapter Purpose: Guide the PPT in pilot study construction activities and considerations

Desired Outcome: As-builts and a construction report that documents installation of all project components

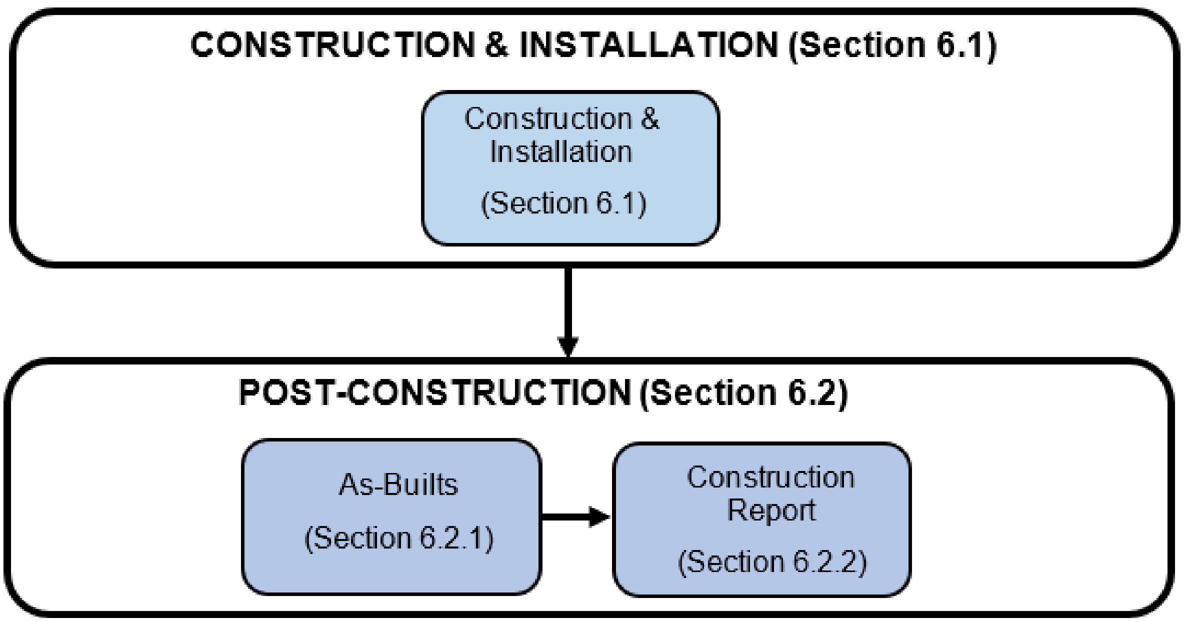


Figure 6-1. Construction Process Flowchart

6.1 Construction and Installation

Under the A-E Contract Delivery Method, construction and installation are performed by an A-E Consultant. When the drawings and specifications are prepared by the same A-E Consultant who is providing the construction/installation services, the A-E Consultant's Resident Engineer and inspection staff bear primary responsibility for the work. The local Caltrans district Construction group may assign a Resident Engineer and/or inspector(s) who will serve as observers. If an EP was necessary (see Section 4.1.1), the local District EP Office provides an EP Officer who acts as a part-time inspector to verify that the work is being performed in accordance with the issued EP.

Responsibilities for the PPT during actual construction and installation activities include:

- Water pollution Control
- Hazardous Waste Management
- Submittals, Requests for Information, and Requests for Clarification
- Progress Documentation

There responsibilities are described below.

6.1.1 Water Pollution Control

To control the discharge of pollutants during construction, Caltrans requires the development and implementation of either a Stormwater Pollution Prevention Plan (SWPPP) or a Water Pollution Control Program (WPCP). The Caltrans General Construction Permit specifies the type and size of projects that require a SWPPP. A WPCP is not required by the permit, but is currently required by Caltrans for all other cases. The SWPPP, or the WPCP, must be certified by the A-E Consultant's resident engineer and reviewed for acceptability by the Caltrans Construction representative. If a SWPP is required, then the A-E Consultant must also prepare a Notice of Construction (NOC), which must be submitted to the Regional Water Board 30 days before construction.

Guidelines, templates, and examples of SWPPPs and WPCPs are available on the Division of Construction website: https://dot.ca.gov/programs/construction/storm-water-and-water-pollution-control.

6.1.2 Hazardous Waste Management

Caltrans manages many contaminants and waste streams. These wastes are frequently encountered or produced by Caltrans construction and maintenance projects or are present on Caltrans property and must be managed when disturbed. The contaminants and waste streams include:

- Aerially Deposited Lead (ADL)
- Asbestos-Containing Materials (ACM)
- Naturally Occurring Asbestos (NOA)
- Perchlorate
- Treated Wood Waste
- AC Grindings
- PCC Grindings
- Petroleum Hydrocarbons

Refer to the Chapter 18 of the PDPM (Caltrans 2019b) for general policies and procedures regarding inspections for and management of contaminants and waste streams. More details can be accessed at the Department Hazardous Waste website: https://dot.ca.gov/programs/environmental-analysis/hazardous-waste. The district hazardous waste office must be consulted on hazardous waste issues.

6.1.3 Submittals, Requests for Information, and Requests for Clarification

Because the A-E Consultant who prepared the drawings and specifications may also be the Contractor, there may or may not be any formal submittals or Requests for Information (RFIs) or Requests for Clarification (RFCs) between the construction staff and the engineering/design staff. However, significant construction-initiated design changes (i.e., shop drawings) should be brought to the attention of the Permit Officer (if an EP was required) and the Caltrans Task Order Manager in a timely manner. Any deviations from the contract plans must be brought to the attention of the Department Task Order Manager immediately.

6.1.4 Progress Documentation

To document and report construction progress, daily construction reports are prepared and submitted to the EP Officer (if an EP was required) and Caltrans Task Order Manager. The Daily Construction Report includes the following information:

- Contract number and task order number
- EP number and site ID/name
- Date of the report
- Names of EP Officer and Caltrans Task Order Manager
- Weather conditions
- Contractor personnel and visitors on site
- Equipment on site
- Summary of daily progress, including photos
- Concerns and issues
- Summary/status of key construction features
- Name and signature (electronic) of individual preparing report

An example Daily Construction Report is presented in Figure 6-2. The concerns and issues entries in the “Construction Issues/Concerns Discussed” section (see example report in Figure 6-2) should be assigned a numerical number to facilitate tracking, and a running log should be maintained and submitted to the Department Task Order Manager during the construction phase. The key construction features described in the “Key Construction Items” section of the Daily Construction Report (see the example report in Figure 6-2) must be approved by the Department Task Order Manager. The report should be submitted within three working days of the date of the report.

In addition, the A-E Consultant responsible for the construction/installation is required to maintain a set of redline plans at the construction site to document design changes made during construction for the purpose of preparing the as-built drawings (see Section 6.2.1). Information to be shown on the redlines is presented in Caltrans Construction Manual: https://dot.ca.gov/programs/construction/construction-manual.

District 12 Route 73 Pilot Study
Daily Construction Report

Date:	10/19/05
Permit Number:	12-0C9854
BMP Site ID:	780R
Permit Officer:	Diba Kazerani
Dept. Task Mgr:	Tim Sobelman

Day	S	M	T	**W**	Th	F	S
Weather	Bright	Sunny	**Over Cast**	Rain	Snow		
Temp.°F	<32	32-50	50-70	**70-85**	85-100		
Wind	**Still**	Moderate	High				
Humidity	Dry	**Moderate**	Humid				

Manpower		
Contractor	**No.**	**Remarks**
Foremen / Superintendents	1	
Laborers	7	
Visitors	1	Permit Officer, David Alderete

Equipment on Site

- 1 – Komatsu PC228USLC Track Hoe
- 1 – Cat 936F Loader
- 1 – Cat 420D Backhoe/Loader
- 1 – Kubota KH-60 Track Hoe
- 1 – Electronic Sign Board
- 1 – Dynapac 4469 6ft Sheepsfoot Roller
- 1 – 2000 Gallon Water Truck
- 1 – Case 60XT Loader
- 2 – Wacker Trench Roller
- 1 – Wacker Vibratory Rammer

Construction Activities

- Limited production due to rain on the 17th
- Completed backfill of CSF area
- Placed and compacted base at inflow monitoring location
- Installed rebar for walls at outflow monitoring location
- Continued trenching for pipe installation

Construction Issues/Concerns Discussed

3. Met with Inspector. Concerns discussed were: 1) unexpected dewatering required for trenching to the concrete channel 2) unexpected dewatering required for excavating for the inflow monitoring station floor

Key Construction Items

Item	Installed yet?	
Outlet Riser	☐ Yes	☒ No
Monitoring	☐ Yes	☒ No
Skimmer	☐ Yes	☒ No

Figure 6-2. Example Daily Construction Report

6.2 Post-Construction

Once construction of the pilot project has been completed, there are only three outstanding responsibilities of the PPT:

- Preparation of as-built drawings (including shop drawings)
- Preparation of the Construction Report
- Certificate of Environmental Clearance

6.2.1 As-Builts

As-built plans (sometimes also referred to as record drawings) are prepared following completion of construction and represent existing field conditions at the completion of the project. As described above, the as-builts are based on the set of redlines maintained by the consultant or contractor during construction. Final as-builts are to be submitted to the Department Task Order Manager and the District. Coordination with the District NPDES Coordinator and Department Task Order Manager are necessary before District submittal.

Preparation of as-builts must be in accordance with the Caltrans CADD Users Manual: https://dot.ca.gov/programs/design/manual-cadd-users-manual. As presented in the CADD Users Manual, as-built plans must include revisions to alignments and rights-of-way, grade revisions, drainage changes, changes to roadway features, and revisions in location of utility crossings and irrigation crossovers. Minor feature changes (such as grade revisions less than 30 mm) and actual construction quantities are typically not reflected in as-built drawings. The Caltrans Construction Manual requires the as-builts to be completed within 60 days of contract acceptance.

6.2.2 Construction Report

The Construction Report, prepared following the completion of construction, documents the differences between what was designed and what was constructed. This documentation is necessary to capture important changes that were made in the field that could affect OM&M activities or future pilot projects.

The Construction Report should include a discussion of the following topics:

- Changes to the plans (reflected in the as-builts and shop drawings)
- Changes to the special provisions (e.g., material substitutions)
- A summary of CCOs (for standard and CCO projects)
- A summary of construction costs
- Revisions to the siting and design criteria presented in the Site Selection TM and the Design Report

- Construction deviations that impact the Study Plan (based on review of revisions to siting and design criteria)
- Construction lessons learned (refer to Section **8.7.2** for details)
- Construction site visit reports and as-builts

Actual construction costs incurred for the pilot study, obtained from the contractors' invoices and related material quantities, must be reported and presented in a tabular format. The construction cost items must be broken down by those items included in the original bid schedule, additional items of work authorized following contract award (e.g., change orders), and state-furnished materials (e.g., biofilter sod). Any construction related monitoring costs must also be included but separated from costs for the construction of the BMPs. Actual construction costs are typically presented for the following categories

- Bid item description, quantity, and unit cost
- Additional work (change orders)
- Actual cost
- Monitoring costs (separate from construction costs)

Site-specific costs that are unique to the pilot study, such as the use of stormwater pumps, the inclusion of guardrails, or unique stormwater sampling/monitoring systems, must be reported separately from the actual construction costs. Other site-specific costs may include items such as location, limited space, utility conflicts, dewatering, and others.

The PPT must prepare draft and final versions of the Construction Report, submitting the draft within two months of completion of construction, and the final in accordance with an agreed upon review period schedule. Follow the outline provided in Appendix I.

6.2.3 Certificate of Environmental Compliance

The purpose of the CEC is to document the Department's environmental compliance efforts at Construction Contract Acceptance (CCA) for all measures specified in final environmental (or other project) documentation, including permits and agreements, and inform all project stakeholders, including regulatory agencies, as to the outcome of the Department's environmental commitment measures.

The information contained within the CEC should be based on the ECR. The CEC, along with the updated ECR, should be filed in the Uniform File System and a copy retained in the project history file as evidence that the Department has met its obligations to fully document environmental compliance efforts for projects. For additional information, see the Caltrans Workplan Standards Guide (Caltrans 2018c) and the SER (Caltrans 2019c).

6.3 Typical Timeline

A typical timeline for the construction process for the A-E contract delivery method is presented in Figure 6-3. Depending on the complexity of the project, the construction phase may take anywhere from 14 to 52 weeks.

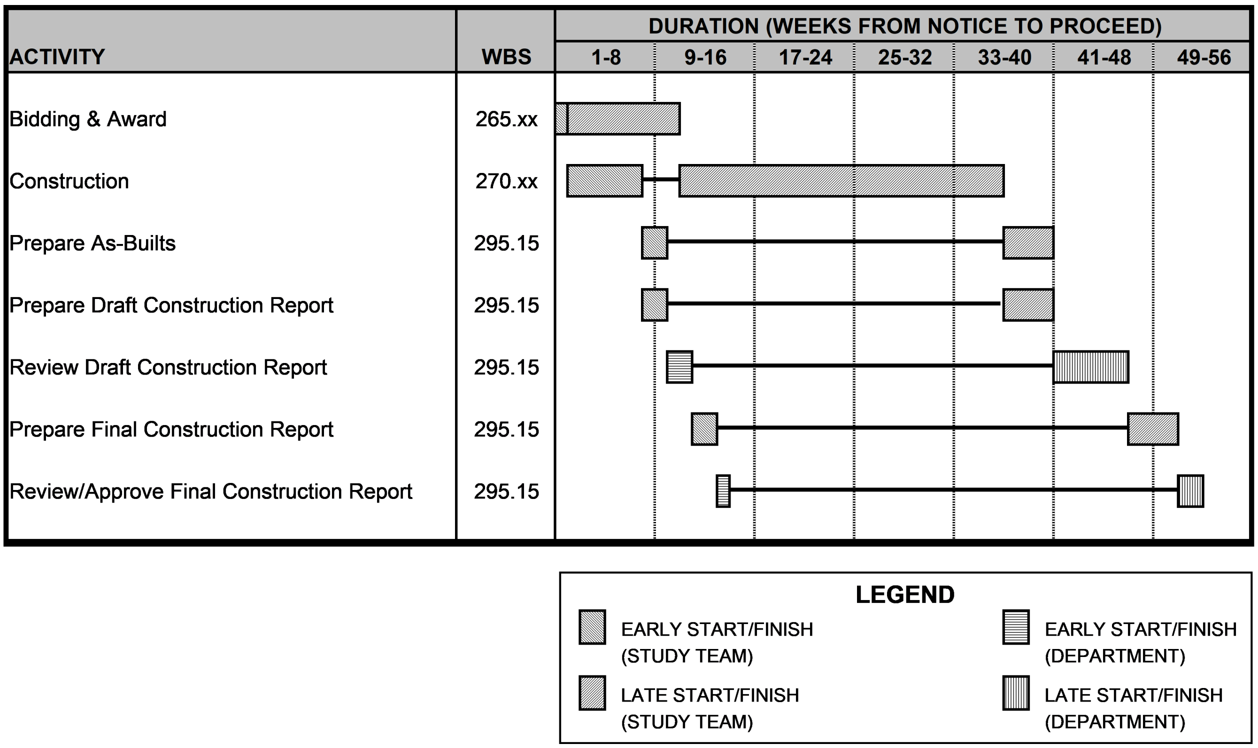


Figure 6-3. Typical Pilot Study Construction Timeline

6.4 Task Order Development

Unless directed otherwise by the Department Task Order Manager, task orders for construction activities should include, at a minimum, the scope elements outlined in Table 6-1.

Table 6-1. Construction Task Order Scope Elements

Scope Element	Brief Description	Deliverable(s)
Project Management	Administration, coordination, scheduling, and quality control	• Meeting minutes • Monthly progress reports • Invoices
Construction	Construction/Installation of pilot facilities	• SWPPP (or WPCP) • Lead compliance plan (if ADL present) • ADL excavation & disposal plan (if ADL present) • Pilot Facilities
Construction Administration	Administration of construction/installation tasks	• Preconstruction meeting minutes • Notice of construction (if SWPPP used)
Construction Inspection	Oversight and inspection of construction activities to ensure compliance with contract documents	• Daily construction reports • Redline markups
Contract Acceptance	Acceptance of construction activities and preparation of final documents	• Pre-final inspection punch list • Notice of completion of construction (if SWPPP used) • Contract acceptance form • As-builts • Construction Report • CEC

Chapter 7 Operation, Maintenance, and Monitoring

Conditions such as erosion, vegetation height, trash, and sediment buildup may directly affect the performance of a BMP, so activities must be conducted to ensure proper performance of the BMP. Likewise, specific monitoring procedures must be conducted to obtain reliable data. This chapter presents the necessary steps (Figure 7-1) related to OM&M of a BMP pilot study.

Chapter Purpose and Desired Outcome

Chapter Purpose: Guide the PPT in proper implementation of maintenance and monitoring procedures

Desired Outcome: Study data that has been collected and documented following Caltrans OM&M procedures

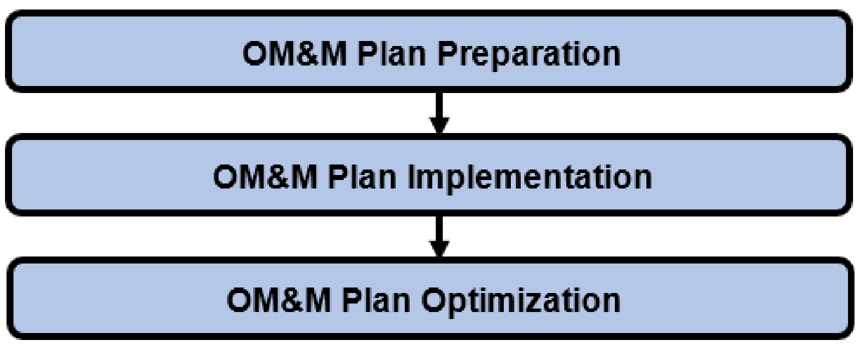


Figure 7-1. BMP Pilot Study OM&M Process Flowchart

7.1 OM&M Plan Preparation

Prepare an OM&M Plan that defines the procedures to operate, maintain, and monitor the BMP(s) in the study. To prepare the OM&M plan, the PPT must:

- Review relevant pilot study documents
- Determine O&M activities
- Specify maintenance threshold indicators
- Identify equipment and tools needed
- Create checklists for inspectors and maintenance personnel
- Describe O&M cost collection procedures
- Identify monitoring and analysis procedures
- Develop a health and safety plan (HSP)

- Document the OM&M plan

The following subsections provide specific details on what is required for each of these activities.

7.1.1 Review Relevant Pilot Study Documents

The PPT must review the Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a) and the Study Plan TM (Chapter 2 of this PSGM), which will form the basis for the OM&M Plan. In addition, review the Design Report and Construction Report before developing the OM&M Plan is developed so that the final site characteristics are established and deviations from the Study Plan are identified. The OM&M Plan acts as a bridging document between the Study Plan TM and the day-to-day operations of the pilot study. It must be practical and clearly define operating procedures. Adherence to the OM&M Plan by the PPT assures that, throughout the study, procedures are consistent and data is properly documented.

7.1.2 Determine O&M Activities

Most, if not all, BMPs require routine maintenance, and inadequate attention to the maintenance schedule poses a high risk of operative failure. Maintenance of BMPs must include regular inspections and, if necessary, appropriate repairs or remedies. The requirements vary from system to system, but the end result must always be to operate and maintain the BMP under a maintenance schedule similar to what the BMP might normally receive in a typical, real-world application. The types and extent of maintenance activities should be as minimum as possible, in order to align with Caltrans O&M criteria listed in the SWMP (Caltrans 2016). Specific types of O&M activities can be categorized as follows:

  - Routine inspection and maintenance requirements
  - Erosion and structural repair
  - Debris and Litter removal
  - Sediment removal and disposal
  - Mowing
  - Nuisance control
  - Vector management

The PPT must discuss the project with District maintenance so there is a clear understanding of who will handle maintenance issues; however, the PPT may have to perform some equipment-specific maintenance.

The frequency of some BMP maintenance activities will depend on the pollutant-loading rate and the ability of the BMP to remove and retain these pollutants. Most factors that influence pollutant-loading rates are site-specific, such as erodibility of native soils and landscaping materials, land-use activities, and flow dynamics. Storage capacity is also a significant factor affecting maintenance frequency.

There may be occasions where emergencies, such as accidents, spills, or other incidents arise, when critical response is needed and nonroutine maintenance is required. On those occurrences, Caltrans crews must respond accordingly, and, if necessary, the BMP may be taken out of service until its functionality can be restored. The pilot study health and safety plan (HSP, see Section 7.1.8) will describe possible types of emergency conditions and appropriate mitigation.

Example of Incorporating O&M into the Pilot Study

For the Caltrans RVTS Study, it was important that Caltrans consultants coordinate with District maintenance. Caltrans consultants wanted District maintenance to continue with their routine activities (e.g., mowing) during the study. As a result, the consultants needed to track District maintenance's activities to determine their influence on the study.

Routine Inspection and Maintenance Requirements

Routine or preventive maintenance refers to procedures that are performed on a regular basis to keep the BMP aesthetic and in proper working order. Routine maintenance may include removing debris, removing silt and sediment, and clearing vegetation around flow control devices to prevent clogging. Sediment and debris removal is also important to ensure that monitoring equipment function properly. Routine maintenance also includes the maintenance of a healthy vegetative cover. Dead turf or other unhealthy vegetative areas must be replaced or restored. If the BMP has battery-operated components, the batteries may need to be recharged or replaced. Connections, fittings, valves, joints, screws, and other mechanical parts may need to be adjusted, repaired, or replaced. BMPs with chemical additives may require removal of spent materials and/or addition of new chemicals.

Inspections must be performed at regular intervals to ensure that the BMP is operating as designed. At a minimum, an annual inspection must be considered but additional inspections following storm events may be appropriate, depending on the design on the BMP and the study goals. For inspections following a major storm, the inspector must attempt to observe whether the BMP is properly passing, retaining, or infiltrating water, and whether the pollutant storage capacity has been exceeded.

Example visual observations during a routine inspection include checking:

- Accumulation of debris and sediment at the inlets and outlets
- Side slopes for signs of erosion, settlement, slope failure, or vehicular damage

- Emergent vegetation zones to ensure that water levels are appropriate for vegetative growth
- Whether vegetative cover is above acceptable limits
- Whether the water level is where it should be
- Evidence of frequent flow bypass

Nonroutine or corrective maintenance refers to rehabilitative activities that are not performed on a regular basis. Examples include flow control structure replacement and the major replacement and cleaning of aquatic vegetation.

Erosion and Structural Repair

Areas of erosion and slope failure should be repaired and reseeded (or sodded) as soon as possible. However, use of compost or fertilizers must be reviewed by the Caltrans Task Order Manager to ensure that the materials used do not affect the study results. Eroded areas near the inlet or outlet of the BMP may also need to be lined with riprap. Major damage to the inlet, outlet, or other structures must be repaired immediately. Delay in such repairs can cause structural failure. When that occurs, the BMP may require total reconstruction, resulting in delays, cost overruns, and potential invalidation of the pilot study. Damage to inlets and outlets can also affect the proper operation of the BMP, biasing the study results.

Debris and Litter Removal and Control

Debris and litter accumulate mostly near the inlet and outlet structures of BMPs and need to be removed when they threaten the proper operation of the BMP. Particular attention must be paid to floatable debris that can clog the outlet control structure or riser. Trash screens or trash racks can be strategically placed near inflow or outflow points to capture debris and assist with maintenance.

Litter and debris from illegal dumping must also be cleaned up immediately. An accurate log must be maintained of all the materials removed and improvements made. Controlling illegal dumping is difficult, although posting signs with a phone number for reporting a violation in progress may help. Notice of enforcement and substantial penalties for illegal dumping and disposal could also be a deterrent.

Sediment Removal and Disposal

Silt and other sediment removal activities can often require many laborers and heavy equipment over several days. Sediment needs to be removed on a regular schedule, but, as noted earlier, the frequency of removal is site specific. Regular inspections will reveal how often sediment must be removed.

Mowing

Side slopes, embankments, emergency spillways, and other grassed areas of BMPs must be periodically mowed to control weeds and prohibit woody growth. Mowing can constitute a large portion of the routine maintenance expense. Any materials used to control weeds and prohibit woody growth must be reviewed by the Caltrans Task Order Manager to ensure that adverse impact to the study results does not occur.

Nuisance Control

Standing water or soggy conditions in a study area or BMP can create nuisance conditions for nearby residents. Odors, vectors, weeds, and litter can be potential problems. Regular maintenance to remove debris and ensure BMP functionality will help control potential nuisance problems.

Vector Management

If the BMP being studied creates standing water, disease-carrying organisms such as insects (mosquitoes and midges) and rodents may have to be monitored and controlled. In such an instance, there will be a need to contract with local Vector Control Districts (VCDs) for inspection and abatement. Records must be kept of the frequency of inspections, the number of insects observed, and the abatement activities performed.

The strategy for the management of potential vector problems requires the following:

- Minimize the opportunities for such vector or nuisance organisms to become a potential problem.
- Develop a monitoring and maintenance program based upon defined and regular observations and inspections that will ensure that unsuitable conditions do not develop that encourage a vector or nuisance problem.
- Define threshold criteria to identify such a problem, and treatment guidelines to correct the problem.

An example vector management plan can be found in the Caltrans BMP Retrofit Pilot Program Final Report (Caltrans 2004).

7.1.3 Specify Maintenance Threshold Indicators

It is paramount, for consistency in operations, to develop specific thresholds for conditions that “trigger” maintenance activities. The thresholds must be determined before the start of the study. Maintenance activities must be defined for those times when the field measurement exceeds the maintenance indicators. These thresholds and associated maintenance activities must be based on existing technical literature and vendor-specific recommendations. For those BMPs that are not designed to contain standing water, regular inspections and maintenance activities must be planned to prevent the incidental formation of pools. For example, trash must be removed if it collects water. Again, records must be kept of the type and frequency of the performance of maintenance activities.

The PPT should prepare a table of BMP Inspection and Maintenance Requirements to summarize specific thresholds that trigger maintenance activities for the study BMP. Include the following information:

- *Design Criteria, Routine Actions* – Specific aspect of the BMP subject to inspection and maintenance (e.g., vegetation, battery voltage, sediment, standing water).
- *Maintenance Indicator* – Field measurement threshold that indicates the need for maintenance activities (e.g., average vegetation height greater than 18 inches; battery voltage drops below 11 volts; more than 85 percent of total volume filled with accumulated material; standing water for more than 72 hours; evidence of erosion; evidence of frequent flow bypass, wet season has ended).
- *Field Measurement* – Description of field measurement required to verify condition (e.g., visual observation; measure depth at apparent maximum and minimum accumulation of sediment and calculate average depth).
- *Measurement Frequency* – Frequency of inspection that potentially triggers maintenance (e.g., annually; 72 hours after a target storm event; before beginning of the rainy season and monthly during the rainy season).
- *Maintenance Activity* – Action that is needed when observed conditions exceed the maintenance indicator.
- *Site-specific Requirements* – Description of any unique site inspection and maintenance requirements.

Examples of maintenance indicator thresholds and maintenance activities are shown Table 7-1.

Table 7-1. Maintenance Indicator Thresholds from the Detention Basin Optimization Study

Design Criteria, Routine Actions	Maintenance Indicator	Field Measurement	Measurement Frequency	Maintenance Activity	Site-Specific Requirements
Basin vegetation	Average vegetation height greater than 18 inches, emergence of trees or woody vegetation	Visual observation and random measurements throughout the side slope area	Once during wet season, once during dry season	Cut vegetation to an average height of 6-inches and remove trimmings only if there is the potential to interfere with BMP outlet performance	Remove any trees or woody vegetation
Inspect for standing water	Standing water for more than 72 hours	Visual observation	Annually, 72 hours after a target storm (0.25 in) event	Drain facility, check and unclog clogged orifice; notify Caltrans Project Coordinator, if immediate solution is not evident	None
Inspection for sediment management and characterization of sediment for removal	Sediment fills 10 percent of basin volume or exceeds 18 inches in depth (evaluate marker on staff gauge)	Measure depth at apparent maximum and minimum accumulation of sediment; calculate average depth	Annually	Remove and properly dispose of sediment; regrade if necessary	None
Inspect for burrows	Burrows, holes, mounds	Visual observation	Annually and after vegetation trimming	Where burrows cause seepage, erosion and leakage, backfill firmly	None
Inspection for trash and debris	Debris/trash present	Visual observation	During routine trashing, per District schedule	Remove and dispose of trash and debris	None
Slope stability	Evidence of erosion	Visual observation	October each year	Contact the Caltrans Project Coordinator to determine the most appropriate erosion control method	None

Design Criteria, Routine Actions	Maintenance Indicator	Field Measurement	Measurement Frequency	Maintenance Activity	Site-Specific Requirements
General maintenance inspection	Inlet structures, outlet structures, side slopes or other features damaged, significant erosion, graffiti or vandalism, fence damage, etc.	Visual observation	Semiannually, late wet season and late dry season	Corrective action before wet season; consult Caltrans Project Coordinator if immediate solution is not evident	None
Inspect for clogged orifices in skimmer	Clogged orifice	Visual observation	72 hours after a target storm	Unclog the orifice	None
Inspect the skimmer connections	Loose hose connections	Visual observation	72 hours after a target storm	Tighten/repair the connection	None
Verify that skimmer floats	Skimmer below water line	Visual observation	During storm (or immediately after target storm)	Consult Caltrans Project Coordinator	None
Inspect skimmer for sediment between float valve and orifice plate	Sediment found between float valve and orifice plate	Visual observation	Just before start of rainy season	Remove sediment from between float valve and orifice plate	Applicable to 808R
Inspect the skimmer's lever connections of the lever operated flap valve outlet	Loose or broken connections	Visual observation	Just before start of rainy season and 72 hours after a target storm	Repair any loose or broken connections	Applicable to 859L
Verify that skimmer floats	Skimmer below water line	Visual observation	During storm (or immediately after target storm)	Consult Caltrans Project Coordinator	None
Check HRV[1] pressure in compressed air tank	Pressure less than 80 psi	Visual observation of pressure gauge	72 hours after a target storm, at beginning of rainy season and monthly during the rainy season	Add compressed air to tank to achieve a pressure of 120 psi	Applicable to 457L and 535L
Check all electrical connections tightness and corrosion in HRV	Loose or corroded connections	Visual observation	Just before the rainy season	Tighten, clean, and replace as necessary	Applicable to 457L and 535L

Design Criteria, Routine Actions	Maintenance Indicator	Field Measurement	Measurement Frequency	Maintenance Activity	Site-Specific Requirements
Check HRV battery voltage	Battery voltage drops below 11 volts	Measure battery voltage with voltage meter	72 hours after a target storm, at beginning of rainy season and monthly during the rainy season	Recharge or replace battery to 12 volts	Applicable to 457L and 535L
Check HRV timing cycle for proper operation	Timer initiates valve closing and holds valve closed for specified timer period	Manually initiate operation and visually check valve operation	Just before the rainy season	Reset timer if necessary; repair and replace if necessary	Applicable to 457L and 535L
Exercise the HRV and ensure that the valve closes tight	Valve does not close tight	Visual observation	Just before start of rainy season	If air pressure is sufficient, check to ensure solenoid valve is operating and repair as necessary; if solenoid valve is operating, check to determine if rubber liner needs to be replaced and replace if necessary	Applicable to 457L and 535L

[1]*HRV: Hold and release valve (Air-operated valve outlet)*

**Table adapted from District 12 State Route 73 Pilot Program Operation, Maintenance, and Monitoring (OM&M) Plan (Caltrans 2003*)

7.1.4 Identify Equipment and Tools Needed

The PPT should identify equipment and tools that are needed by inspection and maintenance personnel. Checklists may include, but are not limited to:

- Survey equipment
- Tape measure or other measuring device
- Flashlight
- Maintenance equipment and tools
- Rain gear
- Safety equipment, including personal protective equipment (PPE)
- Traffic control devices (must comply with California Manual on Uniform Traffic Control Devices: https://dot.ca.gov/programs/traffic-operations/camutcd
- Inspection forms and logbook; pen or pencil
- Encroachment Permit
- Business cards or other identification
- Camera

7.1.5 Create Checklists for Inspectors and Maintenance Personnel

The PPT should provide Site Inspection and Site Maintenance Forms to ensure consistent BMP O&M in conformance with routine inspection and maintenance requirements.

- Site Inspection Forms must include:
  - Date and time of inspection
  - Inspector's name
  - List of all aspects of the BMP subject to inspection and maintenance
  - Cue for required field measurements, with reference to maintenance indicator thresholds
  - Description of photographs
  - General site observations
- Site Maintenance Forms must include:
  - Date and time of maintenance work
  - Supervisor name
  - List of specific maintenance activities
  - Start, end, and total time and resources used for each activity
  - Status of completion of each activity
  - Comments

O&M field logs should be maintained to record all site visits, inspections, and maintenance activities.

7.1.6 Describe O&M Cost Collection Procedures

This section specifically focuses on the types of operation and maintenance cost data that must be collected throughout the pilot study. These data are important for evaluating the cost-effectiveness of the BMP, and for estimating future life-cycle costs for potential installations. As shown in Figure 7-2, life-cycle costs include original construction, regular and irregular maintenance, and major rehabilitation or reconstructing at the end of the design life.

Typical costs associated with BMP O&M include labor, equipment, materials, tools and utilities. During the pilot study, actual costs for operation and maintenance need to be tracked.

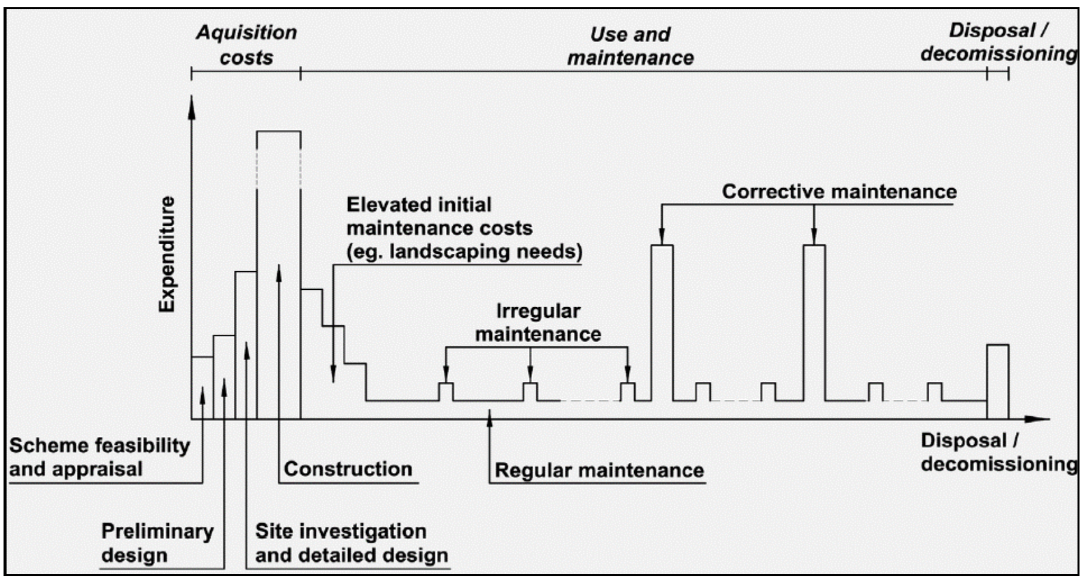


Figure 7-2 BMP Life-cycle Costs

Source: Long-Term Performance and Life-Cycle Costs of Stormwater Best Management Practices (NCHRP 2014)

O&M costs go hand-in-hand with effort needed to keep the BMP functioning (i.e., effort needed to perform activities that are triggered by maintenance thresholds). Forms must be developed for the OM&M Plan. The forms must include sections that allow the inspection and maintenance teams to track the hours and direct costs for each inspection and maintenance activity performed. Costs can be categorized as administrative, O&M, vector control, equipment, and direct costs associated with operation and maintenance. All costs except administrative define the total O&M cost. Described below is each cost category:

- *Administrative costs* are comprised of general program support/follow-up, EPs, travel, and unscheduled events. Travel costs include labor and equipment hours to and from the study site for inspection and maintenance. Costs for unscheduled events include office time to support equipment break-downs, power outages, or storm events.
- *Operation costs* are related to labor and equipment hours used for inspection and field calls. Scheduled inspections include wet season and dry season inspections of the BMP. Unscheduled inspections needed to evaluate the BMP are also included.
- *Maintenance costs* must be categorized under the subheadings of scheduled and unscheduled maintenance, vandalism, acts of God, and landscape maintenance. Scheduled and unscheduled maintenance costs can include irrigation, removal of standing water, removal of sediment, removal of trash, removal of debris, landscape management, management of structural integrity, pump servicing, cleaning of filters, and graffiti removal. Acts of God include costs for repairs to the BMP caused by severe weather, earthquakes, or other extreme acts of nature.
- *Vector control costs* include vector control and abatement and office work related to contracting VCDs.
- *Equipment costs* are associated with the time a piece of equipment is used for BMP maintenance.
- *Direct costs* are associated with VCD supplies, reproduction and postage, field supplies and minor equipment (shovels, gloves, etc.), miscellaneous equipment rental, sediment analysis, sediment disposal, and miscellaneous other direct costs.

Tracking these costs is important for a number of reasons, including evaluating the cost-benefit of a BMP, budgeting future BMP O&M cost expenditures, tracking the level of effort during the year to determine when peak staff effort is required, and for identifying opportunities to adjust maintenance activities. Figure 7-3 is an example of an Operation and Maintenance Cost Accounting Summary Form.

7.1.7 Identify Monitoring and Analysis Procedures

Monitoring and analysis are usually needed to answer the study questions and meet the study objectives. Figure 7-4 is a graphical representation of the processes required to monitor a BMP pilot study, including planning as well as data collection, verification, validation, and management. Refer to the Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a) for specific requirements and guidance.

DISTRICT: 7	LOCATION: I-605/Carson & Del Amo Avenue	SITE NO. 73255	BMP TYPE: Biofiltration Swale	CONSULTANT: Montgomery Watson - Chaudhary

Labor (hours)

TASK	1999 Jun	Jul	Aug	Sep	Oct	Nov	Dec	2000 Jan	Feb	Mar	Total (hrs)	Avg. Rate	TOTAL $
Administration													
General program support/Follow-up	3.0	5.8	8.0	4.8	3.3	7.0	9.3	7.3	4.5	15.4	68.2	$120	$ 8,184
Encroachment Permits											0.0	$87	$ -
Travel	2.7	3.8	15.4	32.4	37.0	5.1	5.5	6.8	6.4	8.6	123.6	$87	$ 10,753
Unscheduled events		1.0	8.3	1.0							10.3	$87	$ 892
Monthly Subtotal (hours)	5.7	10.6	31.6	38.1	40.3	12.1	14.8	14.1	10.9	24.0	202.0		
Monthly Subtotal ($)	$591	$1,114	$3,013	$3,471	$3,609	$1,278	$1,595	$1,465	$1,097	$2,597	$19,828		
												Task Subtotal =	$19,828
Operation													
Wet season inspections					1.0	0.5	1.5	0.8	1.3	1.5	6.5	$55	$ 358
Dry season inspections	0.5	1.0	0.5	1.0							3.0	$55	$ 165
Unscheduled inspections/field calls								1.0			1.0	$60	$ 60
Monthly Subtotal (hours)	0.50	1.00	0.50	1.00	1.00	0.50	1.50	1.75	1.25	1.50	10.50		
Monthly Subtotal ($)	$28	$55	$28	$55	$55	$28	$83	$101	$69	$83	$583		
												Task Subtotal =	$583
Maintenance													
Scheduled maintenance	1.5	1.0	0.5	5.0	11.0		14.5	6.3	6.8	3.5	50.0	$55	$ 2,750
Unscheduled maintenance			38.3	46.0	33.8	10.0	6.0				134.0	$55	$ 7,370
Vandalism											0.0	$55	$ -
Acts of God											0.0	$55	$ -
Landscape Maintenance Contractor											0.0	$0	$ -
Sediment Removal Contractor											0.0	$0	$ -
Vegetation Consultant					0.9	0.9	3.0	1.2			5.9	$75	$ 445
Other Contractor											0.0	$0	$ -
Other Contractor											0.0	$0	$ -
Monthly Subtotal (hours)	1.50	1.00	38.75	51.00	45.62	10.90	23.50	7.41	6.75	3.50	189.93		
Monthly Subtotal ($)	$83	$55	$2,131	$2,805	$2,527	$618	$1,353	$431	$371	$193	$10,565		
												Task Subtotal =	$10,565
Vector Control													
Contract & General administration	1.0	1.3	1.0	0.3				0.1	0.1	0.1	3.9	$120	$ 462
Vector prevention maint. (consultant)											0.0	$65	$ -
Response to VCD calls (consultant)											0.0	$55	$ -
VCD efforts (contracted)	6.2	9.7	8.1	6.5	5.5	5.1	8.0	5.6	6.4	5.0	66.2	$46	$ 3,074
Monthly Subtotal (hours)	7.20	10.99	9.13	6.75	5.45	5.10	8.04	5.71	6.52	5.13	70.02		
Monthly Subtotal ($)	$408	$606	$498	$332	$253	$237	$373	$273	$310	$246	$3,536		
												Task Subtotal =	$3,536

Non-Recurring Equipment (hours)

Equipment	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Total (hrs)	Avg. Rate	TOTAL $
Water Tank with Pump			38.3	46.0	33.8	10.0	6.0				134.1	$5	$ 670
Weed Wacker					2.0						2.0	$5	$ 10
Piece of Equipment 3											0.0	$0	$ -
Piece of Equipment 4											0.0	$0	$ -
Piece of Equipment 5											0.0	$0	$ -
Piece of Equipment 6											0.0	$0	$ -
Monthly Subtotal (hours)	0.00	0.00	38.25	46.00	35.80	10.00	6.00	0.00	0.00	0.00	136.05		
Monthly Subtotal ($)	$0	$0	$191	$230	$179	$50	$30	$0	$0	$0	$680		
												Equipment Subtotal =	$680

Dollars ($)

Direct Costs	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Total $
VCD supplies (direct costs less labor)											$ -
Reproduction	$ 8	$ 9	$ 13	$ 5		$ 2	$ 1	$ 1	$ 1	$ 4	$ 44
Postage/FedEx	$ 8	$ 9	$ 13	$ 5		$ 3	$ 1	$ 1	$ 2	$ 4	$ 45
Lodging			$ 5		$ 9		$ 15			$ 7	$ 35
Per Diem			$ 3		$ 6		$ 10			$ 3	$ 22
Incidentals	$ 8	$ 9	$ 13	$ 5		$ 10	$ 1	$ 1	$ 7	$ 18	$ 71
Vehicle Rental/Lease	$ 50	$ 74	$ 65	$ 10	$ 11	$ 24	$ 48	$ 113	$ 110	$ 6	$ 511
Airfare											$ -
Field Supp./Expendables	$ 10	$ 1	$ 3	$ 1	$ 24	$ 3	$ 3		$ 4	$ 9	$ 58
Equipment Rental											$ -
Sediment Analyses											$ -
Sediment Disposal											$ -
Weed Wacker	$ 15										$ 15
Vegetation Disposal											$ -
Scarifying/Hydroseeding								$2,374			$ 2,374
Storage Container			$ 4								$ 4
Monthly Subtotal	$ 99	$ 102	$ 118	$ 25	$ 50	$ 42	$ 79	$2,491	$ 124	$ 50	$ 3,180
MONTHLY TOTAL	$1,208	$1,932	$5,979	$6,918	$6,673	$2,252	$3,512	$4,760	$1,971	$3,167	**1999/2000 TOTAL = $38,371**

Figure 7-3. Example Operation and Maintenance Cost Accounting Summary Form

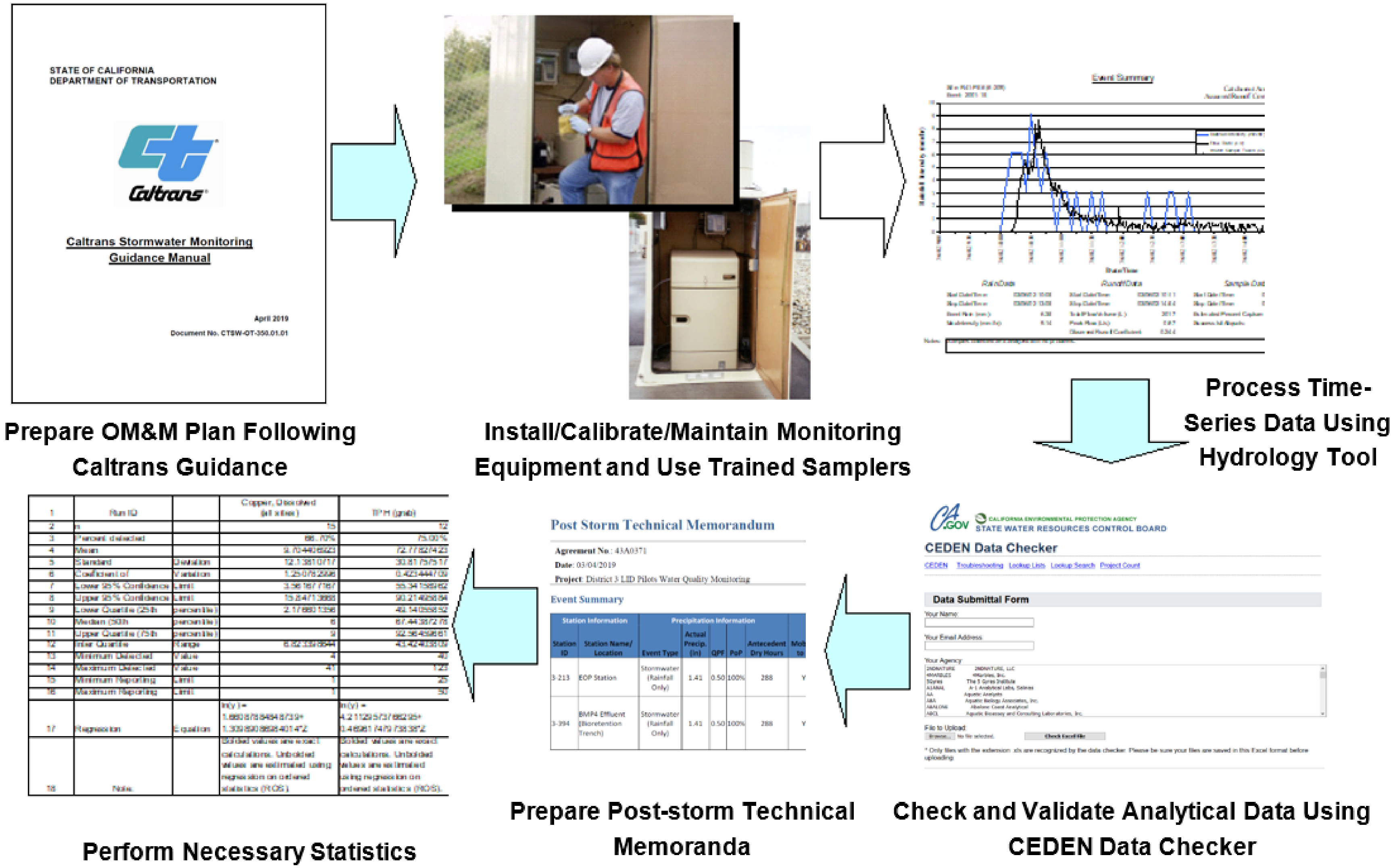


Figure 7-4. BMP Pilot Monitoring Process

7.1.8 Develop HSP

As part of the OM&M Plan, the health and safety of personnel involved in the monitoring program must be considered a high priority. Persons accessing the site for O&M and monitoring must adhere to the requirements of the HSP. Highway BMPs in particular may be placed in challenging locations and planning for human health and safety is a top priority. Some potential considerations include:

- Traffic hazards
- Wet and possible cold weather conditions
- Physical obstructions that complicate access to the site and sample collection point (e.g., steep slopes, vegetation overgrowth)
- Confined spaces (e.g., manholes that might contain toxic, explosive, or otherwise unsafe conditions)
- Flooding and fast-moving water
- Dim lighting
- Slippery conditions
- Contact with water that could be harmful (e.g., caustic, pathogenic)
- Lifting and carrying heavy and bulky pieces of equipment, including carboys and sample bottles filled with water

Based on the hazard assessment, the appropriate equipment and procedures to protect field personnel from the potential hazards must be included in the OM&M Plan. Consider adjusting monitoring locations and/or methods, if necessary, to minimize the risk of health and safety problems. Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a) provides guidance for preparing an HSP. The HSP must be an appendix of the OM&M Plan.

7.1.9 Document OM&M Plan

Once the above activities have been completed, they must be documented in an OM&M Plan. Appendix J includes details that must be provided for various elements of the OM&M Plan. Most of the components of the OM&M plan are straightforward can can be copied from the Study Plan TM, Site Selection TM, and other documents developed for the study, as described throughout this document. Additional site-specific information that should be included in the site decriptiona and access sections of the OM&M Plan include:

- Coordination with appropriate Caltrans personnel
- Caltrans monitoring site ID
- Location and post mile, with site map

- Driving directions from the Caltrans District office
- City or county jurisdiction
- Regional Water Board jurisdiction
- Parking location
- Traffic control requirements
- Gates, locks, keys, combinations
- Sensitive habitat or species
- Safety considerations
- Other features of the site, including distance from roadway(s), slopes, ground cover, overhead concerns, etc. (include photographs of the site)
- Notification requirements
- Other requirements that may be specified in an EP

Also include photographs of the site and site plans/location maps, where possible.

Note that sometimes a Quality Assurance Project Plan (QAPP) and Sampling and Analysis Plan (SAP) are developed for a study, as described in Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a). If such plans exist, they can be attached to the OM&M Plan to cover the monitoring elements of the study. Alternatively, if the only maintenance required for the pilot study is that for monitoring equipment, the QAPP and SAP may be used in lieu of developing an OM&M Plan.

The OM&M Plan must be reviewed and approved by the Caltrans Task Order Manager.

7.2 OM&M Plan Implementation

Upon completion of the OM&M plan, approval by the Caltrans Task Order Manager, and execution of a monitoring task order (see Section 7.4), monitoring equipment must be installed and tested, operations and maintenance must be conducted, samples must be collected and analyzed, and monitoring data must be retrieved, recorded, and documented according to the details described in the OM&M plan and monitoring task order.

7.3 OM&M Plan Optimization

Ideally, the systematic planning process can be iterative for each study. It is possible to modify study parameters, based on collected data and field observations, to optimize data collection to meet study objectives. This can sometimes be done during natural break points of the study (e.g., at the end of the wet season), when data are analyzed and achievement of study objectives is assessed. In some cases, the collected information may indicate the need to modify goals, collection methods, the analytical constituent list, or types/frequency of storms sampled. Interim reports for the study must document such recommendations (see Chapter 8).

As discussed later in this manual (Chapter 8), interim reports are to be written annually for ongoing studies. At that time, the Study Plan TM must be consulted and checked against the field experience. The purpose of this check is to assure that the experimental activities are adhering to the Study Plan TM. Of particular importance is checking whether assumptions made in the Study Plan TM are proving to be true. One example is checking whether the runoff is originating from the highway or whether there is intermingling with extraneous flows. Another example might be an assumption that infiltration or some other parameter is negligible. Preliminary statistical analysis must be performed to check the assumptions made of the variance in the data used to estimate the number of samples needed. If the actual variance is smaller than the assumed value, the study might be shortened. If the variance is larger than assumed, additional sampling may have to be done. Based on field experience, it might also be advisable to modify sample collection methods, the analytical constituents list, or the storm event criteria. Care must be exercised, however, to avoid introducing changes that interrupt the continuity of the data collection, making data collected at different times in the study incompatible. An example of this might be a decision to drop total suspended solids (TSS) in favor of turbidity as a measure of solids in the runoff. Another example might be radically changing the location or protocol of sampling. With proper caution, alternative sampling and analysis designs may improve the quality and/or cost-effectiveness of the study.

7.4 Task Order Development

Task orders for conducting OM&M activities must include the key components described in Table 7-2.

Table 7-2. Monitoring Task Orders and Component Descriptions

Task	Description	Components
1	Kick Off Meeting	Discuss scope, schedule, deliverable due dates, and project budget; distribute relevant documents; introduce responsible personnel (Caltrans, consultants, academia); establish chain-of-command for communication; discuss lessons learned; conduct site walk.
2	Review Existing Documents	Review documents (e.g., PS&Es; Design Reports; Study Plan TM; Site Selection TM; existing permits; BMP post-construction report).
3	Obtain Permits	Obtain permits necessary to install monitoring equipment and conduct the study activities, including encroachment and other agency permits that may not have yet been obtained (see Chapter 4)
4	Prepare OM&M Plan	Prepare OM&M Plan that defines O&M and monitoring procedures.
5	Acquire Monitoring Equipment	Acquire equipment used by Caltrans previously, where available. If sufficient equipment is not available, purchase new equipment.
5.1	Install Monitoring Equipment	Install monitoring equipment. Conduct field calibration checks to ensure equipment operability after the equipment is installed.
5.2	Maintain Monitoring Equipment	Perform routine inspection, maintenance, and calibration. Inform Caltrans of major equipment and/or maintenance issues.
6	Storm Sampling	Identify number of storm events and period of monitoring. Collect time series data throughout the pilot study.
7	Operation and Maintenance of BMP	Conduct BMP O&M in accordance with the study's OM&M Plan
8	VCD Contracting	Contract with a Vector Control District for the inspection and abatement of vectors, if necessary.
9	Data Management	Manage laboratory data and OM&M cost data.
10	Post-Storm Reporting	Prepare and submit a Post-Storm TM to Caltrans for each sampling event. Include storm conditions, equipment performance, hydrographs, and other required records.
11	Equipment Demobilization	Remove the monitoring equipment from the stations at the end of the study period. Coordinate with Caltrans to move the removed monitoring equipment to the Caltrans equipment inventory.
12	Interim and Final Reporting	Submit Interim and Final Reports in accordance with the study's specified report frequency and content. (See Chapter 8).
13	Project Management and Coordination	Schedule meetings with Caltrans as necessary to ensure a common understanding of the project status.

Chapter 8 Interim and Final Reports

The results of a pilot study must be presented in interim reports and a final report. Interim reports are prepared annually at the end of each monitoring season and summarize the season's activities and preliminary data. The Final Report is developed at the end of several seasons of monitoring, when monitoring is complete. It documents a comprehensive review of all activities, data, and conclusions from the entire study period.

Chapter Purpose and Desired Outcome

Chapter Purpose: Guide the PPT in documenting pilot study activities and findings

Desired Outcome: Study report that accurately documents the study objectives, assumptions, activities, results, and conclusions

Typical outlines for an interim or final report include:

- Report cover
- Executive summary
- Introduction
- BMP and site descriptions
- Data collection and analysis
- Data results, evaluations, and findings
- Conclusions
- References
- Appendices

All reports will need to comply with the ADA. Check with the Task Order Manager for the latest Caltrans policies and practices.

8.1 Report Cover

Reports must include the following information on the document cover:

- Caltrans logo
- Title of report
- Date of issue (month and year)
- Document ID number (check with DEA for convention)
- State of California Departmnet of Transportation
    - Division of Environmental Analysis
    - Office of Stormwater Program Development

- 1120 N Street Sacramento, California 95814
- https://dot.ca.gov/programs/environmental-analysis

8.2 Executive Summary

Provide a concise overview of the pilot study. This is typically created after the remainder of the report has been completed, allowing you to copy excerpts from the main report. Doing so saves time and ensures consistency. The executive summary will include:

- The problem statement and study goal
- The study questions and objectives
- Summaries of the site and BMP features
- Summaries of the data collection effort and analytical results
- Conclusions, including if and how the study objectives were met

8.3 Introduction

The introduction must discuss the following items:

- Problem description (or need) and study goal (taken from the Study Plan)
- Study questions and objectives (taken from the Study Plan)
- Type and number of BMPs evaluated
- General location of the BMPs (e.g., statewide, within District 12, along SR 73)
- Report organization

8.4 BMP and Site Descriptions

8.4.1 BMP Description

As was done for the Study Plan TM, describe the BMP, modifying the description to cover any changes from the originally intended design. Reference the Design Report so the design calculations and other design features and assumptions are understood. Include schematics to visually demonstrate the concepts and BMP characteristics, such as:

- Footprint
- Slope
- Media characteristics (e.g., type, grain size, uniformity coefficient, depth, operating range of infiltration rates)
- Depth of standing water
- Vegetation characteristics
- Other BMP characteristics or variables that can impact BMP performance or operation

8.4.2 Site Descriptions

Describe the sites where BMPs were implemented and monitored. This should include:

- Facility address or roadway names and intersections
- City
- Latitude and longitude
- Elevation
- Caltrans District
- Regional Water Board region
- Drainage area and slope
- Traffic volume (AADT)
- Vegetation/vegetative canopy
- Percent imperviousness
- Soil characteristics (hydrologic soil group, classification, saturated conductivity, grain size, compaction, etc.)
- Representativeness of inflows
- 85th percentile, 24-hour storm event
- Other site characteristics or variables identified that can impact the study

Include a plan view (e.g., a screen shot from Google Earth) that depicts where the BMP is located within the site and other relevant features, such as streams, buildings, roadways, or sources of commingled flow.

8.5 Data Collection and Analysis

8.5.1 Data Collection

Describe the following, summarizing and referencing the Sampling and Analysis Plan or Monitoring Plan, as appropriate:

- Study type (include a schematic of monitoring locations relevant to the BMP)
- Mobilization criteria
- Sample collection process (grab or automated) and associated equipment
- Flow and precipitation measurement devices
- Calibration procedures
- Sampling handling
- Chain-of-custody protocols
- Operation and maintenance recordkeeping
- Procedures for collecting life-cycle costs

8.5.2 Data Analysis

Describe the following, summarizing and referencing the Sampling and Analysis Plan or Monitoring Plan as appropriate:

- Laboratory analytical methods (e.g., ASTM) for each constituent of concern
- DQOs and other quality assurance/quality control information
- Procedures for analyzing life cycle costs

8.6 Data Results, Evaluations, and Findings

8.6.1 Results

Summarize the following data, as applicable, including tables with summary statistics and figures where appropriate:

- Data used to evaluate feasibility criteria identified in the Study Plan
- Data used to evaluate O&M criteria identified in the Study Plan
- Data used to evaluate treatment performance criteria identified in the Study Plan
    - Monitoring dates
    - Storm type (e.g., rainfall, snow, snowmelt)
    - Precipitation depths, volumes, and intensities as appropriate
    - Flow rates and volumes as appropriate
    - Laboratory results
- Data used to evaluate life-cycle cost criteria identified in the Study Plan
    - Capital costs
    - O&M costs for the life of the project

8.6.2 Evaluations and Findings

Document how the data were evaluated to address the study objectives and what the evaluations indicate. This includes providing equations, statistical methods, tables, box plots, and other graphics, as well as accompanied narratives to demonstrate your methodology, considerations and reasoning. It often helps to organize this section according to study objectives, with a different subsection devoted to each study objective.

Evaluate the BMP’s life-cycle costs using the method documented in Caltrans Life-cycle Costs Based on 2007 BMP Operation and Maintenance Cost Analysis Technical Memorandum (Caltrans 2008) and summarize the activities, assumptions, and findings. Then, tabulate the proposed BMP’s expected life-cycle costs along with those for approved BMPs, including present worth for capital and O&M costs. Plot costs per acre treated versus expected average annual load reduction for a typical sized BMP for the proposed BMPs and approved BMPs. Describe how results might influence BMP selection guidance. Section 9.0 of Appendix K of this PSGM provides an example.

8.7 Conclusions

8.7.1 Address the study questions and objectives

Based on the results, findings, and criteria stated in the study objectives, answer the study questions, providing qualifying statements or exceptions as appropriate. This should not be a restatement of the study findings, but summaries of what you can conclude based on the data collected, as it relates to each study question and objective. Recall that the study questions and objectives should have been (during development of the Study Plan) based on Caltrans criteria for approving BMPs as stated in 2016 SWMP (Caltrans 2016), specifically in regards to technical feasibility, O&M requirements, treatment performance, and life-cycle costs. Pilot studies may address all BMP approval criteria or a select few, depending on the study scope. To ease review of this section, organize the study questions and objectives according to these categories. This section should state whether the BMP should be recommended for approval and inclusion in the Caltrans BMP toolbox.

8.7.2 Lessons Learned

Summarize the lessons learned regarding siting, design, construction, and monitoring so it may be recorded and disseminated to others. This will help promote the recurrence of desirable outcomes and minimize the recurrence of undesirable outcomes. They should draw on both positive and negative experiences. Examples include good ideas that prevent problems and reduce costs and lessons learned from undesirable outcomes.

General questions to consider when developing the lessons learned include the following:

- What went right (i.e., as planned)?
- What went wrong (i.e., not as planned)?
- What unexpected events occurred?
- What could be done differently in the future (to save time, reduce costs, or improve performance, monitoring, or safety)?
- What should be done differently in the future (to avoid delays, cost escalations, or performance/monitoring problems)?
- Where were significant resources focused ineffectively?
- What areas/tasks received more attention than necessary?
- What areas/tasks need more attention in the future?

Specific lessons learned during the siting, design, construction, and monitoring phases of the pilot study are to be included in the Site Selection TM, Final Design Report, Construction Report, and Post-Storm TMs, respectively. At the conclusion of the pilot study, the lessons learned from these individual documents are reviewed from an overall project viewpoint for completeness and content, and compiled for inclusion in the Pilot Study Final Report. Lessons applicable for future study, should be highlighted for future updates to this document.

Sample lists of some lessons learned from past pilot studies are presented in Table 8-1.

Table 8-1. Example Lessons Learned

Phase	Lesson Learned
Siting	Avoid sites with potential base flow.
Siting	Ensure BMPs have proper aesthetics appeal, especially in the Lake Tahoe area.
Construction	Make sure flumes installed by the Contractor are perfectly level.
Construction	Inform the Resident Engineer (RE) of features not built per plan as soon as they are discovered and not when construction is complete. (Coordinate site visits with the RE on completion of grading and on completion of setting concrete forms and elevations)
Construction	Ensure quality control during construction for drainage items with minimal slopes.
Design	Present PS&E submittals to District staff in a meeting to facilitate the review process.
Design	For projects with long duration construction schedules, incorporate appropriate escalation factors into the engineer's estimate.
Design	Non-roadway specialty items may take a long time to procure and should be considered when preparing the Special Provisions.
Design	Do not include vegetation removal in the PS&E without prior coordination with District environmental staff.
Design	Specify locally-available material to avoid procurement problems.
Design	Engage the monitoring consultant early in the design phase to facilitate the proper location, design, and O&M of the BMP.
Design	Account for soil uplift and vegetation growth in the transition between impervious and pervious surfaces to maintain the desired flow path
Monitoring	Flumes sized for maximum monitoring events were not capable of rating low flows. Consider the use of smaller flumes to monitor smaller, more typical flows. Consider using larger flumes or weirs to measure higher flows. Use the Caltrans SOPs for flow measurement verification (Caltrans 2010c)
Monitoring	Use caution if modifying mobilization criteria during the study to avoid biasing dataset (systematic exclusion of regulated storms of a certain type)
Monitoring	Ensure laboratory is able to meet project reporting limits.
Monitoring	Use customized coolers to prevent bottle breakage.
Monitoring	Place sampler close to monitoring point so it can adequately draw samples within the capability of the monitoring pump.

8.8 References

Cite all references used in conducting the study, including design, construction, monitoring, analysis, and data evaluation. Common references may include:

- Study Plan TM (and lessons learned, if created)
- Site Selection TM
- Design Report (and lesson learned, if created)
- Construction Report (and lessons learned, if created)
- As-Built Plans
- OM&M Plan
- Data validation results
- O&M lessons learned
- Caltrans Stormwater Monitoring Guidance Manual (Caltrans 2019a)

8.9 Appendices

Include appendices to provide copies of raw data and supporting documents developed during the course of the pilot study. At a minimum, the appendices should include, as applicable:

- Design Report
- Construction Report
- Stormwater Data Report
- Hazardous Waste/ADL Compliance report
- Monitoring reports (with monitoring data and PSTMs as appendices)
- Construction cost data and cost summary
- Maintenance cost data and cost summary
- O&M data and inspection forms (including vector abatement)

Chapter 9 Stormwater Advisory Team Package

The results of the pilot study must be summarized into a package to be reviewed by the various SWATs for potential approval. The package must include sufficient information to demonstrate to the SWATs that the BMP meets all BMP evaluation criteria identified in the 2016 SWMP (Caltrans 2016; copied in Section 2.0 of Appendix C of this PSGM).

Chapter Purpose and Desired Outcome

Chapter Purpose: Guide the PPT in documenting BMP characteristics, performance and life-cycle costs.

Desired Outcome: SWAT package that provides all information needed by the SWATs to approve a BMP for use in the Caltrans ROW.

To do so, the package must include the following:

- Problem statement
- BMP fact sheet
- Performance summary
- Siting recommendations
- Design recommendations
- Recommendations for specifications and special provisions
- O&M recommendations
- Life-cycle costs

9.1 Problem Statement

Copy the problem statement developed in Step 1 of the study planning process (Section 0), modifying it as appropriate to address adjustments that were made due to changing conditions, such as new regulatory requirements.

Section 1.0 of Appendix K of this PSGM provides an example.

9.2 Pilot Study Summary

The pilot study summary gives the SWATs an overview of the BMP that was studied and the amount and type of data collected. Details to provide include:

- Background
    - Location
    - Schematic
    - Project plan sheets and redlines

- Design calculations for NPDES permit compliance

- Monitoring activities
- Results (refer to Section 9.4, Performance Summary, below)
- Alternatives (if relevant)

Section 2.0 of Appendix K of this PSGM provides an example.

9.3 BMP Fact Sheet

The fact sheet provides an overview of the BMP to give SWATs an initial understanding of its function and layout, as well as design, construction, and O&M requirements. The fact sheet includes:

- A schematic of the BMP
- A narrative description
- Relevant treatment mechanisms
- A treatment performance summary table
- Advantages
- Constraints
- General siting (placement), design, construction, and O&M considerations
- References

Section 3.0 of Appendix K of this PSGM provides an example, as does Caltrans Treatment BMP Technology Report (Caltrans 2018b).

9.4 Performance Summary

Provide a summary of the BMP's performance that includes the following:

- Results of Caltrans pilot studies
    - Study locations
    - Period of study and number of storms
    - At least one of the following:
        - Table of percent removal of each consistent of concern for each monitoring event, including mean and median percent removals for all events for each constituent and a comparison of influent values to regional or statewide discharge characterization statistics (see Caltrans 2009)
        - Table of influent (or control) and effluent summary statistics, including common percentiles, and comparison of influent quality to regional or statewide discharge characterization statistics (see Caltrans 2009)

- Narrative describing which constituents were effectively removed, moderately removed, poorly removed, or variably removed by the BMP
- Results of performance-based ranking analysis for each relevant constituent (see Appendix L)
- Limitations or qualifiers related to the study findings (e.g., unrepresentative influent flows, unexpected base flow, BMP liner damaged, BMP construction issues, BMP operational issues)
- Results of performance reported by others (e.g., International BMP database)

Section 4.0 of Appendix K of this PSGM provides an example.

9.5 Siting Recommendations

Caltrans Division of Design Stormwater Program will use the siting recommendations to incorporate the BMP into BMP selection processes in the PPDG (Caltrans 2019e). Recommendations typically include:

- A narrative of target applications that describes the intended use of the BMP (statewide, certain TMDLs, where other approved BMPs are infeasible, etc.)
- Siting constraints
  - Drainage area and runoff sources
  - Slope/elevation change/hydraulic head
  - Power availability
  - Soils and groundwater
  - Site geometry and available space
  - Traffic
  - Proximity to structures (e.g., buildings, roads, utilities)

Section 5.0 of Appendix K of this PSGM provides an example.

9.6 Design Recommendations

Caltrans Division of Design Stormwater Program will use the design recommendations to incorporate the BMP's design elements into the PPDG (Caltrans 2019e) and Design Guidelines. The recommendations typically include:

- A schematic and narrative description of the BMP (copy from the fact sheet)
- Typical plan and cross section, with appropriate elements labeled
- Table of factors affecting preliminary design (i.e., siting and design criteria)
  - O&M access requirements
  - Pretreatment requirements
  - Components

- Sizing
- Construction considerations (e.g., specialized equipment needed to minimize soil compaction or meet strict slope tolerances)
- Limits to allowable backwater inundation (see Caltrans HDM, https://dot.ca.gov/programs/design/manual-highway-design-manual-hdm)

Section 6.0 of Appendix K of this PSGM provides an example.

9.7 Recommendations for Specifications and Special Provisions

Identify which of Caltrans existing specifications and SSPs (https://dot.ca.gov/programs/design) are applicable to the BMP and list them in the Recommendations for Special Provisions section of the SWAT Package. If there are none, or they need to be highly modified to meet the design, maintenance, or other required aspects of the BMP (i.e., when the payment and measurement clauses are changed or different materials or installation methods are specified), nSSPs must be developed, if not done so during design. The nSSPs provide instructions on how SSPs need to be modified or added to, and should identify unique testing and submittal requirements. If nSSPs were developed during design of the pilot study, modify them as appropriate based on the study results.

Section 7.0 of Appendix K of this PSGM provides an example.

9.8 O&M Recommendations

O&M guidance will be used by Caltrans Division of Maintenance Stormwater Program to develop appropriate materials for Caltrans Maintenance Staff Guide (contact the Task Order Manager for the latest version). The guidance must include:

- ***Introduction***: State that the section summarizes the BMP's O&M needs
- ***Appropriate applications***: Reference tables that summarize the inspection and O&M requirements to maintain the minimum water quality function that is necessary to justify the use of the BMP
- ***Preventive Maintenance Table***: Table summarizing the preventive maintenance activities. The table should include three columns:
    - *Frequency*: frequency of the activity
    - *Routine action:* specific action to be performed
    - *Activity cut sheet:* reference existing Caltrans Activity Cut Sheets (Appendix B of Caltrans Maintenance Staff Guide) or summarize how to conduct the activity
- ***Corrective Maintenance Table***: Table summarizing the corrective maintenance activities. The table should include three columns:

- *Maintenance indicator:* condition under which the activity would be required
- *Inspection frequency:* frequency for which the condition should be checked
- *Corrective maintenance activities*: activities to be conducted to correct the condition. Several activities may be required for one condition.

- ***Schematic, plan, and/or cross section:*** copy from fact sheet and siting and design guidance

Section 8.0 of Appendix K of this PSGM provides an example.

9.9 Life-cycle Costs Comparison

Summarize the life-cycle costs evaluated and documented in the final study report, including tables of the proposed BMP's expected life-cycle costs along with those for approved BMPs, including present worth for capital and O&M costs. Also provide plots of costs per acre treated versus expected average annual load reduction for a typical sized BMP for the proposed BMPs and approved BMPs. Describe how results might influence BMP selection guidance.

Section 9.0 of Appendix K of this PSGM provides an example.

9.10 References

List all relevant references used to develop the SWAT package.

Section 10.0 of Appendix K of this PSGM provides an example.

23 U.S.Code §327. *Surface Transportation Project Delivery Pilot Program*. 2010.

American Society of Civil Engineers and US Environmental Protection Agency (ASCE and EPA 2000). *Determining Urban Stormwater Best Management Practice (BMP) Removal Efficiencies, Task 1.1 – National Stormwater BMP Databased Data Elements*. July 2000.

American Society of Civil Engineers (ASCE 2002). *Urban Stormwater BMP Performance Monitoring*. EPA/821/B-02/001. U.S. Environmental Protection Agency, Office of Water. 2002

California Department of Transportation. (Caltrans 2002). *Memorandum on Guidelines on the Issue of Encroachment Permits for Architectural and Engineering Consultants*. 2002.

California Department of Transportation. (Caltrans 2003a). *District 12 State Route 73 Pilot Program, Operation, Maintenance, and Monitoring (OM&M) Plan, Version 1.0*. CTSW-RT-03-078.51.35. October 2003.

California Department of Transportation (Caltrans 2003b). *Discharge Characterization Study Report*. CTSW-RT-03-065.42. November 2003.

California Department of Transportation. (Caltrans 2004). *BMP Retrofit Pilot Program Final Report*. 2004.

California Department of Transportation (Caltrans 2007). *Project Planning and Design Guide (PPDG).* CTSW-RT-07-172.19.1. March 2007.

California Department of Transportation (Caltrans 2007b). *Final Linear Filtration Alternatives Pilot Study Plan Technical Memorandum.* CTSW-TM-07-172.27.5. October 2007.

California Department of Transportation. (Caltrans 2008). *Caltrans Life Cycle Costs Based on 2007 BMP Operation and Maintenance Cost Analysis Technical Memorandum. 2008*. CTSW-TM-08-176-11.1.

California Department of Transportation. (Caltrans 2009). *Discharge Characterization Report Study Update*. CTSW-RT-09-182.44.2. January 2009.

California Department of Transportation (Caltrans 2010a). *Performance-Based Ranking of Treatment BMPs*. CTSW-RT-10-239.01. February 2010.

California Department of Transportation (Caltrans 2010b). *Project Planning and Design Guide (PPDG).* CTSW-RT-10-254.03. July 2010.

California Department of Transportation. (Caltrans 2010c). *Standard Operating Procedures for Stormwater Flow Measurement Verification*. CTSW-RT-10-211.19.01. November 2019.

California Department of Transportation. (Caltrans 2016). *Caltrans Statewide Stormwater Management Plan*. 2016.

California Department of Transportation (Caltrans 2017). *Project Planning and Design Guide (PPDG)*. CTSW-RT-17-314.24.1 July 2017.

California Department of Transportation. (Caltrans 2018a). *Encroachment Permits Manual*. April 2018. https://dot.ca.gov/programs/traffic-operations/ep/ep-manual

California Department of Transportation. (Caltrans 2018b). *Stormwater Treatment BMP Technology Report*. 2018.

California Department of Transportation. (Caltrans 2018c). *Workplan Standards Guide for the Delivery of Capital Projects, Release 12.0*. 2018. https://dot.ca.gov/programs/project-management/workplan-standards-guide

California Department of Transportation. (Caltrans 2018d). *Caltrans Plans Preparation Manual*. 2018. https://dot.ca.gov/programs/design/manual-plans-preparation-manual-ppm

California Department of Transportation. (Caltrans 2018e). *Caltrans Highway Design Manual*. 2018. https://dot.ca.gov/programs/design/manual-highway-design-manual-hdm

California Department of Transportation (Caltrans 2018f). *Performance-Based Ranking of Treatment BMPs, Addendum #1.* CTSW-RT-18-350.04.01. June 2018.

California Department of Transportation. (Caltrans 2019a). *Stormwater Monitoring Guidance Manual*. CTSW-OT-19-350.04.01. 2019.

California Department of Transportation. (Caltrans 2019b). *Project Development Procedures Manual*. September 2019. https://dot.ca.gov/programs/design/manual-project-development-procedures-manual-pdpm

California Department of Transportation. (Caltrans 2019c). *Caltrans Standard Environmental Reference*. Accessed 2019. https://dot.ca.gov/programs/environmental-analysis/standard-environmental-reference-ser

California Department of Transportation. (Caltrans 2019d). *Construction Contract Development Guide*. Version 5.0. July 2019. http://ppmoe.dot.ca.gov/hq/esc/oe/rtl/ccdg/CCD-Guide.pdf

California Department of Transportation. (Caltrans 2019e). *Project Planning and Design Guide*. Updated April 2019. https://dot.ca.gov/programs/design/manual-project-planning-design-guide

California Department of Transportation. (Caltrans 2019f). *Caltrans Design Memorandum webpage*. Accessed 2019. https://dot.ca.gov/programs/design/design-memoranda

California State Water Resources Control Board. (State Water Board 2012). Order No. 2012-0011-DWQ-EXEC NPDES Permit No. CAS000003, *National Pollutant Discharge and Elimination System (NPDES) Statewide Stormwater Permit Waste Discharge Requirements (WDRs) for State of California Department of Transportation*. 2012.

California State Water Resources Control Board. (State Water Board 2017). Order No. 2012-0011-DWQ, as amended by Order WQ 2014-0006-EXEC, Order WQ 2014-0077-DWQ, Order WQ 2015-0036-EXEC, and Order WQ 2017-0026-EXEC NPDES Permit No. CAS000003, *National Pollutant Discharge and Elimination System (NPDES) Statewide Stormwater Permit Waste Discharge Requirements (WDRs) for State of California Department of Transportation*. 2017.

Coffey, S.W., J. Spooner, and M.D. Smolen. (Coffey et al. 1993). *The Nonpoint Source Manager's Guide to Water Quality and Land Treatment Monitoring*. NCSU Water Quality Group, Department of Biological and Agricultural Engineering, Raleigh, NC. 1993.

Currier, B., Bonham, J. (Currier and Bonham 2019). *Load Removal Efficiency: Derivation of Load Removal Efficiency from Volume and Concentration Removal*. *Water Math Concepts*, No. 6. Office of Water Programs at Sacramento State. Sacramento, CA. 2019.

Gibbons, Robert D. (Gibbons 1994). *Statistical Methods for Groundwater Monitoring*. John Wiley & Sons, Inc., New York

Gilbert, Richard O. (Gilbert 1987). *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Hassig, N.L., R.O. Gilbert, J.E. Wilson, B.A. Pulsipher, and L.L. Nuffer (Hassig et al. 2005). *Visual Sample Plan 4.0 User's Guide.* PNHL-15247. Pacific Northwest National Laboratory, Richland, Washington, prepared for the U.S. Department of Energy. http://dqo.pnl.gov/VSP

Helsel, Dennis R. (Helsel 2005). *Non-detects and Data Analysis. Statistics for Censored Environmental Data*. John Wiley & Sons. New Jersey

ISCO. (ISCO 2006). *Isco Open Channel Flow Measurement Handbook*. 2006.

National Cooperative Highway Research Program (NCHRP 2014). *Long-Term Performance and Life-Cycle Costs of Stormwater Best Management Practices*. Report 792. 2014.

U.S. Department of Energy. (USDOE 2005). *Visual Sample Plan Version 4.0 User's Guide*. PNNL-15247.

U.S. Environmental Protection Agency (USEPA 1994). Statistical Methods for Evaluating the Attainment of Cleanup Standards. EPA 230-R-94-004.

U.S. Environmental Protection Agency (USEPA 2006a). *Data Quality Assessment: Statistical Methods for Practitioners*. EPA QA/G-9S.

U.S. Environmental Protection Agency (USEPA 2006b). *Guidance on Systematic Planning Using the Data Quality Objectives Process*. EPA QA/G-4.

US Environmental Protection Agency (USEPA), Water Environment Research Foundation, Federal Highway Administration, Environmental and Water Resources Institute of the American Society of Civil Engineers (USEPA et al. 2009). *Urban Stormwater BMP Performance Monitoring*. October 2009.

U.S. Geological Survey (USGS 2005). *Computer Program for the Kendall Family of Trend Tests*. Scientific Investigations Report 2005-5275.

U.S. Geological Survey (USGS 2020). *Statistical Methods in Water Resources: U.S. Geological Survey Techniques and Methods, Book 4, Chapter A3*. 2020. https://pubs.er.usgs.gov/publication/tm4A3

Washington State Department of Ecology. (ECY 2008). *Guidance for Evaluating Emerging Stormwater Treatment Technologies, Technology Assessment Protocol – Ecology (TAPE)*. Revised January 2008.

Appendix B: Abbreviations and Acronyms

AADT	Average Annual Daily Traffic
ACM	Asbestos-Containing Materials
ADA	Americans with Disabilities Act
ADL	Aerially Deposited Lead
ASCE	American Society of Civil Engineers
ASTM	American Society for Testing and Materials
BCDC	Bay Conservation and Development Commission
BEES	Basic Engineering Estimating System
BIA	Bureau of Indian Affairs
BMP	Best Management Practice
Caltrans	California Department of Transportation
CCA	Construction Contract Acceptance
CCC	California Coastal Commission
CCD	Construction Contract Development Guide (Caltrans 2018a)
CCO	Contract Change Order
CDFW	California Department of Fish and Wildlife
CE	Categorical Exemption or Categorical Exclusion
CEC	Certificate of Environmental Compliance
CEE	Chief Environmental Engineer
CEQA	California Environmental Quality Act
CRZ	Clear Recovery Zone
cy	cubic yards
DCM	Design Changes Memoranda
DEA	Caltrans Division of Environmental Analysis
DOD	Caltrans Division of Design
DQO	Data Quality Objective
DTSC	California Department of Toxic Substances Control
DWR	California Department of Water Resources
EA	Environmental Assessment
ECR	Environmental Commitments Record
ED	Environmental Document
EIR	Environmental Impact Report
EIS	Environmental Impact Statement
EP	Encroachment Permit

EPA	US Environmental Protection Agency
ft	feet
FHWA	Federal Highway Administration
HDM	Highway Design Manual
hr	hour
HQ	Caltrans Headquarters
HSP	Health & Safety Plan
IDF	Intensity Duration Frequency
IS	Initial Study
ISA	Initial Site Assessment
km	kilometer
LCCA	Life-Cycle Cost Analysis
m	meter
MEP	Maximum Extent Practicable
mm	millimeter
MMRR	Mitigation Monitoring and Reporting Record
MOU	Memorandum of Understanding
mph	miles per hours
NEPA	National Environmental Policy Act
NOA	Naturally Occurring Asbestos
NOC	Notice of Construction
NPDES	National Pollutant Discharge Elimination System
NRCS	Natural Resources Conservation Service
nSSP	Non-Standard Special Provisions
O&M	Operations and Maintenance
OM&M	Operations, Maintenance and Monitoring
PAM	Permits, Agreements and Mitigation
PCR	Post-Construction Report
PDPM	Project Development Procedures Manual
PPDG	Caltrans Project Planning and Design Guide
PPE	Personal Protective Equipment
PPT	(BMP) Pilot Project Team
PS&E	Plans, Specifications and Estimate
PSGM	Pilot Study Guidance Manual

PSTM	Post-Storm Technical Memorandum
QAPP	Quality Assurance Project Plan
RCRA	Resource Conservation and Recovery Act
Regional Water Board	California Regional Water Quality Control Board
RFC	Request for Clarification
RFI	Request for Information
RVTS	Roadside Vegetated Treatment Site
SAP	Sampling and Analysis Plan
SER	Standard Environmental Reference
SEZ	Stream Environment Zone
SHS	State Highway System
SR	State Route
SSP	Standard Special Provisions
SWAT	Stormwater Advisory Team
SWDR	Stormwater Data Report
SWMP	Stormwater Management Plan
SWPPP	Stormwater Pollution Prevention Plan
State Water Board	California State Water Resources Control Board
TM	Technical Memorandum
TMC	Traffic Management Control
TMDL	Total Maximum Daily Load
TRPA	Tahoe Regional Planning Agency
TSS	Total Suspended Solids
USACOE	US Army Corps of Engineers
USFWS	US Fish and Wildlife Service
VCD	Vector Control District
WDR	Waste Discharge Requirements
WPCP	Water Pollution Control Program
WQF	Water Quality Flow
WQV	Water Quality Volume

Appendix C Contents

1.0 Treatment Mechanisms

Table C- 1 tabulates common mechanisms applicable to Caltrans BMP implementation, as well as key factors that affect each mechanism.

Table C- 1. Treatment Mechanisms and Key Controlling Factors

(Likely [L] or Unlikely [U])

Treatment Mechanism	BMP Footprint [1]	Hydraulic Loading Rate	Media Characteristics [2]	Subsurface Infiltration Rate[3]	Influent Water Quality	O&M Practices	Antecedent Dry Period	Slope	Vegetation Type	Extent of Vegetative Coverage
Infiltration	L	L	L	L	L	L	L	L	L	L
Trash Capture	L	L	U	U	L	L	U	U	U	L
Filtration	L	L	L	L	L	L	L	L	L	L
Adsorption	L	L	L	L	L	L	U	U	U	U
Evapotranspiration	L	L	L	L	U	L	L	L	L	L
Rainfall Harvest[4]	L	L	U	U	U	L	U	U	U	U
Sedimentation	L	L	U	U	L	L	U	L	U	L
Flotation	L	L	U	U	L	L	U	U	U	U
Plant Uptake	L	L	U	L	L	L	L	U	L	L
Biochemical Transformation	L	L	L	U	L	L	L	U	L	L

[1] *Length, width, area*
[2] *Type, infiltration rate (3), depth, grain size, uniformity coefficient, age, etc.*
[3] *For design, often represented by saturated hydraulic conductivity*
[4] *Rain barrel, cisterns, and other storage for later user*

2.0 Selecting Relevant BMP Evaluation Criteria

Table C- 2 lists the BMP evaluation feasibility, O&M, treatment performance, and/or life-cycle cost criteria from Caltrans SWMP that apply to the study goal and BMP characteristics. Not all criteria may apply.

Table C- 2. Caltrans BMP Evaluation Criteria (Section 4.2.1 of 2016 SWMP)*

**Text below is exact citation from 2016 SWMP*

Criteria Type	Feasibility Criteria
Feasibility	• The BMP will function under one or more climatic, geological, and topographical conditions encountered at Department roadways and facilities. Except for initial installation and vegetation establishment periods, irrigation, or supplemental water should not be required. • The BMP will be able to be sited so it complies with the safety requirements of the Caltrans Highway Design Manual. • Products that require handling by Caltrans employees must comply with applicable health and safety regulations, and the Office of Health and Safety Services must approve their use. • The site, design, and operation of a BMP will not produce any adverse environmental impacts, and comply with applicable environmental regulations.
O&M	• The BMP will operate passively during storm events. No personnel are required to be on site prior to or during a storm event to initiate operation of the BMP or perform maintenance to keep the BMP operational. • Maintenance requirements for a BMP are well understood and defined with respect to scope and frequency. The goal for BMPs is to use field maintenance personnel to perform inspections and maintenance tasks using available equipment and resources where possible. • Maintenance personnel or Contractors must be able to perform operations and maintenance (O&M) inspections and tasks safely. • Long-term maintenance requirements and costs for the BMP are identified. • The treatment BMP can be designed and operated in a manner that does not create a public nuisance or health hazard. Specifically, there is not a concern with regard to potential disease vectors, such as mosquitos. Treatment BMP design and prescribed O&M are adequate to ensure BMP operation and meet water quality goals, while at the same time reducing potential vector concerns to an acceptable level. • Operation and maintenance records and costs for the evaluation period are documented and reported with the study results.
Treatment Performance[1,2]	• Study provides results from full-scale field testing of a stabilized (erosion-free) post-construction transportation-related impervious drainage area. • Sampling and analysis was conducted in accordance to the Caltrans Stormwater Monitoring Guidance Manual, or other recognized protocol, such as the Urban Stormwater Runoff Monitoring Guidance (USEPA et al. 2009). • Testing was conducted at flow rates and volumes typical of Caltrans drainage areas (areas vary, but usually are between 0.1 and 15 acres. Flow and volumes can be found by using Caltrans Basin Sizer).

Criteria Type	Feasibility Criteria
	• Mean (influent) concentrations from study were below the 90th percentile of statewide characterization data. • Data was collected from at least eight storm events over a minimum period of two years, and demonstrate a statistically significant removal ($p \leq 0.1$), which may require monitoring additional storm events. • Particle size distribution (PSD) during study was similar to the proposed field conditions (e.g., state whether traction sand was applied). • The study's mean removal estimate corroborates the performance claim. • The rainfall record for the study area or its vicinity during the evaluation period is documented and reported with the study results. • Not all of the above criteria are met, but at least one of the following apply: ○ Statistically significant (p-value ≤ 0.1) constituent removal was established from independent stormwater field monitoring for at least one year ○ Removal efficiency based on best professional evaluation of unit operations and processes that are well established for treatment of other waters ○ Load reduction of nutrients or BOD due to partial infiltration ○ Statistically significant (p-value ≤ 0.1) constituent removal was established from independent laboratory testing that follows the Technology Assessment Protocol – Ecology (TAPE) from Washington State (ECY 2008), and testing used a volume of water equivalent to one year of runoff for a typical installation. Alternatively, a laboratory loading using actual stormwater could be used as with the Tahoe Small Scale Research Facility
Life-Cycle Cost	• Pollution control benefits have a reasonable relationship to the life-cycle costs to implement the BMP within Caltrans transportation infrastructure. • Estimated life-cycle costs are comparable to the established life-cycle costs of other approved BMPs that target the same constituents in runoff.

[1]*Performance criterion applies to statewide approvals. Influent data for regional BMP approval may be compared to regional datasets (Caltrans 2009).*

[2]*The 2016 SWMP states "Appendix A of the Treatment Technology Report Provides specific performance criteria that are used for BMP evaluation." Criterion taken from Caltrans 2018b.*

3.0 Types of Study Variables

Study variables generally fall into one of four categories:

- Site characteristics
- BMP characteristics
- Storm characteristics
- Maintenance practices

Examples of variables for each category and how they can affect BMP function are described below.

3.1. Site Characteristics

This category consists of variables that affect the quality and quantity of runoff. Examples include:

- Drainage area and slope
- Traffic volume
- Vegetation and vegetation canopy
- Percent imperviousness, soil type, and soil compaction
- Representativeness of inflows

3.1.1. Drainage Area and Slope

Data collected from assessments within a watershed can vary depending on various factors such as the location of sampling points, and size and slope of the watershed. It is important to select sampling locations where slopes, vegetation, channel width, etc., are relatively uniform and similar to the rest of the study area.

3.1.2. Traffic Volume

Traffic volume has the potential to affect water quality. Annual average daily traffic (AADT) is the total volume for the year divided by 365 days. The traffic count year is from October 1 through September 30. Very few locations in California are actually counted continuously. Traffic counting is generally performed by electronic counting instruments moved from location throughout the state in a program of continuous traffic count sampling. The resulting counts are adjusted to an estimate of annual average daily traffic by compensating for seasonal influence, weekly variation and other variables, which may be present.

3.1.3. Vegetation and Vegetative Canopy

The amount of vegetation and vegetative overhanging canopy can change during the course of a study and may alter the results. Riparian vegetation along a creek can indirectly affect temperature and biologic parameters in receiving waters. Vegetation along a road can contribute to the gross solids loading, which might affect clogging and BMP maintenance frequency.

3.1.4. Percent Imperviousness, Soil Type, and Soil Compaction

Similar to other geographical or topographical variations, percent imperviousness, soil type, and soil compaction in the study area can yield variations in the data that need to be addressed during the planning of the study.

3.1.5. Representativeness of Inflows

Another variable to consider is whether the inflows are representative of Caltrans runoff. One factor that might affect inflow characteristics is the presence of source control BMPs or base flow that might cause the inflow to be cleaner than usual, or the presence of local sources that may make runoff more polluted than normal Caltrans runoff. An example of the latter in the Tahoe Basin is the contribution from a snow mobile rental business that resulted in much higher turbidities than expected at one media filter site. Another factor to look for is the intermingling of Caltrans runoff with runoff from other sources. In urban areas, it is common for Caltrans to pipe its runoff to city stormwater sewers. In these cases, the runoff must be sampled before the pipe connection. In rural areas, the reverse sometimes happens. Runoff from local non-Caltrans facilities mixes with highway runoff in a drainage ditch in the highway right-of-way.

3.2. BMP Characteristics

The follow aspects of a BMP will affect its performance:

- Footprint (e.g., length, width, area)
- Slope
- Media characteristics
- Depth of standing water
- Vegetation characteristics
- Hydraulic loading rate
- Effluent orifice sizing

3.2.1. Footprint

BMP size is a fundamental design variable that has an impact on treatment performance. The length or area available for treatment is usually the key design variable for treatment BMPs.

3.2.2. Slope

Slope affects the velocity of runoff across a surface, which influences the amount settling, filtration, and infiltration that can occur, as well as the amount of erosion or rilling that can impede the intended function of the BMP.

3.2.3. Media Characteristics

These variables include properties of the soil, media, or other material that provide the treatment such as type, grain size, uniformity coefficient, media depth, and infiltration rate. Note that infiltration rate is often approximated by using saturated hydraulic conductivity, although other variables such as moisture content, suction head, or initial deficit.

3.2.4. Depth of Standing Water

The amount of water ponded atop a BMP just before a rainfall event can influent the infiltration rates and overflows rates of a BMP.

3.2.5. Vegetation Characteristics

The type and amount of coverage influence the performance of a BMP. Different plant types can have different root systems that influence infiltration, different foliage that can influence evapotranspiration, and different cell structures that influent uptake. BMP's that are supposed to have good vegetated coverage but do not will not perform to their full potential and may become a source of pollutant transport if erosion and rilling occur.

3.2.6. Hydraulic Loading Rate

BMP's are typically designed to accept a minimum or maximum rate of runoff entering it. Whether or not the actual influent rates meet the design rates will influence the BMP's performance.

3.2.7. Effluent Orifice Sizing

The size of the effluent orifice or other component that throttles the effluent rate of the BMP can affect the BMP's performance by controlling the retention time, thereby influencing adsorption and settlement.

3.3. Storm Characteristics

Storm Characteristics that affect a BMP's performance include:

- Precipitation intensity and duration
- Inter-event timing
- Antecedent dry period
- Seasonal timing

3.3.1. Precipitation Intensity and Duration

Variations in rainfall intensity and duration affect runoff rate, volume, pollutant wash off rate, in-channel flow rate, and other phenomena that determine the pollutant concentrations, pollutant forms, and stormwater flow rates observed in a study.

3.3.2. Inter-event Timing

Time between monitoring events affects the BMP performance. Between monitored events, there will be different rainfall intensities, durations, rate, and volumes, as well as different amounts of evapotranspiration, pollutant buildup, and infiltration. Some or all of these conditions will after the BMP performance, depending on its applicable treatment mechanisms.

3.3.3. Antecedent Dry Period

This is the time between rainfall events, whether or not they were monitored. Similar to inter-event timing, concentrations of pollutants may vary widely depending on the antecedent dry period. Infiltration, evapotranspiration, storage, and other treatment mechanisms dictating the BMP's performance are also affected.

3.3.4. Seasonal Timing

If the study area is subject to seasonal changes or fluctuations due to particular events, this will affect the timing of data collection. Rainfall patterns vary over wide regions and consideration will be given to the months of the year designated as "rainy season" and "non-rainy season."

3.4. Maintenance Characteristics

Maintenance involves the following variables:

- Type of activity
- Frequency of activity
- Adequacy of practice

3.4.1. Type of Activity

Maintenance activities may include cleaning and repairing equipment, vegetation control, algae reduction, sediment removal/dredging, litter/debris control, and inlet/outlet cleaning. The type of maintenance activity will influence the cost and feasibility of a BMP.

3.4.2. Frequency of Activity

How frequent a maintenance activity is required will also affect the cost and feasibility of a BMP.

3.4.3. Adequacy of Practice

The degree to which the required maintenance activities are actually performed (i.e., whether they were conducted on time and sufficiently) can affect the BMP's performance.

4.0 Evaluating Relevant Assumptions

BMP studies involve making assumptions, consciously or unconsciously, about sites, BMP O&M, monitoring conditions, and other study aspects. Some of these assumptions can result in important variables being overlooked. Common assumptions related to the selected study sites, BMP design, BMP operation, and monitoring are listed below.

4.1. Site Assumptions

- Runoff is representative of Caltrans runoff.
- There are no unaccounted sources (e.g., nutrient inputs from fertilizer or compost use are minimal and can be ignored).
- There is no base flow or groundwater intrusion into the BMP.
- Soil and vegetation characteristics are known (e.g., soil infiltration rates inferred from maps are representative of actual rates).

4.2. BMP Design Assumptions

- The full range of flows up to the design storm enters the BMP
- The outlet/overflow elevation is correctly designed and constructed
- The slope is correctly graded
- The BMP is correctly sized

4.3. BMP Operation Assumptions

- Installation of an impermeable liner will prevent infiltration losses from the base and sides of the BMP and migration of existing contaminated soils into the BMP. (This may not be the case if the liner is installed improperly or has significant tears that occur during installation.)
- There is no short-circuiting of flows within the BMP.
- O&M practices are carried out as needed and do not have an adverse impact on BMP operation (i.e., O&M is done following required frequency and all tasks are done thoroughly and properly. For example, vegetation is maintained such that decay and resulting nutrients have negligible impacts on treatment performance).

4.4. Monitoring Assumptions

- Data collected from few sites and over a relatively short time span will accurately represent how the BMP works.
- The monitoring plan accounts for all inflows and outflows from the BMP.

5.0 Study Types

In conducting a BMP study, several different monitoring approaches (study types) can be chosen. The selection is based on the study questions, the type of BMP, the study constraints, and the current and historic conditions of the study area. Each type of study has associated strengths and weaknesses as described below.

5.1. Influent–Effluent Approach (In-and-Out)

Comparison of influent and effluent water quality is the method most often used in studies of treatment BMPs. This method is used to estimate the pollutant removal capability of an individual BMP or a series of in-line BMPs (i.e., a treatment train). The typical monitoring layout strategy for the influent-effluent approach is illustrated in Figure C-1. The monitoring layout for a treatment train would look similar, except that the effluent from the first unit process also serves as the influent to the second unit process in the series. Typically, the effluents from both unit processes would need to be monitored.

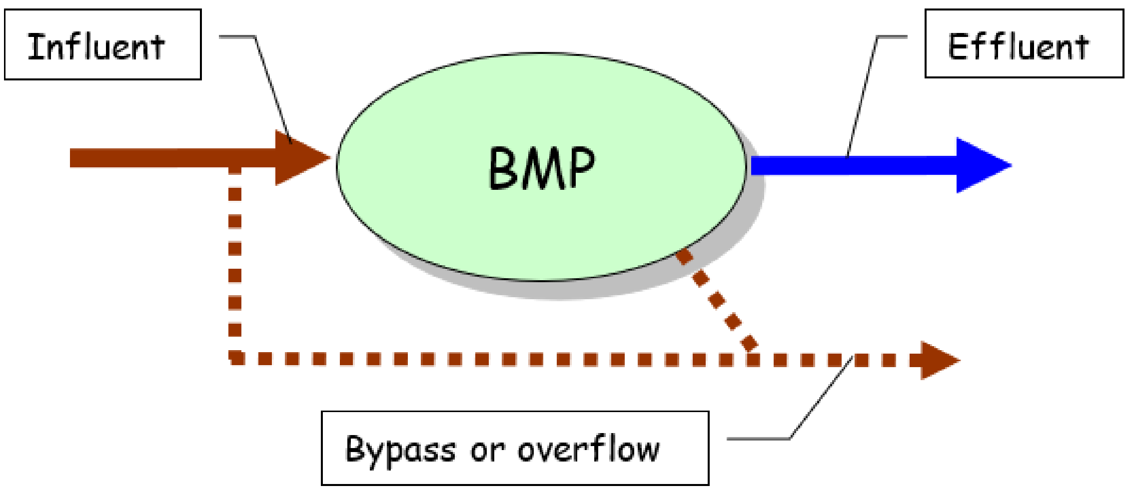


Figure C- 1. Influent–Effluent Approach Showing Monitoring Locations

There are several benefits in applying the influent-effluent approach for BMP efficiency:

- Environmental factors are better controlled in this approach and statistical variability in the data is generally less.
- The cost of monitoring is substantially less than that of watershed approaches (discussed below) because fewer data points are needed.
- The time required for monitoring can be substantially less than that required for watershed approaches because a calibration period prior to a monitoring program is not required.
- If climate is not a major factor affecting performance, the experimental results for a particular type of BMP can be extrapolated to other physiographic regions.

A limitation to the influent-effluent approach is that downstream benefits of BMP implementation cannot be established without additional data collection. Influent-effluent studies reveal pollutant removal rates and effluent flows and concentrations, but do not directly measure a BMP's effects on aquatic or riparian communities. Similarly, hydromodification impacts (i.e., increased erosion of channel beds and banks, sediment pollutant generation, or other impacts due to increased erosive forces) can be inferred, but are not directly measured.

5.2. Watershed Approaches

Watershed approaches to BMP evaluation are used where (1) discrete inflows and outflows cannot be monitored, (2) the BMPs are of a dispersed nature such as porous pavement applications, (3) the BMPs in question involve source control activities such as street sweeping and public outreach programs, and (4) the study questions relate to BMP effects on the environment rather than just BMP performance. Watershed approaches include upstream downstream, before-after, and paired watersheds.

5.2.1. Upstream–Downstream Approach

In contrast to the influent-effluent approach, the upstream-downstream approach entails a comparison of data collected from in-stream locations upstream and downstream of a BMP application. Figure C- 2 shows a schematic of an upstream-downstream approach. Station A is sited to monitor the in-stream concentration of constituents upstream of the land treatment area; Station B is sited below the BMP treatment area. Monitoring at the upstream location accounts for incoming pollutant sources that are unrelated to those that arise from within the study area. This method is more complex than the influent-effluent approach because the BMP is no longer isolated. Rather, its effectiveness must be discerned out of a data set that includes naturally occurring climatic and environmental conditions. For example, the occurrence of tributaries between the two data collection points, or changes in geology or land use can introduce variations in stream characteristics that may mask or overwhelm the effect of the BMP.

At the very least, a relatively large data set may have to be collected to discern statistically significant effects. Yet, sampling over a long time period brings up other issues. Year-to-year and seasonal variability in water quality constituent concentrations under certain conditions may also surpass the changes caused by the BMP over any given time period. To account for some of this variability, a monitoring period of at least two to three years is recommended for both pre- and post-BMP evaluations. Care should be taken in siting monitoring locations to minimize confounding influences. Also, if time-dependent changes are anticipated (e.g., rapid urbanization of the watershed), consideration should be given to increasing the number of sampling sites to minimize the sampling period. Despite these potential drawbacks, if this method of monitoring is conducted properly, the results can produce evidence of BMP influence or lack of influence on the study watershed (Coffey et al. 1993).

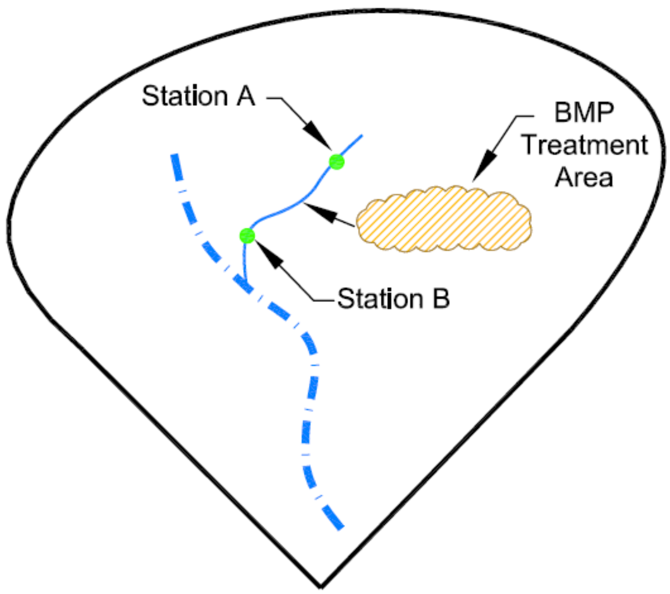


Figure C- 2. Upstream–Downstream Approach

5.2.2. Before–After Approach

In the before-after approach, data are collected at some location, a change is made (i.e., a BMP is implemented or modified), and additional data are then collected at the same location. Figure C- 3 shows an example.

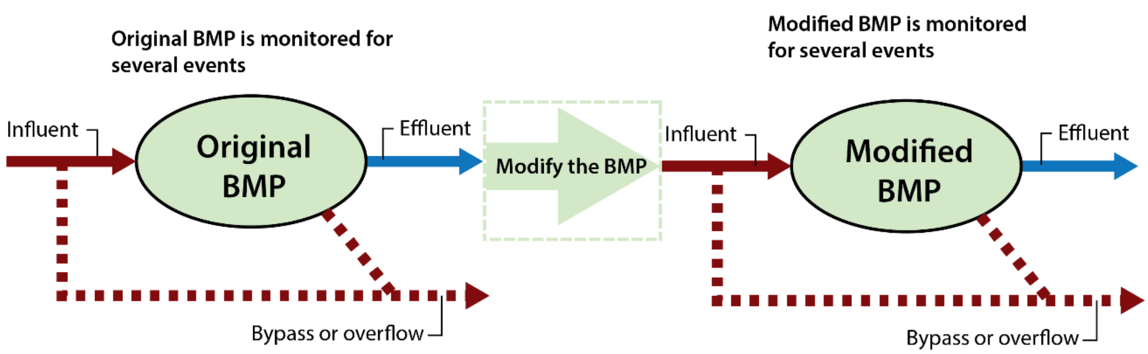


Figure C- 3. Before-After Approach Schematic

This approach can be used to evaluate a BMP at a single location or a watershed-wide BMP program. As in the upstream-downstream approach, year-to-year and seasonal differences can have a significant influence on the results. A two- to three-year pre- and post-BMP monitoring period is recommended to account for this variability (Coffey et al. 1993). The effect of longer-term climatic trends on hydrologic variability may, however, still mask the effectiveness of a BMP program. An additional shortcoming of this approach is that once the BMPs are implemented the baseline data characterization cannot be improved upon. Unrepresentative conditions during the baseline period can lead to erroneous conclusions by researchers comparing "before" and "after" data sets. For example, drought or extreme seasonal precipitation during baseline monitoring may make the baseline data unrepresentative. Also, to substantiate a cause-and-effect relationship, the predictor variable (e.g., erosion rate) must be adjusted for year-to-year changes in hydrologic conditions. Because of these problems, some experts prefer to combine this method with that of the upstream-downstream approach to strengthen the results of the findings (Coffey et al. 1993). Because hydrologic variability can occur over longer periods of time, comparative analysis of data collected using a before-after approach over the short term may actually be dealing with two distinct populations of hydrologic conditions.

5.2.3. Paired Watersheds Approach

In the paired watersheds approach, water quality data from two or more similar watersheds are compared. At least one is established as the control (undisturbed) watershed while the others include the BMPs being studied. Figure C- 4 shows a schematic of paired watersheds BMP monitoring. Data are collected in concurrent time periods, and changes in the data are taken as being indicative of BMP influence. In a typical paired watersheds approach, it is often desired to switch the roles of each watershed, from "control" to "treatment" and vice versa, to eliminate potential for bias. The switch may be performed on a yearly basis, or as often as needed. If properly implemented, this method provides reliable results and is perhaps the most effective watershed approach for monitoring BMP program effectiveness (Coffey et al. 1993). One weakness of this approach; however, is that it depends on the watersheds being truly similar except for the BMP. The watersheds to be compared must be in close proximity for climatic homogeneity, and have similar geology and land uses that are stable over the study period. Finding such watersheds can be difficult. Even if the watersheds appear similar, there is no guarantee that the runoff quality will be the same. For this reason, it may be advisable to undertake a calibration season in which the watersheds are sampled without any BMPs installed. While this increases the cost of the study, it also exposes any differences, which could cause misinterpretation of the data from the BMP study. Finally, in this kind of study, it is important that concurrent samples be collected for comparison. The sampling program must be very reliable, which can be a challenge, especially if the watersheds are located some distance from each other.

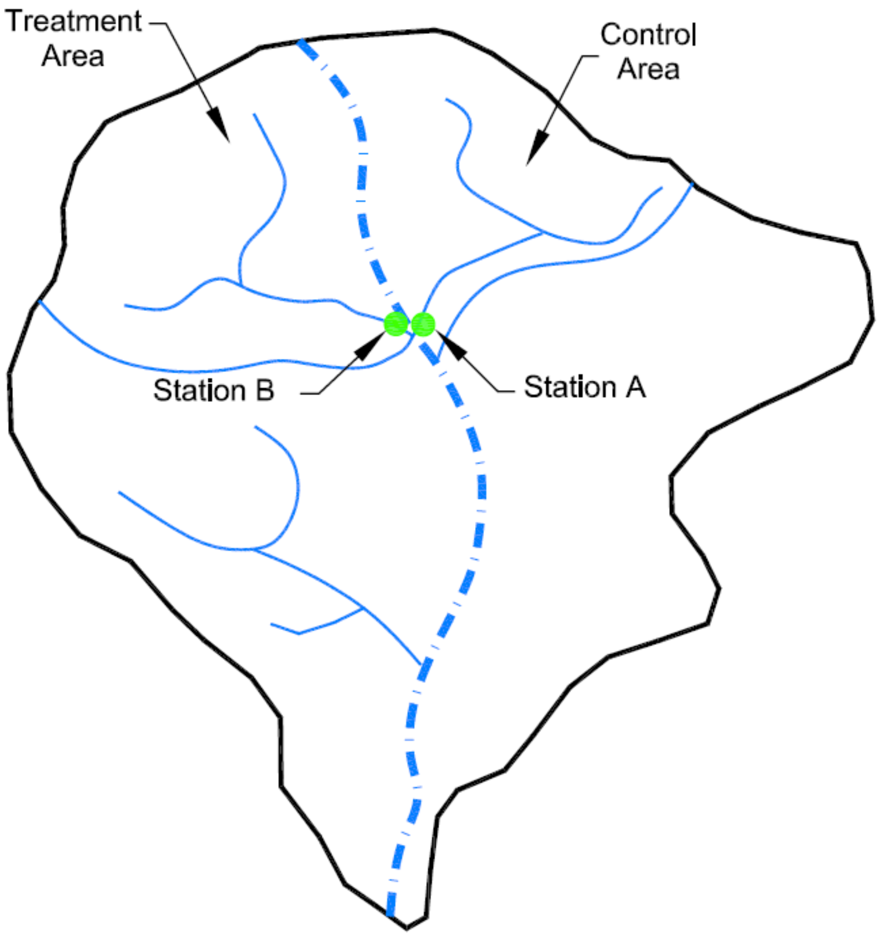


Figure C- 4. Paired Watersheds Approach

6.0 Specifying Which Variables to Monitor and Which to Control

6.1. Variables to Be Monitored or Recorded

Typical variables to be monitored or recorded include:

- Constituent concentrations
- Runoff volumes and flow rates
- Precipitation type, intensity, and duration
- O&M activities, frequencies, and completeness
- Vegetation/vegetative canopy
- Depth of ponding water
- Antecedent dry period (time between storm events)
- Inter-event timing (time between monitored events)
- Precipitation depth and probability of occurrence (for mobilization decisions)

6.2. Variables to Be Fixed or Controlled

Typical variables to be fixed or controlled include:

- Drainage area and slope
- Traffic volume
- Percent imperviousness
- Soil type and compaction
- Representativeness of flows (amount of runoff at a site that is from Caltrans)
- Representativeness of inflows (whether influent water quality is representative of typical Caltrans influent water quality)
- Constituent types
- Seasonal timing (time of year)
- Number of sites
- Number of samples

7.0 Controlling Relevant Study Variables

The subsections below describe how different study variables can be controlled. The variables are categorized as follows:

- Site characteristics
- BMP characteristics
- Construction considerations
- Monitoring considerations

7.1. Site Characteristics

During site selection, the need for controls should be considered. For study questions oriented toward general characterization of BMP performance, providing controlling variables may not be an essential part of the study. In these cases, making sure that conditions are representative or typical of Caltrans applications is more important. For study questions focused on how treatment mechanisms work, or the effects of varying design parameters, it is important to control as many of the variables as possible (ideally all of them) except for the variable under study. For example, in the SR 73 Detention Basin Study, basin size and operation mode (in-line vs. off-line) were the variables being studied. Four basins of different sizes were designed in in-line mode to isolate basin size as the study variable. To study off-line operation required four additional basins of the same sizes as the first four but in the off-line mode. To control for weather, the eight basins were located as close to each other as practical. As this example shows, the need to provide controls will influence the number of basins needed for the study.

7.1.1. Drainage Area and Slope

These can be constrained as necessary to meet the study objectives. If the objective is to study the effect of a design parameter, for instance, only a few representative drainage areas need to be studied. If widespread testing of the BMP under the full range of Caltrans conditions is desired, a range of drainage areas can be specified.

7.1.2. Traffic Volume

Traffic volume should be considered when selecting sites because it has the potential to affect inflow quality. Sites with more than 30,000 vehicles per day are characterized as high AADT; sites with 30,000 or less vehicles per day are characterized as low AADT.

7.1.3. Vegetation and Vegetative Canopy

If the presence of vegetation is likely to adversely affect the study results, sites should be selected where vegetation is minimal, or maintenance procedures should be initiated to control the level of vegetation during the study period. Alternatively, sites with very similar vegetation can be chosen and maintained (or not) in identical fashion.

7.1.4. Percent Imperviousness, Soil Type, and Soil Compaction

These can be constrained as necessary to meet the study objectives. Soil characteristics are critical in siting infiltration type BMPs – site selection for these should follow Caltrans procedures specified in the Project Planning and Design Guide (PPDG) (Caltrans 2007).

7.1.5. Representativeness of Inflows

It is desirable to select sites where inflows are representative of Caltrans runoff. This may require field surveys to ensure that potential sites meet this requirement, since some inflow sources may be seasonal. It also requires consideration of traffic volume (see above). In the Highway 50 Activated Alumina Filter Pilot Study, for example, one site had seasonal base flow that significantly affected inflow quantity and quality.

7.1.6. Site Characteristics Affecting Monitoring

Monitoring sites must facilitate representative sampling and flow measurement. The following criteria should also be considered in the selection of monitoring sites:

- Locate monitoring sites where flows (inflow, outflow, overflows, bypasses, or in-stream flows) are relatively uniform and stable or can be made relatively uniform and stable using control features so that accurate flow measurements can be made.
- Avoid sites likely to be affected by backwater and tidal conditions.
- Ensure monitoring sites will be accessible, well secured, and large enough to accommodate monitoring equipment such as flumes.
- Locate monitoring sites need where field personnel can be as safe as possible.

Further discussion of siting criteria can be found in Section 3.2.1 of the ASCE Urban Stormwater BMP Performance Monitoring Manual (ASCE 2002), and Chapter 3 of the PSGM.

7.2. BMP Characteristics

This category includes the physical design features identified in Step 4.

7.2.1. Treatment Area

Most BMPs are sized based on either water quality volume or water quality flow. The basis for sizing the BMP should be stated clearly, together with deviations from standard sizing methods. For example, the media area at one of the Highway 50 Activated Alumina Filter Pilot Study sites was reduced to one-third of the calculated design area (based on standard Austin-style sand filter design) to purposely increase loading and provide a quicker determination of media life. Refer to the Caltrans PPDG for sizing parameters, safety features, and other issues (Caltrans 2007).

7.2.2. Other BMP Parameters

The characteristics of the soil, media, plants, or other material that provide the treatment, and that reduce the flow volume and rate should be defined as necessary to meet the objectives of the study. Soil characteristics include texture, erosion potential (erodibility), infiltration qualities, and the ability to establish vegetation (e.g., fertility and amount of organic matter). Media characteristics include type (e.g., sand, peat, compost, perlite, zeolite, carbon, and other "exotic" media), and its physical and chemical properties. Plant characteristics include type, cover, need for irrigation, etc. For studies intended to determine how variations in design parameters affect performance, care should be taken to control the variables not under study. An example is given in the text box for a hypothetical pilot study comparing two alternative media filters.

Example of Controlling Variables: Unusual Hydrology

Unknown or unusual flows affect the interpretation of experimental results. For the Los Angeles River trash TMDL, litter was to be captured from all storms up to the 1-year, 1-hour event. In the Gross Solids Removal Device (GSRD) Studies, flow monitoring equipment was not installed because water quality samples were not collected. Researchers could see that the GSRDs occasionally overflowed because there were solids in the overflow bags, but they could not determine what size storm caused overflow, and therefore, could not determine if the devices were adequately meeting the TMDL standard.

In the SR 73 Detention Basin Studies, the rainfall for the first monitoring year (2004-05) was unusual in the sense that the rain fell predominantly during large and intense storms. Not only was the rainfall unrepresentative of the project site, but the solids loading was also unusually high because the slopes in most catchments had not stabilized after construction.

7.3. Construction

Construction parameters include key construction features that need to be identified to ensure they are constructed or installed carefully. A review of the treatment mechanisms described in the PSGM may be helpful here. For example, soil compaction must be minimized at sites where infiltration may be an important treatment mechanism. This can be done by prohibiting the use of heavy machinery and by specifying alternative construction methods. Another example is the requirement that manufacturer's guidelines be followed for media installation where filtration is the treatment mechanism. Poor media installation has caused settling to occur and may have been responsible for short-circuiting of flows in previous pilot studies.

Further discussion on key construction features is provided in Chapter 6 of the PSGM.

7.4. Monitoring

Monitoring parameters include the storm characteristics previously identified and the frequency and number of storms. Seasonal timing and frequency of monitoring activities is based on the appropriate number of random samples that need to be collected to characterize the rainfall and seasonal patterns associated with the monitoring site.

Estimates of the number of samples required to yield statistically valid monitoring results are also necessary for making decisions about the nature and extent of monitoring efforts. The appropriate number is the number of samples required to discern a significant difference between influent and effluent or between effluent and a numeric limit.

7.4.1. Seasonal Timing

Rainfall and seasonal patterns vary greatly in California and consideration shall be given to the months of the year that are designated “rainy season” and “non-rainy season” within each region of the state. To account for the various rainfall patterns (e.g., time frame, intensities, and amounts), the state is separated into several rainy seasons, as follows:

- Northwestern and southwestern California, rainy season: October 1 through May 1.
- Northern and central California, rainy season: October 15 through April 15.
- Eastern California, rainy seasons: August 1 through October 1, and November 1 through May 1.

7.4.2. Rainfall Type and Intensity

Establishing appropriate storm selection criteria can be a challenge. Ideally, one would want to obtain data from several phases of each storm for as long a study period as possible. The reasons for doing this are:

- To learn how the BMP performs during periods of low, medium, and high flows. The performance of some BMPs can vary dramatically with throughput rate.
- To provide statistical confidence in estimating performance based on widely varying runoff flows.
- To characterize the water quality of dry weather flows for those BMPs that rely on base flow (e.g., constructed wetland) or standing water (e.g., wet ponds). This is particularly important when the water quality volume of the BMP is large relative to storm events. In such a facility, comparing inflow to outflow during a storm event is not valid because the outflow may have little or no relationship to the incoming storm.

Setting storm event criteria is a complex process that is affected by the study goals, local climatic factors, permit requirements, and the BMP itself. The first consideration is whether the study goals require “representative” or “worst-case” events to be monitored.

Deciding storm representativeness requires consideration of storm size and antecedent dry period. If representative storms are desired, most storms occurring in the study period (except unusually large ones) would be monitored. If the objective of the monitoring is to consider a “worst-case” scenario, it would be desirable to select storms with the highest pollutant concentrations rather than a representative mix of storms. Worst-case conditions are likely to occur after long antecedent dry periods (72 hours to 14 days). Therefore, if feasible, storms would be selected with antecedent periods greater than 72 hours. Biasing storm selection to the “worst case” would not provide a representative sample of the population of all types of storm events and will not be used to estimate statistically derived exceedance frequencies. Local climatic factors also need to be considered in judging representativeness. In some climates, the majority of the rainfall may come in small events; in others, the majority may arrive in large events.

Monitoring plans and equipment may be different for each situation. Flow measuring equipment generally has a limited range of effectiveness. Primary flow monitoring devices (flumes and weirs) should be sized to rate the expected range of flows. Primary flow monitoring devices should be sized to rate the lowest flow possible. Primary flow measurement devices should also be sized for the flood design event (refer to the Highway Design Manual [Caltrans 2018e]) so that they do not cause an obstruction in the stormwater conveyance system, which could cause flooding or result in flows overtopping the device. The Isco Open Channel Flow Measurement Handbook, Sixth Edition (ISCO 2006) is an excellent resource that can be used to select and size a primary flow measurement device. Flumes sized to measure large flows, for instance, usually cannot accurately measure very small flows and vice versa. Regulatory requirements can be another consideration. In many locations, stormwater treatment requirements may apply to storms only up to a given size. Therefore, determining performance in large storms may not be of interest.

Finally, it should be remembered that the BMP itself might influence the storm criteria. For instance, in very small events, detention basins will often not produce a measurable effluent because of infiltration.

Lacking any criteria for storm volume selection to capture the worst-case conditions, and acknowledging that storm characteristics are highly dependent on climatic region, the following criteria may be used as a starting point in storm selection:

- Rainfall volume: 0.10 inch minimum, no fixed maximum.
- Rainfall duration: No fixed maximum or minimum, typical range 6 to 24 hours.

For specific details, refer to Caltrans Monitoring Guidance Manual (Caltrans 2019).

7.4.3. Number of Storms and Frequency of Monitoring

Because of the variability of rainfall and runoff quality, it is necessary to sample a number of storms to generate statistically reliable answers to the study questions. The number of samples needed depends on the variability in the data, the magnitude of the effect being studied, and the degree of confidence desired in the answer. For instance, suppose the study question involves the effectiveness of a BMP and differences between influent and effluent samples are being examined. Other factors being equal:

- More samples are needed when the data are variable than when they are not.
- More samples are needed when the differences between influent and effluent are small than when they are large.
- More samples are needed to establish a high degree of confidence than a lower one.

Appendix D1 presents a statistical methodology for determining the number of samples needed. Once the number of samples needed is calculated, the length and number of study sites can be determined. Table C-3 summarizes the number of storms and average rainfall depth per storm for several locations around California. Note that the average number of storms occurring each year decreases with increasing minimum storm size, increasing minimum dry time between events, and decreasing latitude. Lacking more site-specific data, this table can be used as a planning tool. After choosing the minimum storm size to be monitored, the minimum dry period defining the break between storms, and the location, estimate the number of storms per year and the average storm size using the table. For example, if the storm event criteria are a 0.25-inch minimum storm size with at least 24 hours between storms, then an average of 15 storms per year could be expected for a Sacramento site, and an average of 10 storms per year could be expected for a San Diego site. If the statistical calculations suggest that 30 storm events are needed, a study based in Sacramento would require two years, while a study based in San Diego would require three years. In general, Caltrans has conducted three-year pilot studies to ensure data collection during representative weather, although study durations should depend on data variability and statistical considerations. On the lower end, study periods will not be less than one year at any location so that the variability from seasonal changes or fluctuations can be captured.

Table C- 3. Average Number of Storms and Average Rainfall Depth per Year

Rain Gauge Location	Gauge No.	6 hr ADP[a]: Avg. # of Storms	6 hr ADP[a]: Avg. Storm Depth (inch)	24 hr ADP[a]: Avg. # of Storms	24 hr ADP[a]: Avg. Storm Depth (inch)	72 hr ADP[a]: Avg. # of Storms	72 hr ADP[a]: Avg. Storm Depth (inch)
Min. Storm Depth	-	0.1 in	0.1 in	0.1 in	0.1 in	0.1 in	0.1 in
Redding	7295	39	0.68	24	1.07	15	1.73
Sacramento	7630	28	0.54	20	0.77	13	1.21
Oakland	6335	29	0.54	21	0.78	13	1.27
San Francisco	7769	33	0.58	23	0.86	14	1.43
Fresno	3257	25	0.41	18	0.58	13	0.83
Los Angeles	5114	18	0.65	14	0.85	11	1.14
San Diego	7740	19	0.48	14	0.65	11	0.87
Min. Storm Depth	-	0.25 in	0.25 in	0.25 in	0.25 in	0.25 in	0.25 in
Redding	7295	24	1.01	18	1.41	11	2.26
Sacramento	7630	19	0.74	15	0.98	10	1.48
Oakland	6335	19	0.75	15	1.01	10	1.58
San Francisco	7769	21	0.81	17	1.13	11	1.77
Fresno	3257	14	0.59	13	0.76	10	1.04
Los Angeles	5114	12	0.90	10	1.10	8	1.42
San Diego	7740	12	0.68	10	0.86	8	1.12

[a] 6, 24, and 72 hr minimum antecedent dry period (dry time between storms)
Source: Scott Meyer (OWP at Sacramento State) based on data from Hydrosphere Data Products 2005.

Example of Number of Data Points Needed to See Differences

Where data sets contain large variations, many samples are required to demonstrate statistically significant results. For instance, in the Retrofit Pilot Study, the mean effluent values for coliform bacteria were 90 percent lower than the influent values, but the removal could not be considered to be statistically significant because only four coliform bacteria samples were collected. Choosing the right number of data points can be a complicated decision. In the DICE Study, litter discharges from drain inlets were initially judged to be the same regardless of whether the inlets were cleaned or not (based on about 50 samples). The fact that litter accumulated in the cleaned inlets (and was therefore removed from the flow) was not reflected in the discharge statistics. Another year of litter monitoring was required to establish a clear statistical difference.

7.4.4. Number of Sites

The number of sites that need to be monitored depends on program objectives, type of study, need for control sites, drainage basin size and complexity, and resources (time, personnel, funds) available for monitoring. In addition, the frequency of sampling at each location must be considered. Depending on objectives, resources, and logistical considerations, many locations may be sampled infrequently, or fewer locations more frequently. Sampling many locations is generally better for evaluating geographic variability as it affects climate and/or runoff characteristics. Employ this strategy if a study goal is to assess the wide-scale applicability of a BMP. Sampling a few locations for longer periods is generally better for evaluating BMP effectiveness over time and for characterizing specific monitoring locations (e.g., sensitive water bodies). Based on the number of sites proposed for the study, use statistical analysis (Appendix D) to estimate the number of data points and sites required for a given length of study.

7.4.5. Study Period vs. Number of Test Sites

Statistical considerations dictate the number of samples needed to discern an effect given certain variability in the data (see Appendix D). Whether these samples are collected at a few sites for a long time or at many sites for a short time is a Study Plan decision. Sometimes the study questions include geographic and temporal aspects that fix the number of test sites or the length of the study period. For instance, it may be a study goal to determine how a particular BMP performs in a variety of typical California climates. Alternatively, a goal may be to finish the study within a certain time to meet a regulatory requirement. If the study does not have these kinds of restrictions, however, there may be flexibility in planning pilot facilities – many installations with a short monitoring period or fewer installations with a longer monitoring period.

In calculating sampling costs, consideration must also be given to the number of unproductive events that are likely to occur. Sometimes rainfall does not occur as predicted or does not occur in sufficient quantities to justify sampling. Samples can also be missed because of mechanical problems with automatic samplers. For planning purposes, assume that one out of four sampling attempts will be unsuccessful.

8.0 Formulating the Study Questions

Example study questions are provided below.

8.1. Technical Feasibility Example Questions

According to the 2016 SWMP, BMPs must function under conditions typical of Department roadways and facilities and comply with Department drainage and safety requirements. Relevant study questions include:

- Can the BMP meet Caltrans vector control standards?
- Will the BMP interfere with highway or facility drainage?
- Will the BMP pose safety issues?
- What are the siting constraints?
- Is there a need for irrigation after initial installation and vegetation establishment?

8.2. O&M Requirements Example Questions

According to the 2016 SWMP, Caltrans personnel must able to be effectively operate and safely maintain a BMP throughout its intended life. Study questions include:

- What are the O&M requirements and are they reasonable?
- Can the requirements be met with Caltrans existing O&M practices?
- Can the BMP be modified so its O&M better fits within Caltrans existing practices?

8.3. Performance Capabilities Example Questions

The 2016 SWMP requires that pollutant removal effectiveness must be equal to or greater than those of currently approved BMPs. Study questions include:

- Is the BMP proposed for statewide use or restricted to a unique scenario (e.g., TMDL watershed, traction sand use area)?
- What level of treatment does the BMP provide under typical operating conditions?
- How does performance vary for different pollutants of concern?
- How does the BMP's performance compare to approved BMPs?
- Does the BMP protect or degrade downstream beneficial uses?
- How do design or site variables, influent quality, O&M practices, and/or precipitation characteristics affect performance?
- Does effectiveness improve, decrease, or remain unchanged over time?

8.4. Life-Cycle Costs Example Questions

Per the 2016 SWMP, a BMP's pollution control benefits must have a reasonable relationship to its implementation costs. Study questions include:

- What are the life-cycle costs (labor and materials for installation and O&M)?
- How do the life-cycle costs compare with those for approved BMPs?
- Can a BMP component be changed to reduce life cycle costs without adversely (or excessively) affecting performance?

9.0 Modifying the Study to Match Available Budget

Examples for modifying the study to match the available budget are provided below.

9.1. Adjust the Monitoring Scheme (and Study Objectives)

Reevaluate the monitoring details and study objectives to see if they can be changed to provide statistically meaningful data that answer the study objectives and stay within the allotted budget. Suppose, for example, that the costs of a study involving four sites and 30 samples per site needed to be reduced by half. One strategy would be to reduce the number of samples to 15 at each site. Doing this, though, would increase the probability of drawing incorrect conclusions from the experimental results. An alternative strategy would be to drop two of the sites and still collect 30 samples at each. Doing this, though, would limit the ability to determine the BMP's performance under the different conditions represented at different sites (e.g., climatic factors, runoff characteristics, traffic volumes, BMP design features). If a BMP's performance is expected to vary greatly with differing site conditions, it might be preferable to collect fewer samples at more sites. Generally, it is better to produce highly reliable results at fewer sites than less reliable results at many sites. Other potential strategies to reduce monitoring costs include:

- A phased approach wherein subsets of the overall study are addressed sequentially and the study stopped when sufficient data are collected.
- Limiting the number of constituents evaluated instead of reducing the number of storm events monitored. Some constituents (e.g., metals) may be treated similarly by the BMP, so results from only a few representative constituents may be needed.
- Using available data from published sources to supplement the study, thereby reducing the data collection effort.

After each modification of the study scope, the key question that needs to be answered is, "Will the study provide the information needed to reliably address the study objectives?" If the answer is "no", then the study objectives need to be modified or reduced. If that is not sufficient, then the possibility of postponing the study until adequate resources are available should be considered.

9.2. Integrate BMP Studies with Larger Construction Projects

This strategy applies to retrofits and new construction, and offers these benefits:

- More opportunities to locate BMPs in conjunction with other features (e.g., drainage systems, interchanges)
- Giving engineering staff experience with respect to stormwater BMP design, construction, and O&M
- Reduced mobilization, traffic-control, and equipment costs, as well as economies of scale during the construction process
- Regulatory compliance cost savings through the use of single permits for the entire project

9.3. Modify Existing BMPs

Testing new components (e.g., filter media) for existing BMPs can offer substantial costs savings. Existing BMPs can also be modified to allow side-by-side testing of alternatives, which reduces analysis costs since a single influent can be used.

9.4. Create Cross-Jurisdictional Partnerships

Cost sharing and cooperation between Caltrans and other agencies in constructing joint stormwater evaluation facilities can result in greater cost-effectiveness. However, different agencies may have different study objectives. Be wary of modifying the study scope in a way that compromises the ability of the study to meet its objectives, and ultimate Caltrans needs.

9.5. Consider Small-Scale Testing

This is a good option when the BMP is not well defined. Many approaches can be tested more quickly and cost effectively at the small-scale than the full-scale. With this approach, researchers can discard concepts that do not have merit before the expense of full-scale pilots.

Appendix D Introduction and Contents

This appendix describes statistical methods that can be used to analyze the performance of Best Management Practices (BMPs). Table D-1 provides a cross-reference to the title of each appendix and the topics covered in it. The intended audience for these appendices is engineers and scientists who have little or limited background in statistics. Accordingly, the focus is not on the theory of the statistical methods, but rather on the selection of an appropriate statistical method, interpretation of results, understanding of the limitations of the analysis methods, and drawing of valid conclusions. Flowcharts of step-by-step procedures are provided to guide the user to appropriate decisions or actions. References are provided regarding the theoretical basis of the various methods.

It is assumed that the user will have access to a statistical analysis software package and therefore will not need to use any equations to calculate results. The software package will be used instead to perform all calculations and generate typical output reports. As a result, few equations are included in the appendices. References are provided which include detailed equations for the various methods.

Emphasis is placed on graphical and numerical data summaries that facilitate the understanding of salient features of the input and output. Illustrative examples related to BMP studies are provided to show the organization of input data and output reports for a typical software package. Important portions of the reports are annotated and explained.

The specific commercial statistical software package used to analyze the illustrative examples in this manual is JMP, developed by the SAS Institute (www.jmp.com). Other software packages, including Minitab (www.minitab.com), SAS (www.sas.com), IBM SPSS® Statistics (www.spss.com), and System Status (www.systat.com), are also available. Any of these packages will be adequate for the types of statistical analysis that may be used in typical BMP studies. Although the specific procedures and commands vary for the different packages, typical parts of the output reports are similar. The annotated output reports from JMP, included in the appendices, should be useful in understanding similar reports from other software packages.

Table 1-1 Topics Covered in Each Statistical Appendix

Appendix	Title	Typical Study Questions Addressed
D1	How to Estimate an Adequate Number of Samples	How many samples would I need to achieve a desired confidence in the conclusions? After one or two years of sampling, how do I decide whether I would need more samples than initially planned?
D2	How to Examine Data Quality and Detect Possible Outliers in the Data	How do I prepare graphical and numerical data summaries to understand the salient data features and identify potential outliers?
D3	How to Examine Data Quality in the Presence of Non-detect Values	How do I account for non-detect results?
D4	How to Verify Common Assumptions for the Selection of an Appropriate Statistical Test	How do I verify whether data are normally distributed? How do I verify that the data variability of two or more groups is similar?
D5	How to Estimate Probabilities Using Data for a Single Variable	How do I estimate how often the average BMP effluent concentration would meet a legal limit? How do I estimate the BMP percentage removal of a pollutant with a specified confidence level?
D6	How to Compare Two Independent Data Sets	In an upstream-downstream watershed approach or paired watersheds approach, how do I decide whether a given BMP is effective in removing a pollutant? How do I compare the effectiveness of two pilot BMPs at a given geographic location?
D7	How to Compare Two Paired Data Sets	In an influent-effluent approach or before-after approach, how do I decide whether a given BMP is effective in removing a pollutant?
D8	How to Compare Three or More Independent Data Sets	How do I compare the effectiveness of three or more pilot BMPs at a given geographic location?
D9	How to Develop a Linear Regression Equation	How does BMP effectiveness vary as a function of such other factors as storm characteristics, BMP design variables, and operation/maintenance practices?
D10	How to Evaluate Time Trends in BMP Monitoring Data	How can I tell if the effectiveness of my pilot BMP is changing over time?
D11	How to Compare and Rank BMPs Monitored at Multiple Sites	How do I evaluate whether the treatment performance of BMPs is significantly different when each BMP type has been monitored at multiple locations?

Appendix D1 Purpose and Organization

D1.1 Purpose and Organization

This appendix presents statistical methods to estimate an adequate number of samples ("sample size") needed to answer various questions of interest in BMP pilot studies. If the sample size is too small, the variability of a data parameter (e.g., the mean pollutant concentration) could be large and mask the true effectiveness of a BMP. The smaller the sample size, the greater the risk of failing to detect a true BMP difference. In developing a statistics-based sampling plan, the sample size should be large enough so that the risk of failing to detect some minimum specified difference is sufficiently small.

This appendix presents procedures for estimating an adequate number of samples when planning BMP studies, as well as procedures used for interim data reviews. The latter reviews would be conducted after data have been collected for one year to check the adequacy of the original sample size estimate and decide whether additional data would be needed to meet the study objectives. Examples are included to illustrate the use of the statistical procedures.

D1.2 Statistical Procedures for Estimating Sample Size in the Planning Stage

Statistical procedures for estimating the sample size depend on the specific questions being addressed in the BMP study. To estimate sample sizes, group the study questions based on the number of independent data sets being analyzed for a given measure of water quality or BMP effectiveness. Gilbert (1987) defined independent data sets as those for which there is no natural way to pair the data. For BMP studies, pairing could occur because of spatial or temporal association between data points. For example, in an influent-effluent approach, pollutant concentrations may be measured in runoff entering and exiting a BMP at several BMP locations. In this case, the influent and effluent concentrations would be paired by the BMP location. Concentration differences or percent removals can be calculated to define a single data set. In a before-after approach, concentrations may be measured at several locations before and after implementing a particular BMP maintenance action. In this case, concentration differences or percent removals would define a single data set and the data again would be paired by the sampling location. In a paired watershed approach data may be collected at multiple locations in each of two watersheds – control (undisturbed) and treatment (disturbed), resulting in two independent data sets;. There is no natural pairing between a particular sampling location in one watershed and a sampling location in the other watershed. Table D 1-1 shows typical study questions in each group. The following sections describe the statistical procedures to use for estimating the sample size for each group.

Table D 1-1 Typical Study Questions for Comparison of Independent Data Sets

Number of Independent Data Sets	Typical Study Questions
1	For pollutant concentrations measured with an influent-effluent, upstream-downstream, or before-after monitoring approach, does a BMP remove a pollutant by more than a specified percentage? Does the effluent pollutant concentration exceed a specified legal limit?
2	For pollutant concentrations measured with a paired watershed monitoring approach, does a BMP remove a pollutant by more than a specified percentage? Is the performance of two BMPs in a similar environment setting significantly different? Is the performance of a given BMP significantly different in two different environmental settings?
3 or more	Given influent and effluent data for three pilot BMPs of the same type, how can one tell whether they are all operating in an equivalent manner, or whether there is some site-specific factor that causes one BMP to operate differently than the others? Is the performance of a given BMP significantly different in three different environmental settings?

D1.3 Study Questions Involving a Single Data Set

A single data set could be an independent data set in its own right, or it could be the combined result of two related data sets to form a paired data set. For example, the effluent concentration of a given chemical constituent downstream from a treatment or source control BMP would be a single independent data set. An example of a paired data set would be influent and effluent concentrations of a given pollutant upstream and downstream from a BMP. This is a paired data set because the effluent quality depends on the influent quality. A logical way to combine the two sets of data in a paired data set is to calculate the percentage removal of a pollutant or difference in concentrations, which results in a single data set.

The statistical procedure for estimating the sample size for this case focuses on the minimum change, Δ, in the mean value of the parameter of interest. For example, in the case where the percent pollutant removal is of interest, one could specify Δ of 20 percent. This would mean that the sample size should be large enough to detect a pollutant removal of 20 percent or more with high confidence. Another study question for this data set is whether the average concentration of a pollutant constituent exceeds a specified limit (e.g., a legal water quality standard). For this study question, *Δ* will be the minimum absolute or percentage difference in the average concentration above the standard that should be detected with high confidence.

The procedure to estimate the sample size for this case also depends on the distribution that could be assumed for the data set. Three different cases of data distribution could be considered – normal distribution, lognormal distribution, or neither.

D1.3.1 Single Data Set with Normal Distribution

The sample size, *n*, depends on the following factors:

- The standard deviation of the data
- The minimum change or difference in the average value of the parameter of interest, *Δ*
- The desired confidence that the selected statistical test would reach a correct conclusion when the assumed baseline condition is true
- The desired confidence that the statistical test would be able to detect a change of *Δ* from the baseline condition

Each of these factors is discussed below.

The standard deviation of the data, s: Since no data would have been yet collected at the planning stage, the standard deviation is typically estimated based on prior data at similar sites or professional judgment. Table D 1-2 through Table D 1-4 contain influent means and standard Table D 1-2 deviations from past Caltrans studies. BMP percent removal statistics are tabulated in Table D 1-5, categorized by BMP type.

In the interim data review stage (i.e., after collecting data for one or two years), the initial estimate of the standard deviation can be checked and revised if necessary.

The minimum change in the average value of the parameter of interest (e.g., average removal of a specified pollutant), Δ: This factor may be based on applicable legal or regulatory standards or policy decisions. Typically, Caltrans BMP studies look for a minimum change of 50 percent. In some studies, such as those where BMP effluents are close to regulatory limits, or where the BMPs being studied are thought to be working poorly, detecting smaller Δ values might be appropriate. To be considered for approval, however, a new BMP needs to show a practical level of effectiveness. In most cases, new BMPs would not be approved if they would not be expected to remove at least half of the constituent of concern.

The desired confidence that the selected statistical test would reach a correct conclusion when the assumed baseline condition is, in fact, true: Typically, the assumed baseline condition would be that the BMP provides no benefit. For example, the average pollutant concentrations in the influent and effluent are the same, or the pollutant removal is 0 percent. The desired confidence is commonly denoted by (1-α), and α is called the probability (or risk) of making a Type I error, or the probability of false rejection (i.e., the probability of rejecting the baseline when it is correct). In the context of a typical BMP study, α is the chance that one would claim a BMP is working when, in fact, it does not. Typical values of α for a Caltrans stormwater study are 5 percent and 10 percent. Based on common practice, a value of 10 percent is recommended.

The desired confidence that the statistical test would be able to detect a change of Δ from the baseline condition: This confidence is referred to as the power of the statistical test and commonly denoted by (1-β). The parameter β is the probability (or risk) of failing to detect a specified change. It is also referred to as the probability of making a Type II error, or the probability of *false acceptance* (i.e., the probability of accepting the baseline when it is incorrect). In the context of typical BMP studies, β measures the willingness to miss an opportunity to identify a BMP that does work. Typical values of (1-β) for Caltrans stormwater study are 80 percent and 90 percent. Based on common practice, a value of 80 percent is recommended for (1-β).

Table D 1-2 Influent Statistics for Highway and Maintenance Facilities

(Caltrans 2003b)

Constituent	Unit	Highway Facilities, Mean	Highway Facilities, Std Dev	Highway Facilities, n	Mntc. Facilities, Mean	Mntc. Facilities, Std Dev	Mntc. Facilities, n
TOC	mg/L	21.8	29.2	635	20.6	23.0	107
DOC	mg/L	18.7	26.2	635	18.2	18.2	75
EC	µS/cm[1]	96.1	73.4	634	80.9	110.6	56
Hardness as $CaCO_3$	mg/L	36.5	34.2	635	26.7	28.7	106
Chloride	mg/L	266	388	32			
TDS	mg/L	87.3	103.7	635	68.9	78.1	106
TSS	mg/L	112.7	188.8	634	96.4	95.0	106
Turbidity	NTU				144.8	92.2	29
Oil & Grease	mg/L	4.95	11.4	49	-	-	-
TPH (Diesel)	mg/L	3.72	3.31	32	-	-	-
TPH (Gasoline)	mg/L	-	-	-	-	-	-
TPH (Heavy Oil)	mg/L	2.71	3.4	20	-	-	-
As, dissolved	µg/L	1.0	1.4	635	9.5	17.3	106
As, total	µg/L	2.7	7.9	635	12.8	23.1	107
Cd, dissolved	µg/L	0.24	0.54	635	0.27	0.22	106
Cd, total	µg/L	0.73	1.61	635	0.69	0.63	107
Cr, dissolved	µg/L	3.3	3.3	635	1.4	1.0	106
Cr, total	µg/L	8.6	9.0	635	5.1	4.3	107
Cu, dissolved	µg/L	14.9	14.4	635	14.3	17.6	106
Cu, total	µg/L	33.5	31.6	635	29.5	37.6	107
Hg, dissolved	µg/L				27.7	51.4	7
Hg, total	µg/L	36.7	37.9	23	65.4	83.7	8
Ni, dissolved	µg/L	4.9	5.0	635	3.7	4.0	106
Ni, total	µg/L	11.2	13.2	635	7.9	7.7	107
Pb, dissolved	µg/L	7.6	34.3	635	1.6	3.0	106
Pb, total	µg/L	47.8	151.3	635	21.3	26.5	107
Zn, dissolved	µg/L	68.8	96.6	635	21.3	26.5	107
Zn, total	µg/L	187.1	199.8	635	245.6	259.3	107
Fecal Coliform	MPN per 100 mL	1132	1621	32	-	-	-
Total Coliform	MPN per 100 mL	13438	34299	32	-	-	-
NH_3-N	mg/L	1.08	1.46	8	-	-	-
NO_3-N	mg/L	1.07	2.44	634	0.74	1.13	107
Ortho-P, dissolved	mg/L	0.11	0.18	630	0.09	0.40	105
P, total	mg/L	0.29	0.39	631	0.23	0.20	106
TKN	mg/L	2.06	1.90	626	1.79	1.72	105
Chlorpyrifos	µg/L	-	-	-	-	-	-
Diazinon	µg/L	0.13	0.29	34	0.12	0.30	23
Diuron	µg/L	4.60	18.24	367	-	-	-
Glyphosate	µg/L	19.6	27.0	541	-	-	-

[1] *Conductivity unit = µmhos/cm*

Table D 1-3 Influent Statistics for Park and Ride Facilities and Construction Sites (Caltrans 2003b)

Constituent	Unit	Park & Ride Facilities, Mean	Park & Ride Facilities, Std Dev	Park & Ride Facilities, n	Constr. Sites, Mean	Constr. Sites, Std Dev	Constr. Sites, n
TOC	mg/L	18.6	20.6	179	12.8	9.9	47
DOC	mg/L	18.0	28.6	179	11.1	8.4	47
EC	µS/cm[1]	63.5	65.8	179	370.7	1659.8[3]	88
Hardness as $CaCO_3$	mg/L	36.6	45.9	179	185.2	885.7	118
Chloride	mg/L	-	-	-	-	-	-
TDS	mg/L	61.7	78.3	179	327.1	1448.4	117
TSS	mg/L	68.5	59.3	179	539.3	995.7	118
Turbidity	NTU	-	-	2	685.0	2098.3	19
Oil & Grease	mg/L	-	-	-	0.67	0.90	30
TPH (Diesel)	mg/L	-	-	-	-	-	-
TPH (Gasoline)	mg/L	-	-	-	-	-	-
TPH (Heavy Oil)	mg/L	-	-	-	-	-	-
As, dissolved	µg/L	0.7	0.6	179	2.1	1.7	47
As, total	µg/L	1.4	5.9	179	4.5	4.0	47
Cd, dissolved	µg/L	0.12	0.12	179	IDD	IDD	118
Cd, total	µg/L	0.30	0.30	179	0.58	1.17	118
Cr, dissolved	µg/L	1.0	0.9	179	5.7	6.4	118
Cr, total	µg/L	4.0	4.2	179	38.6	70.5	118
Cu, dissolved	µg/L	8.7	8.8	179	7.3	5.9	117
Cu, total	µg/L	17.1	15.2	179	37.2	92.8	118
Hg, dissolved	µg/L	-	-	-	-	-	-
Hg, total	µg/L	57.3	73.6	11	-	-	-
Ni, dissolved	µg/L	3.3	3.9	179	3.1	2.6	118
Ni, total	µg/L	6.2	4.8	179	57.4	283.2	118
Pb, dissolved	µg/L	1.3	2.7	179	1.1	4.3	118
Pb, total	µg/L	10.3	11.5	179	56.4	277.6	118
Zn, dissolved	µg/L	10.3	11.5	179	45.5	433.8	118
Zn, total	µg/L	154.3	157.1	179	190.3	555.8	118
Fecal Coliform	MPN per 100 mL	-	-	-	1777	4268	25
Total Coliform	MPN per 100 mL	-	-	-	3915	12023	26
NH_3-N	mg/L	-	-	-	0.29	0.47	116
NO_3-N	mg/L	0.57	0.83	10	0.96	0.79	71
Ortho-P, dissolved	mg/L	0.15	0.19	10	0.16	0.24	85
P, total	mg/L	0.33	0.42	10	1.98	13.47	115
TKN	mg/L	2.28	2.20	10	2.11	2.53	116
Chlorpyrifos	µg/L	-	-	-	0.05	0.12	108
Diazinon	µg/L	-	-	-	0.24	0.40	108
Diuron	µg/L	-	-	-	-	-	-
Glyphosate	µg/L	-	-	-	30.7	35.6	13

Table D 1-4 Influent Statistics for Rest Areas and Vehicle Inspection Facilities (Caltrans 2003b)

Constituent	Unit	Rest Areas, Mean	Rest Areas, Std Dev	Rest Areas, n	Veh. Insp. Facilities, Mean	Veh. Insp. Facilities, Std Dev	Veh. Insp. Facilities, n
TOC	mg/L	22.2	40.5	53	20.0	16.9	31
DOC	mg/L	19.9	39.6	53	18.5	15.9	31
EC	μS/cm^2	78.2	132.0	53	113.3	137.3	31
Hardness as $CaCO_3$	mg/L	33.0	81.2	53	33.5	22.1	31
Chloride	mg/L	-	-	-	-	-	-
TDS	mg/L	61.2	130.0	53	84.8	92.1	31
TSS	mg/L	63.3	54.4	53	83.4	53.0	31
Turbidity	NTU	-	-	-	-	-	-
Oil & Grease	mg/L	-	-	-	-	-	-
TPH (Diesel)	mg/L	-	-	-	-	-	-
TPH (Gasoline)	mg/L	-	-	-	-	-	-
TPH (Heavy Oil)	mg/L	-	-	-	-	-	-
As, dissolved	μg/L	1.4	3.3	53	1.0	0.4	31
As, total	μg/L	3.6	11.4	53	3.4	16.1	31
Cd, dissolved	μg/L				0.20	0.16	31
Cd, total	μg/L	0.32	0.53	53	0.56	0.40	31
Cr, dissolved	μg/L	1.9	2.5	53	1.8	1.2	31
Cr, total	μg/L	4.8	3.8	53	8.1	4.8	31
Cu, dissolved	μg/L	9.6	12.0	53	15.6	13.3	31
Cu, total	μg/L	16.0	14.2	53	33.6	24.1	31
Hg, dissolved	μg/L	-	-	-	-	-	-
Hg, total	μg/L	-	-	-	-	-	-
Ni, dissolved	μg/L	3.2	5.8	53	3.5	2.4	31
Ni, total	μg/L	7.3	8.3	53	8.4	4.7	31
Pb, dissolved	μg/L	1.2	1.7	53	2.7	3.9	31
Pb, total	μg/L	7.7	8.0	53	21.9	37.7	31
Zn, dissolved	μg/L	82.5	263.7	53	88.2	79.1	31
Zn, total	μg/L	142.4	298.9	53	244.5	151.6	31
Fecal Coliform	MPN per 100 mL	-	-	-	-	-	-
Total Coliform	MPN per 100 mL	-	-	-	-	-	-
NH_3-N	mg/L	-	-	-	-	-	-
NO_3-N	mg/L	0.96	0.88	53	0.89	0.81	31
Ortho-P, dissolved	mg/L	0.44	1.67	52	0.13	0.12	30
P, total	mg/L	0.47	0.53	53	0.28	0.16	31
TKN	mg/L	4.37	14.04	53	2.16	2.72	30
Chlorpyrifos	μg/L	-	-	-	-	-	-
Diazinon	μg/L	-	-	-	-	-	-
Diuron	μg/L	-	-	-	-	-	-
Glyphosate	μg/L	-	-	-	-	-	-

Table D 1-5 Descriptive Statistics for BMP Percent Removals

Constituent	Filtration BMPs, Mean Std Dev (Range)	Filtration BMPs, Mean COV (Range)	Sedm. BMPs, Mean Std Dev (Range)	Sedm. BMPs, Mean COV (Range)	Biofiltr. BMPs, Mean Std Dev (Range)	Biofiltr. BMPs, Mean COV (Range)
TOC	39 (11 – 55)	16 (7 – 32)	18 (12 – 26)	3.3 (1.9 – 4.8)	28 (10 – 403)	2.3 (2.3 – 3.5)
TSS	26 (5.6 – 52)	0.4 (.06 - 1.0)	26 (9 – 42)	0.5 (0.1 – 1.5)	20 (9 – 186)	2.4 (0.1 – 3.5)
TDS	73 (8 – 136)	1.2 (0.2 – 2)	31 (10 – 48)	11 (0.5 – 23)	40 (23 – 1281)	3.3 (0.7 – 7.5)
TPH (Diesel)	37 (26 – 51)	1.0 (0.5 - 2)	43 (21 – 52)	4.0 (0.6 – 6.5)	31 (N/A)	0.5 (N/A)
TPH (Heavy Oil)	25 (13 – 32)	0.5 (0.2 – 1.5)	47 (8 – 63)	3.0 (0.2 – 4)	15 (4.5 – 76)	2.6 (0.1 – 3.6)
Cd, total	27 (19 – 29)	0.6 (0.3 – 1.6)	35 (27 – 52)	1.4 (0.5 – 2.6)	14 (10 – 61)	0.7 (0.2 – 1.4)
Cr, total	80 (16 – 134)	8.3 (0.4 – 22)	22 (12 – 28)	0.8 (0.3 – 1.8)	13 (9 – 40)	56 (0.1 – 141)
Cu, total	28 (24 – 39)	0.9 (0.6 – 2.3)	24 (22 – 28)	0.9 (0.4 – 2.0)	15 (7 – 61)	21 (0.1 – 80)
Ni, total	55 (27 – 95)	25 (1 – 50)	29 (20 – 37)	0.8 (0.3 – 1.8)	16 (12 – 74)	1.3 (0.2 – 2.6)
Pd, total	22 (6.4 – 45)	0.3 (0.1 – 1.3)	28 (15 – 47)	1.6 (0.2 – 3.4)	13 (8 – 42)	0.4 (0.1 - 0.8)
Zn, total	27 (6.6 – 54)	0.5 (0.1 – 1.6)	24 (13 – 37)	0.6 (0.3 – 1.6)	13 (7 – 47)	0.4 (.01 – 1.2)
Fecal Coliform	160 (43 – 436)	2.0 (0.6 – 3.6)	47 (37 – 66)	1.2 (0.7 – 2.5)	20 (14 – 134)	5.7 (1.2 – 16)
P, total	48 (23 – 64)	1.7 (0.5 - 3)	52 (27 – 77)	2.0 (0.6 – 3.4)	19 (40 – 135)	13 (4.4 – 19)
TKN	46 (25 – 76)	1.3 (0.6 – 2.5)	32 (18 – 47)	4.2 (2.7 - 11)	16 (21 – 53)	21 (3 - 47)

Note: Data are for mean standard deviation (and range for individual sites) and mean coefficient of variation COV (and range for individual sites). Descriptive statistics are for the actual percent removals and not for influent or effluent data.
Source: Sacramento State Office of Water Programs.

For given values of *s*, Δ, α, and β, one can calculate the minimum sample size, *n*. JMP and other similar software packages, such as Visual Sample Plan developed by the U.S. Department of Energy (USDOE 2005), provide exact methods. However, calculating *n* requires the knowledge of some advanced U.S. Environmental Protection Agency (USEPA) guidance document EPA QA/G-9S (USEPA 2006a). The approximate method also assumes a normal distribution. Valid computational simplifications are made as described in the USEPA guidance document EPA QA/G-4 (USEPA 2006b). The USEPA document EPA QA/G-9S gives the following Equation D 1-1 to calculate *n*:

Equation D 1-1

$$n = \frac{s^2 \left(z_{1-\alpha} + z_{1-\beta}\right)^2}{\Delta^2} + \frac{z_{1-\alpha}^2}{2}$$

where *n* = sample size

s = sample standard deviation

$z_{1-\alpha}$ = value of standard normal variate at 1-*α* probability

$z_{1-\beta}$ = value of standard normal variate at 1-*β* probability

Δ = minimum change in the average value of the parameter of interest

The USEPA has developed a software program called DEFT (Decision Error Feasibility Trials) that can be used to calculate *n* using Equation D 1-1. This program can be downloaded from https://www.epa.gov/quality/resources-planning-new-data-collections. Equation D 1-1shows plots of *n* derived from Equation D 1-1 for typical values of other parameters. Note that *Δ* is normalized by *s* in Equation D 1-1.

When several constituents are of concern, the above procedure is used for each constituent and the largest sample size is selected.

Example D1-1

Assume that effluent zinc concentrations are of concern in a BMP study. The study question of interest is: Does the mean effluent zinc concentration exceed a specified standard by more than 2 mg/L? The expected number of annual independent storm events to be monitored is 10 and the monitoring period will be 3 years. Based on data on a similar BMP at a similar site, it was assumed that the effluent zinc concentration would follow a normal distribution with a standard deviation of 6 mg/L. The desired confidence for reaching the correct conclusion when the mean effluent zinc concentration was equal to the standard was 0.95 and the desired power of detecting the minimum change of 2 mg/L over the standard was 0.8. How many pilot sites should be selected for this study? Note that for Caltrans studies, the 90 percent confidence level should be used (i.e. a value of ($1-\alpha$) of 0.90 should be used instead of the 0.95 value used in this example). For this example, $s = 6$, $\Delta = 2$, $1-\alpha = 0.95$, and $1-\beta = 0.8$. Using Equation D 1-1, the sample size, n, is calculated to be 57. Because 3 years of monitoring is planned with 10 events per year, 2 pilot sites should be selected, which would yield a total of 60 data points. Alternatively, a single pilot site may be monitored for 6 years to obtain the necessary data. If multiple sites are selected, they should be located so as to provide independent, and not redundant, data.

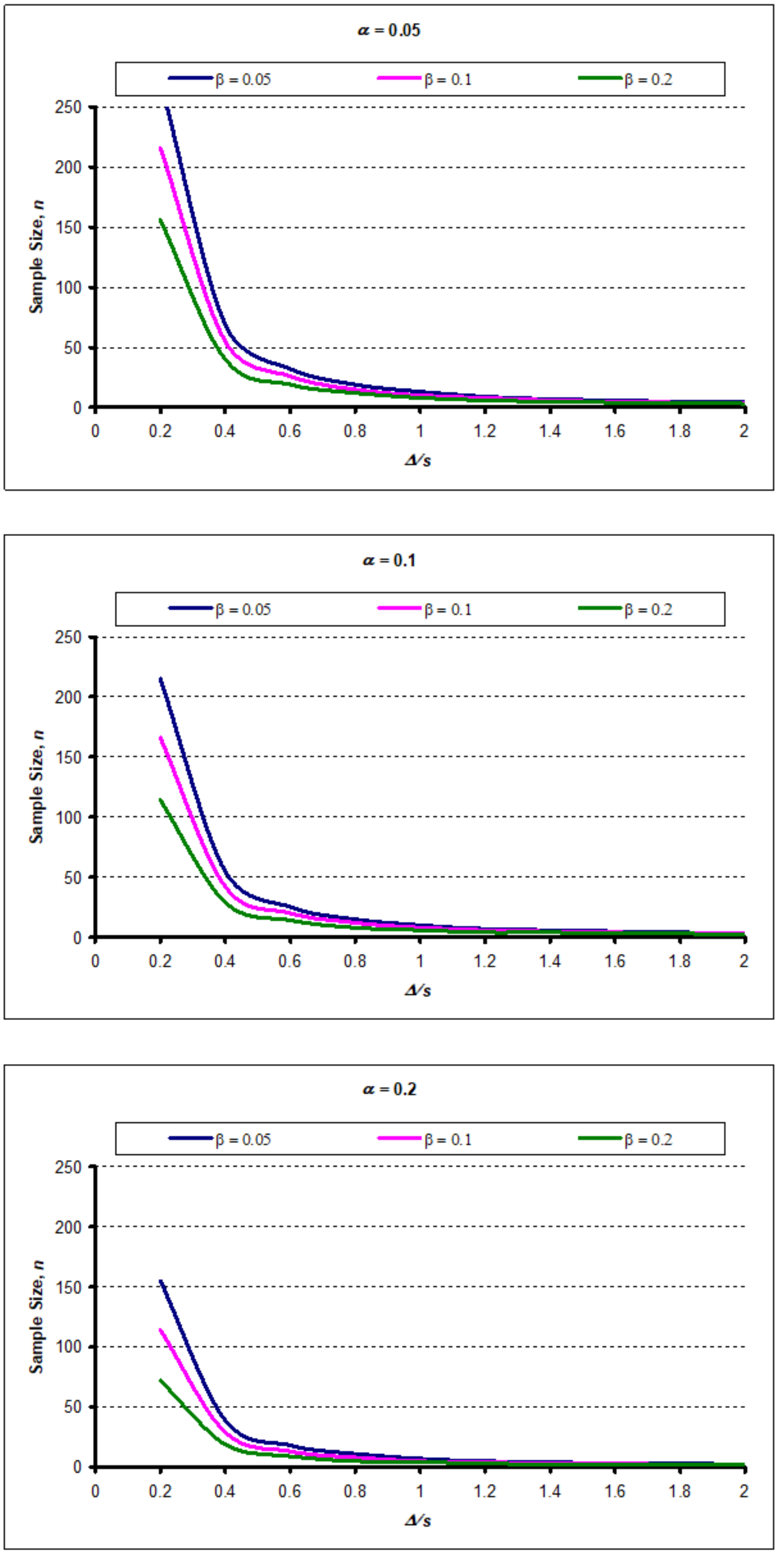


Figure D 1-1 Sample Size, n, Versus Ratios of Δ/s for Typical Values of α and β – One-Population, Normal Distribution

It is important to understand the influence of the four factors on the sample size selection:

(1) *Influence of s.* The sample size (n) increases as s increases. This means that if the data were highly variable, it would take a larger number of samples to detect a specified change. It is critical to eliminate or reduce any variability that is under the control of the Pilot Project Team. For example, variability due to different monitoring personnel, equipment, sampling protocols, or analytical laboratories should be controlled through quality control procedures and training.

(2) *Influence* Δ. The sample size (n) increases as Δ decreases. This means that one needs more data if a smaller change or difference is to be detected. The choice of the minimum change to be detected should be made with careful thinking. Comparing current discharges to legal or regulatory standards may provide a logical method to choose this parameter.

(3) *Influence of* $1-\alpha$. The sample size (n) increases as the $1-\alpha$ increases. Values of $(1-\alpha)$ and $(1-\beta)$ should be selected by considering the consequences of reaching an incorrect conclusion. Typically, the baseline condition is defined such that the consequences of falsely rejecting this condition would be more severe than the consequences of failing to detect the specified change. For BMP studies, the consequences of assuming a BMP is effective when it is not, might be considered to be worse than the consequences of concluding a BMP is ineffective when, in fact, it reduces the pollutant concentration by Δ. If this assumption is reasonable, one should specify a higher value for $(1-\alpha)$ than for $(1-\beta)$. Equivalently, α should be less than β.

(4) *Influence of Power of Detecting the Specified Minimum Change,* $(1-\beta)$. A higher power of detecting a specified minimum change requires a larger sample size.

D1.3.2 Single Data Set with Lognormal Distribution

Environmental variables, such as pollutant concentrations, often exhibit non-symmetric, right-skewed data distributions and hence may be better modeled as a lognormal distribution, rather than the symmetric, normal distribution. If a variable, X, is lognormally distributed, its log-transform ($Y = ln(X)$) would be normally distributed. Thus, the procedures described in the preceding section can be applied to the log-transformed variable. The variability of a lognormal variable is commonly expressed in terms of its coefficient of variation, CV_X, which is defined as the ratio (standard deviation, s_X, divided by mean, $\bar{x}$). For this case, the sample size, n, depends on the following factors:

- Coefficient of variation, CV_X. In the planning stage, this parameter is estimated based on prior data at similar sites or professional judgment. Refer to Table D 1-5 for the coefficient of variation data for BMP percent removals. The standard deviation of Y can be calculated as a function of CV_X using Equation D 1-2:

Equation D 1-2

$$s_y = \sqrt{\ln(1 + CV_X^2)}$$

- Minimum percent change, p, in the average value of the parameter of interest (e.g., average percent removal of a specified pollutant). The change in the log space of Y, Δ_y, can be calculated from p using Equation D 1-3:

Equation D 1-3

$$\Delta_y = \ln(1/(1-p)) = -\ln(1-p)$$

- Desired confidence level, $(1-\alpha)$. This parameter is defined in Section D1.3.
- Desired power of detecting the specified change, $(1-\beta)$. This parameter is defined in Section D1.3.

The parameters s_y, Δ_y, α, and β can then be used in Equation D 1-3 to calculate the required sample size, n. Alternatively, the USEPA program DEFT could be used to calculate the sample size. Figure D 1-2 shows plots of the sample size, n, for typical values of other parameters.

<u>Example D1-2</u>

An influent-effluent monitoring approach is planned to determine if a BMP can remove aluminum concentration by more than 20 percent. The expected number of independent storm events is 10 per year with a monitoring period of 3 years. Based on data from past similar BMP studies, the percent aluminum removal is assumed to follow a lognormal distribution with a coefficient of variation of 1.2. The desired confidence in reaching the correct conclusion when the percent removal is 0 percent is 0.95, and the desired power of detecting the minimum percent removal of 20 percent is 0.8. How many pilot sites should be selected? ** For Caltrans studies, the 90 percent confidence level should be used instead of the 0.95 value used in this example).

Using Equation D 1-2, Equation D 1-3, s_y is 0.944 and Δ_y is 0.223. 1-α = 0.95; and 1-β = 0.8. Using Equation D 1-1, the sample size is 113. Because 3 years of monitoring is planned with 10 events per year, 4 pilot sites should be selected, which would yield a total of 120 data points. Alternatively, 2 pilot sites may be monitored for 6 years to obtain the necessary data.

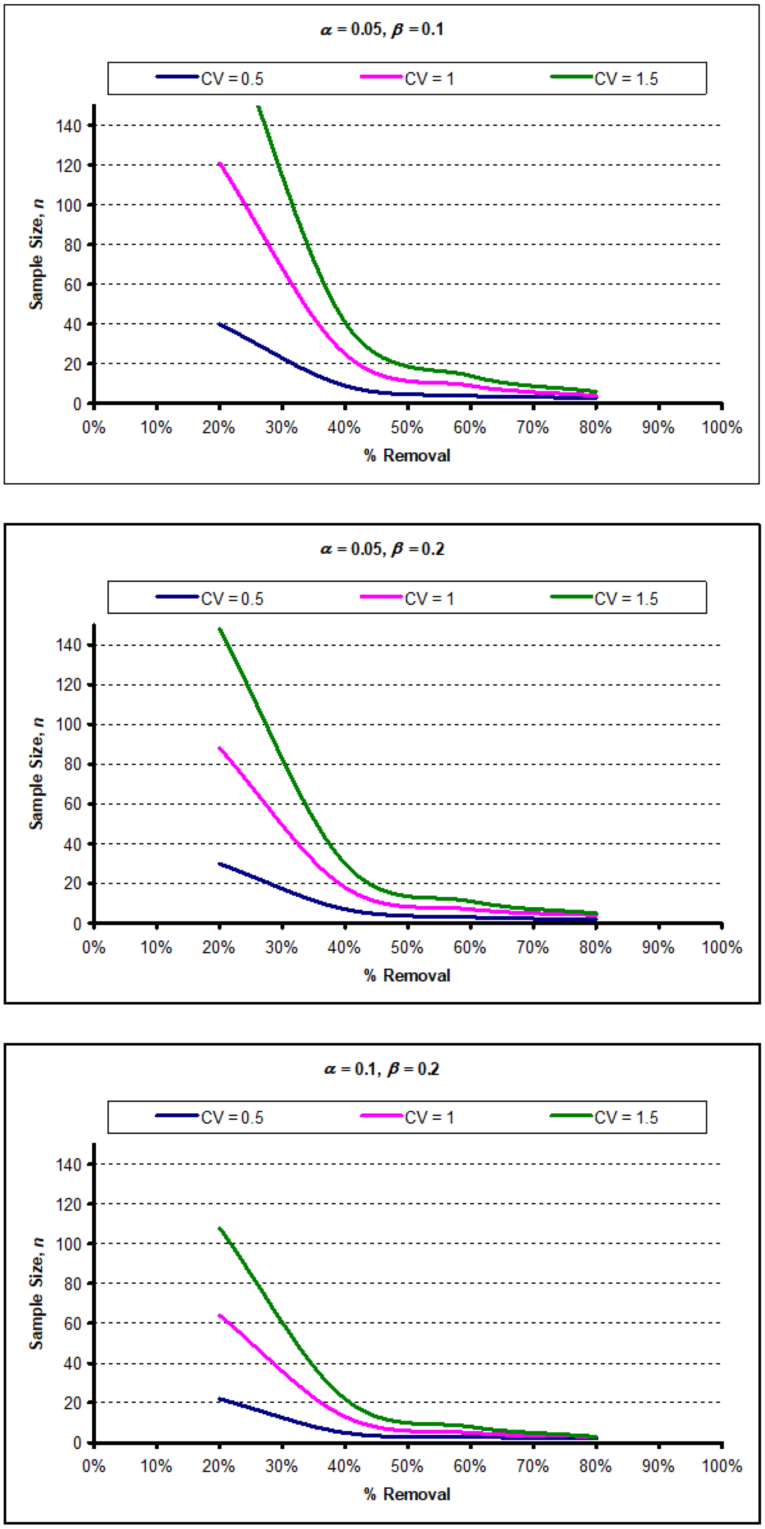


Figure D 1-2 Sample Size, n, Versus Percent Removal for Typical Values of α and β – One-Population, Lognormal Distribution

D1.3.3 A Single Data Set with Neither Normal nor Lognormal Distribution

No simple procedure is available to estimate the sample size in this case. Consultation with a statistician may be needed.

D1.4 Study Questions Involving Two Independent Data Sets

A typical study question related to two independent data sets is as follows: for pollutant concentrations measured with a paired watershed monitoring approach, is the average pollutant concentration higher in one of the watersheds? The paired watershed approach in this case entails comparison of water quality data from two similar watersheds: one control (undisturbed) watershed and one treatment (disturbed) watershed. However, the specific monitoring stations in the two watersheds are not paired with each other. Thus, the data from the two watersheds define two independent data sets. The procedure to estimate the sample size depends on the data distribution (normal, lognormal, or neither) and whether the variances of the two data sets are equal. As described in Appendix D4, the Shapiro-Wilk W test could be used to test the normality assumption for both the original and log-transformed data. The Levene test, also described in Appendix D4, could be used to test equality of variances. The following cases are considered in this section:

(1) Both data sets follow a normal distribution and have an equal variance.

(2) Both data sets follow a lognormal distribution and have an equal variance in log space (which means an equal coefficient of variation in the original arithmetic space).

(3) Neither normal nor lognormal distribution can be assumed for one or both data sets; or variances are unequal in both original and log-transformed space.

D1.4.1 Two Independent Data Sets with Normal Distribution and Equal Variance

The sample size, *n*, depends on the following factors:

- The common standard deviation of the two data sets, s_p. The following equation can be used to calculate s_p:

Equation D 1-4

$$s_p = \sqrt{\frac{(m-1)s_X^2 + (n-1)s_Y^2}{m+n-2}}$$

where s_p = common sample standard deviation of the two data sets

m = sample size of *X*

n = sample size of Y

s_X = sample standard deviation of X

s_Y = sample standard deviation of Y

- Since no data would have been collected at the planning stage, the standard deviation is typically estimated based on prior data from similar sites or professional judgment. In the interim data review stage (i.e., after collecting data for one or two years), the initial standard deviation estimate can be checked and revised if necessary. It should be noted that for typical variables of interest in BMP studies (e.g., pollutant concentrations), the assumption of equal variances is often not valid. For example, if the BMP is effective in reducing pollutant concentration, the effluent concentrations would be lower and the variance is also likely to be smaller. In such a case, the coefficient of variation is more likely to be similar, and the procedure described in the next section may be more appropriate to use.
- The minimum change in the mean value between the two data sets that should be detected, Δ.
- Desired confidence level, (1-α). This parameter is defined in Section D1.4.
- Desired power of detecting the specified change, (1-β). This parameter is defined in Section D1.4.

For given values of s_p, Δ, α, and β, one can calculate the minimum sample size, n. Again, an exact method is available in JMP and other similar software packages. However, its use requires knowledge of advanced statistical concepts. A much simpler procedure that is approximate is provided in the USEPA document QA/G-9S. The approximate method also assumes a normal distribution and equal variance for the two data sets. Valid computational simplifications are described in EPA QA/G-4 (USEPA 2006b). The USEPA document EPA QA/G-9S gives Equation D 1-5 to calculate n:

Equation D 1-5

$$n = \frac{2s_p^2\left(z_{1-\alpha} + z_{1-\beta}\right)^2}{\Delta^2} + \frac{z_{1-\alpha}^2}{4}$$

where n = sample size

s_p = common sample standard deviation of the two data sets

$z_{1-\alpha}$ = value of standard normal variate at 1-α probability

$z_{1-\beta}$ = value of standard normal variate at 1-β probability

Δ = minimum change in the mean value between the two data sets to detect

Note that *n* is the sample size for each of the two groups. The USEPA program DEFT can calculate the sample size using the above equation. Figure D 1-3 shows plots of *n* for typical values of other parameters. Note that Δ is expressed in its normalized form; i.e., in terms of the ratio Δ/s_p.

Example D1-3

A paired-watershed BMP monitoring approach is planned for a BMP study to address the following study question: Is the average aluminum concentration in a treatment (disturbed) watershed higher than that in the control watershed by more than 10 µg/L? The expected number of storm events at the study site is 20 per year. Based on data from past similar BMP studies, the control and treatment watershed data on aluminum concentrations are assumed to follow a normal distribution with a common standard deviation of 20 µg/L. The desired confidence in reaching the correct conclusion when the average control and treatment watershed aluminum concentrations are the same is 0.95, and the desired power of detecting a difference in the average aluminum concentration difference between the two watersheds of 10 µg/L is 0.8. How many years of monitoring should be planned? Note that for Caltrans studies, the 90 percent confidence level should be used.

For this example, s_p = 20 µg/L, Δ = 10 µg/L, α = 0.05, and β = 0.2. From Figure D 1-3 the sample size is calculated to be 51 for each of the two groups. Since 20 events per year are expected, 3 years of monitoring should be planned, which would yield a total of 60 data points in each group.

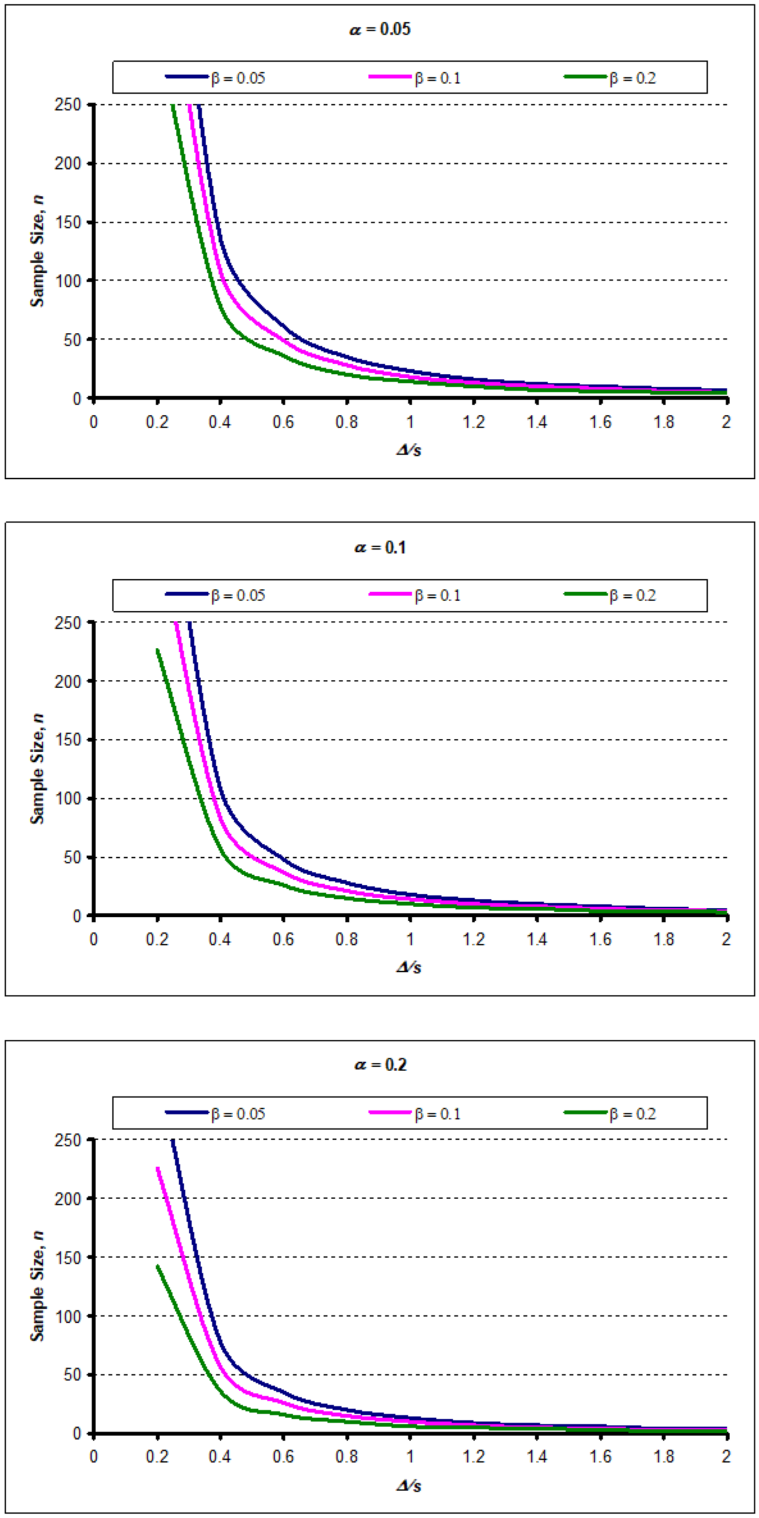


Figure D 1-3 Sample Size n, Versus Ratios of Δ/s for Typical Values of α and β – Two-Population, Normal Distribution

D1.4.2 Two Independent Data Sets with Lognormal Distribution and Equal Coefficient of Variation

As noted previously, data sets in BMP studies are often displayed as right-skewed distributions and can be better modeled as lognormal variables. In addition, the standard deviation of a data set is likely to be proportional to the mean, which means that the coefficient of variation of the two data sets is likely to be the same. If a variable, *X*, is lognormally distributed, its log-transform ($Y = \ln(X)$) would be normally distributed. If the coefficient of variation of two lognormally-distributed variables is the same, the variances of the log-transformed variables would be the same. Therefore, the procedures described in the preceding section can be directly applied to the log-transformed variables.

The sample size, *n*, depends on the following factors:

- The common coefficient of variation of the two data sets, CV_X. In the log space, the common standard deviation of the two data sets, s_{pL}, can be calculated as a function of *CV* using the following equation (Equation D 1-6):

Equation D 1-6

$$s_{pL} = \sqrt{\ln(1 + CV_X^2)}$$

- The minimum percent change, *p*, in the mean value that should be detected with high confidence. The change in the mean value in the log space, Δ_L, can be calculated as a function of *p* using the following equation (Equation D 1-7):

Equation D 1-7

$$\Delta_L = \ln(1/(1-p)) = -\ln(1-p)$$

- Desired confidence level, $(1-\alpha)$. This parameter is defined in Section D1.4.
- Desired power of detecting the specified change, $(1-\beta)$. This parameter is defined in Section D1.4.

For given values of s_{pL}, Δ_L, α, and β, the required sample size can be calculated from Table D 1-5 or using the USEPA program DEFT. Figure D 1-44 shows plots of *n* for typical values of the other parameters.

Example D1-4

This example is similar to Example D1-3, except that each data set is assumed to be lognormally distributed with a common coefficient of variation of 1.2. The study question of interest is: How many years of monitoring should be planned to detect a difference in the average aluminum concentration of the treatment watershed and the control watershed of more than 30 percent?

For this example, CV = 1.2 (and hence s_{pL} = 0.944), Δ_L = -ln(1-0.3) = 0.357, α = 0.05, and β = 0.2. From Equation D -5, the sample size is calculated to be 88 for each of the two data sets. Since 20 events per year are expected, 5 years of monitoring should be planned, which would yield a total of 100 data points in each group.

D1.4.3 Two Independent Data Sets with Neither Normal nor Lognormal Distribution; or with Unequal Variances and Coefficients of Variations

The estimation of the sample size for this case requires the use of advanced methods for which consultation with a statistician may be necessary. The USEPA guidance document EPA 230-R-94-004 (Section 6.2 of USEPA 1994) describes a rigorous method to estimate the sample size for this case. A more approximate method is included in USEPA guidance document EPA QA/G-9S (Box 3-18 on Page 64).

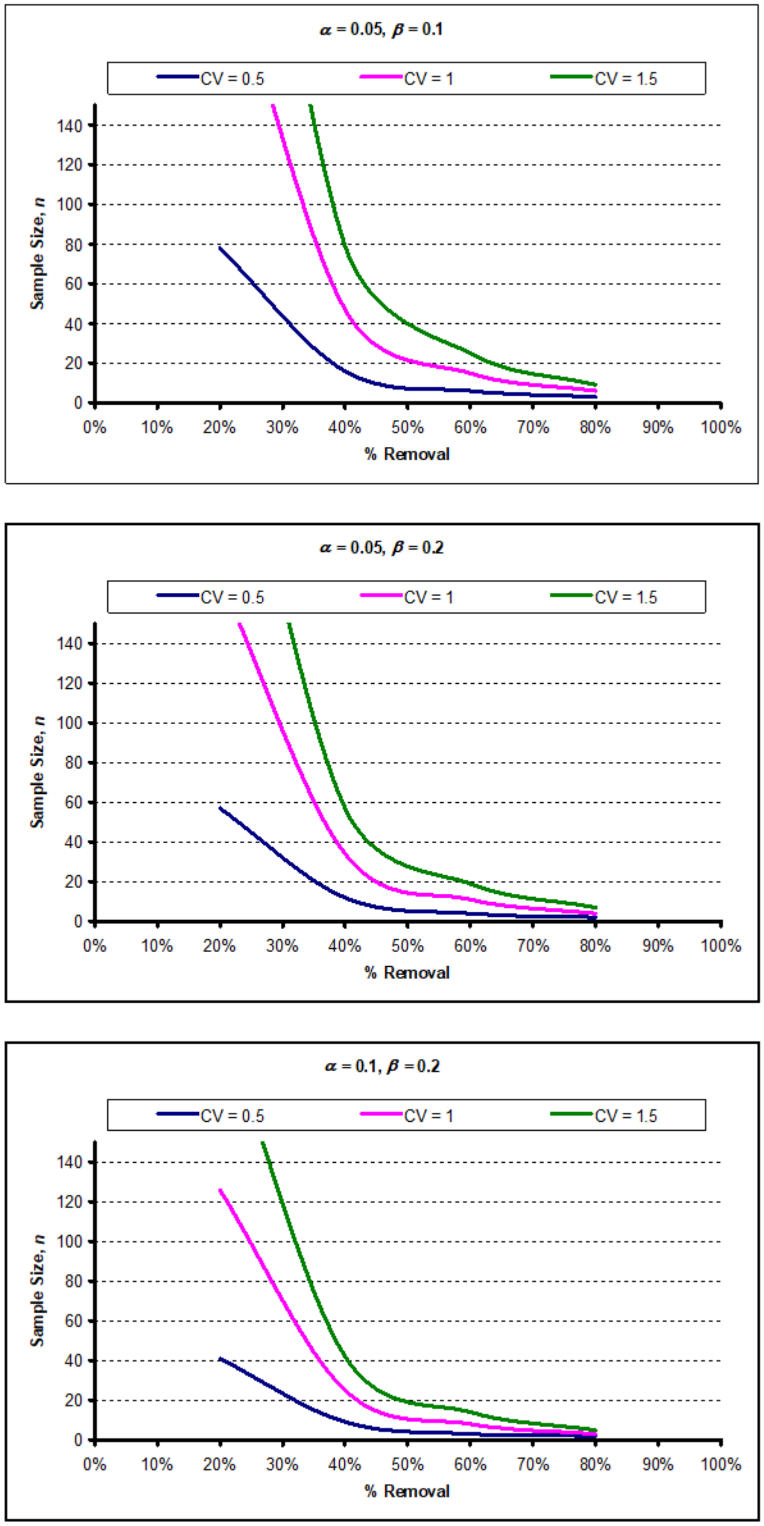


Figure D 1-4 Sample Size, n, Versus Percent Removal for Typical Values of α and β – Two-Population, Lognormal Distribution

D1.5 Study Questions Involving Three or More Independent Data Sets

This topic requires the use of advanced methods, which are not covered in this document. Consult with a statistician to estimate the sample size.

D1.6 Statistical Procedures for Estimating Sample Size in the Interim Data Review

At the planning stage, no actual data would have been collected. Consequently, data from previous similar studies or professional judgment may be used to make assumptions about data distributions and to estimate data variability. After one or two years of sampling, actual data would be available that should be used for an interim data review to check on the assumptions made during the planning stage and determine whether additional samples, beyond those already planned, would be necessary to meet the study objectives. A minimum of eight interim data points will be needed to verify assumptions of data distribution and data variability.

For one-population studies (i.e., those involving a single independent data set), the interim data should be used to check whether the data might be assumed to follow a normal or lognormal distribution, or neither. If a normal distribution is appropriate to assume, the interim data should be used to estimate the sample standard deviation. If a lognormal distribution is appropriate to assume, the interim data should be used to estimate the sample coefficient of variation. Using the appropriate data distribution and the estimated standard deviation or coefficient of variation, one should recalculate the necessary sample size and decide whether more samples than initially planned would be necessary to answer the study question of interest.

Example D1-5

We will continue with Example D1-4 involving two independent data sets. During the planning stage, each data set was assumed to be lognormally distributed with a common coefficient of variation of 1.2. After one year of sampling, 20 sample data points were available for each of the two watersheds. The Shapiro-Wilk W test showed that the assumption of a lognormal distribution was reasonable for each data set. Furthermore, the coefficient of variation of each data set was similar, which was estimated to be 1.5. In contrast, the common coefficient of variation was assumed to be 1.2 in the planning stage to calculate the sample size. With the revised estimate of the coefficient of variation, Equation D 1-5 yields a sample size of 116 for each watershed. Since 20 samples were collected in the first year, an additional 96 samples would be necessary for each watershed. This would suggest that monitoring would have to continue for 5 more years, instead of 4 more years as initially planned.

Alternatively, one may relax the specification of the minimum percent difference in aluminum concentration that should be detected. For example, specify the minimum percent difference in the average aluminum concentration to be 40 percent, instead of 30 percent specified during the planning stage. With that change, the sample size is calculated to be 57 for each watershed. Thus, an additional 37 samples would be needed for each watershed. This would suggest that, if the revised detection threshold is acceptable, two more years of sampling would be adequate.

Appendix D2 How to Examine Data Quality and Detect Possible Outliers in the Data

D2.1 Purpose and Organization

This appendix presents a collection of graphical and numerical methods that should be used to evaluate data quality and consistency and identify potential outliers in a given dataset. Potential outliers are measurements that are extremely large or small relative to the rest of the data, and, therefore, are suspected of not belonging to the population whose characteristics are being evaluated.

For example, suppose one is trying to evaluate the effectiveness of a particular BMP and the data compiled for the statistical analysis are obtained mostly from that BMP. Also, suppose that a few measurements from another BMP are included in the dataset unknowingly. If the effectiveness of the two BMPs is substantially different, the few measurements from the second BMP may be much larger or smaller than the bulk of the data from the first BMP. This appendix's graphical methods help identify such anomalous measurements. Note that the analysis of the combined dataset from the two BMPs may produce results that are representative of neither BMP.

Potential outliers could also result from transcription errors, data-coding errors, or measurement system problems. Outliers may also represent true extreme values of a distribution that result from a greater variability in the data than expected.

It should be emphasized that outlying or influential observations should not be removed from a data set without explicit confirmation of a measurement error or of other factors that identify the measurement as extraneous to the population of interest. Both the failure to remove true outliers and the removal of false outliers are undesirable outcomes because both could lead to erroneous conclusions. Therefore, the decision of whether to exclude any data should be made with great caution and care.

The next section provides a decision flowchart that identifies the sequence of key steps and decisions to be made based on the results of certain steps. Subsequent sections describe the analysis to be performed in each step, provide guidance in making appropriate decisions, and illustrate the implementation of the step with typical examples drawn from BMP pilot studies.

Throughout Appendix D, use is made of the statistical software package called JMP, developed by the SAS Institute (https://www.jmp.com/en_us/home.html), to produce the graphical plots and numerical results for the analysis of test examples. Any other statistical software package could also be used with generally the same sequence of steps.

D2.2 Decision Flowchart

Figure D 2-1 shows a flowchart of the key steps and decisions in examining data quality and identifying potential outliers. The seven key steps are:

1. Compile data in a format suitable for the selected statistical software package.
2. Prepare and interpret graphical displays.
3. Prepare and interpret numerical summaries.
4. Assess whether results of Step 2 or 3 suggest the presence of potential outliers.
5. Check potential outliers identified in Step 4 for specific errors or not being representative.
6. Assess whether any of the potential outliers could be confirmed as actual outliers.
7. Correct or exclude the confirmed outliers from further analysis.

A description of each step follows.

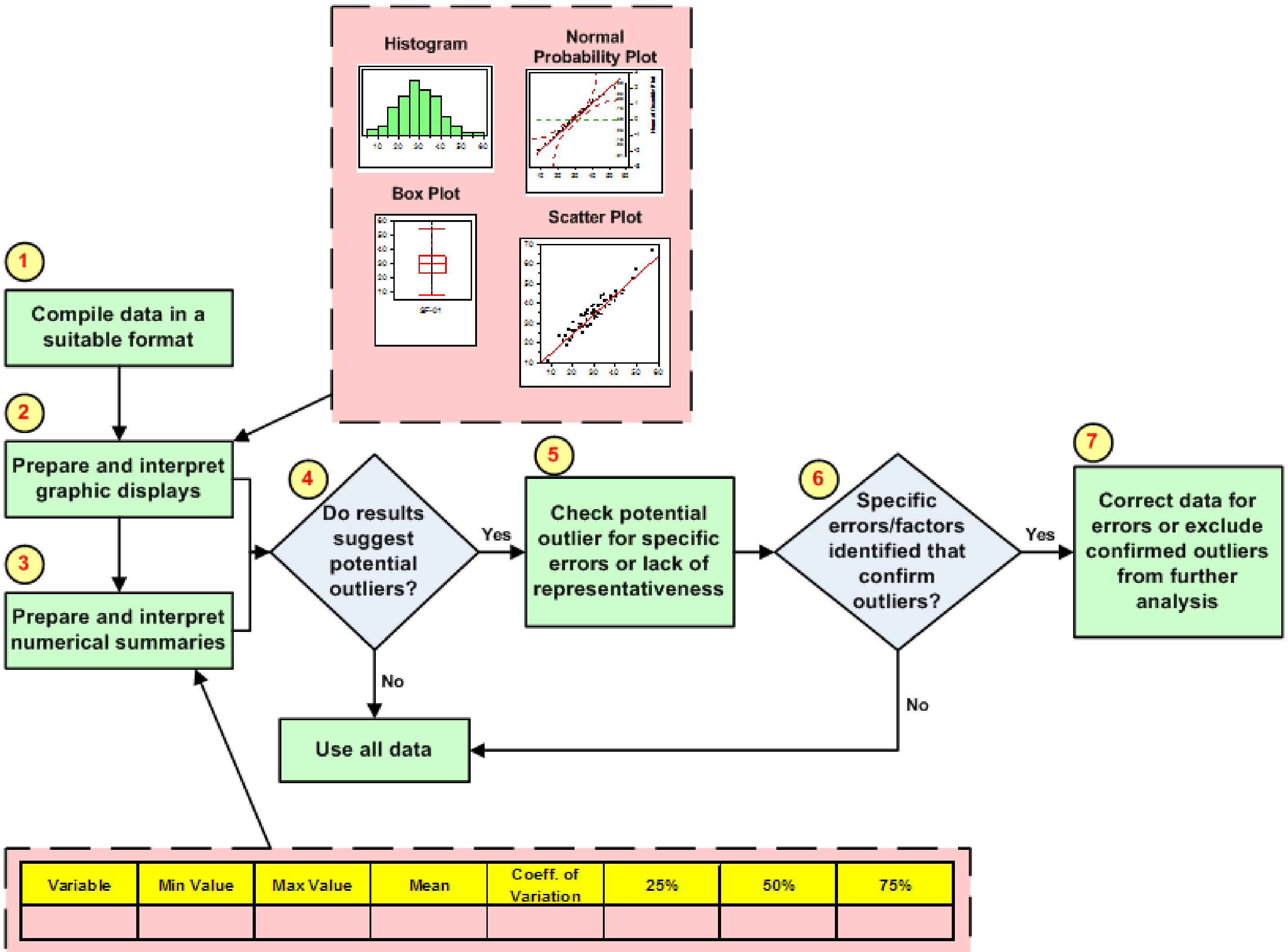


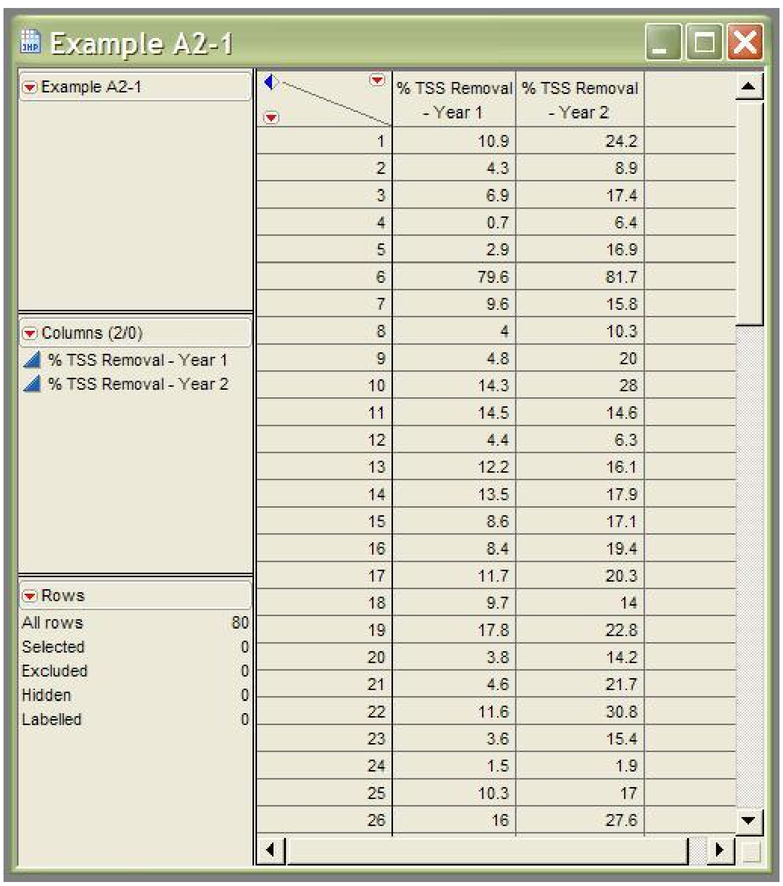
Figure D 2-1 Decision Flowchart for Investigating Data Quality and Potential Outliers

D2.2.1 <u>Step 1. Compile Data in a Format Suitable for the Selected Statistical Software Package</u>

Statistical software packages typically use a tabular format for inputting data. The columns represent the different variables that are to be analyzed for quality and each row represents one sample data point. Example D2-1 illustrates the input data table prepared for the JMP software. The example will also be used to illustrate subsequent steps.

<u>Example D2-1</u>

Suppose TSS data were collected from the influent and effluent of 80 sand filters in the Los Angeles Basin during Year 1 and Year 2. At the end of the Year 1 storm season, the sand filters underwent certain design modifications to improve their effectiveness. The JMP input data table below comprises one row for percent reduction data from each sand filter and two columns – one for Year 1 and one for Year 2. Figure D 2-2 shows the JMP input data.



	% TSS Removal - Year 1	% TSS Removal - Year 2
1	10.9	24.2
2	4.3	8.9
3	6.9	17.4
4	0.7	6.4
5	2.9	16.9
6	79.6	81.7
7	9.6	15.8
8	4	10.3
9	4.8	20
10	14.3	28
11	14.5	14.6
12	4.4	6.3
13	12.2	16.1
14	13.5	17.9
15	8.6	17.1
16	8.4	19.4
17	11.7	20.3
18	9.7	14
19	17.8	22.8
20	3.8	14.2
21	4.6	21.7
22	11.6	30.8
23	3.6	15.4
24	1.5	1.9
25	10.3	17
26	16	27.6

Figure D 2-2 JMP Input Data Table

D2.2.2 Step 2. Prepare and Interpret Graphical Displays

Graphical displays are prepared in the form of data plots. Three common data plots for a single variable of interest (e.g., TSS removal) are histogram, box plot, and normal probability plot. Conversely, when a relationship between two variables is of interest, a scatter plot is used. For example, one may be interested in analyzing the relationship between TSS removal in two different monitoring years. In this case, a scatterplot between TSS removal for the two years would be of interest. The preparation and interpretation of each data plot are described below. Different data plots often provide information about the same data features. The conclusions drawn from one data plot should be confirmed using information from other data plots. Multiple lines of evidence supporting a particular conclusion increase the confidence in that conclusion.

D2.2.2.1 Histograms

A histogram is a type of bar chart in which the data range is divided into bins, the data are sorted into the bins, and the number of data points (and/or the proportion of data points) in each bin is displayed. Figure D 2-3 shows the histogram (including annotated features) generated by JMP for the Year 1 data shown in Figure D 2-2.

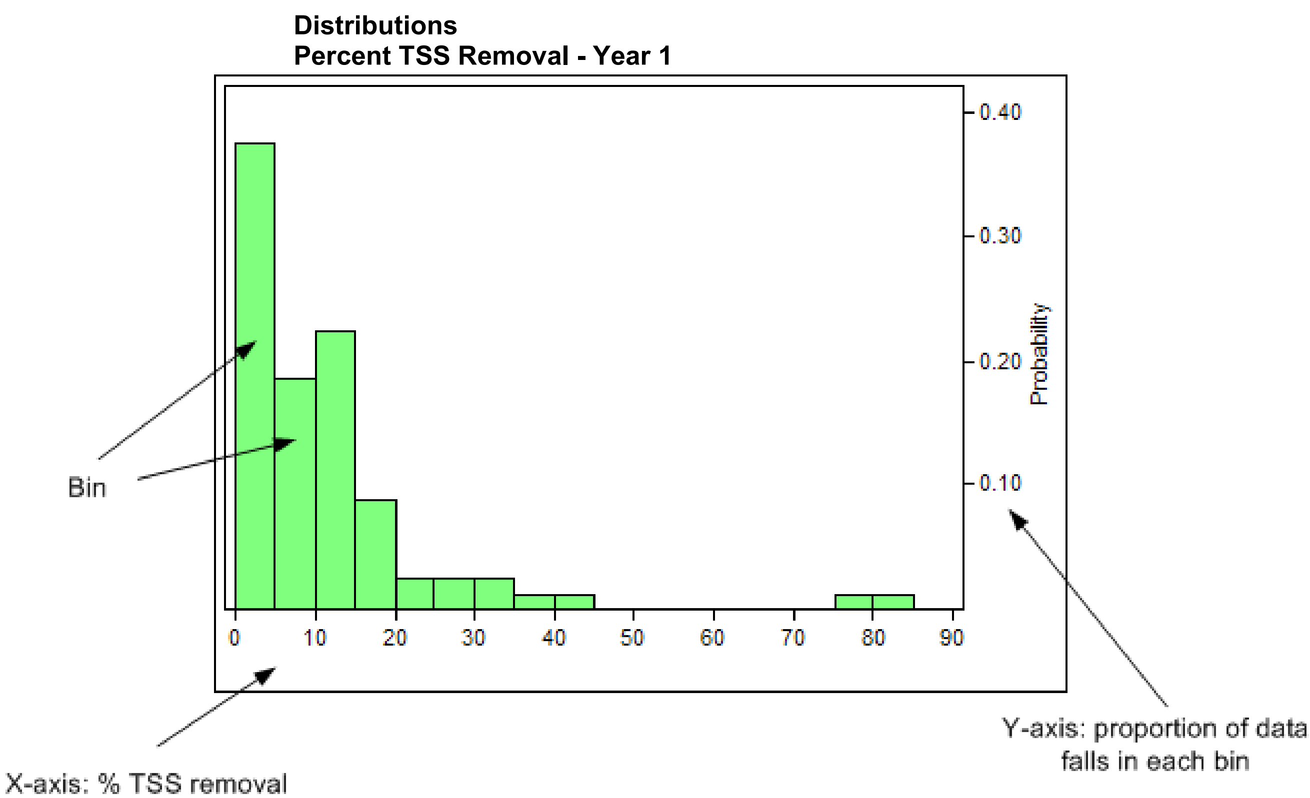


Figure D 2-3 Histogram for "Year 1" Data

A histogram is useful in answering the following questions:

1. Does the data distribution have a single "mode" or more than one "mode"?

The mode of a distribution is the value that occurs most frequently and is indicated by the peak of the histogram. Figure D 2-4 shows examples of several histograms to highlight different conclusions that could be drawn. Examples of one-mode (unimodal) and two-mode (bimodal) distribution are included.

A bimodal distribution generally suggests the data may be a combination of two separate populations. The use of standard statistical methods that assume a unimodal distribution (such as normal or lognormal) would be inappropriate for a bimodal distribution.

2. Is the distribution symmetric or skewed?

If the histogram indicates a unimodal distribution, one can further check whether the distribution is symmetric or skewed. A skewed distribution typically shows a longer tail on one side of the mode than the other. For water quality data, the distribution tends to be "right" skewed; that is, a longer tail toward higher values. Figure D 2-4 shows histograms of symmetric and right-skewed distributions. A normal distribution is symmetric around a single mode. The use of standard statistical methods that assume a normal distribution would be inappropriate for a skewed distribution. A common remedy in case of right-skewed distributions is a data transformation such as log transformation (i.e., taking natural logarithms of the raw data) that would make the distribution symmetric.

3. Are there potential outliers?

A histogram may show a fairly well-behaved distribution except for some isolated data points that are far different from the rest of the data. Figure D 2-4 shows an example of a histogram that suggests one potential outlier.

<u>Interpretation of the Histogram for Example D2-1</u>

The histogram shown in Figure D 2-3 for Example D2-1 reveals the following features:

- The data show a single mode in the range of 0 to 5 percent TSS removal.
- The data distribution is highly nonsymmetric with a longer tail toward the higher values.
- Two values greater than 75 percent are well separated from the rest of the data and may be considered to be potential outliers. Some follow-up on those two sand filters and some attempt to confirm the validity of these two measurements would certainly be in order.

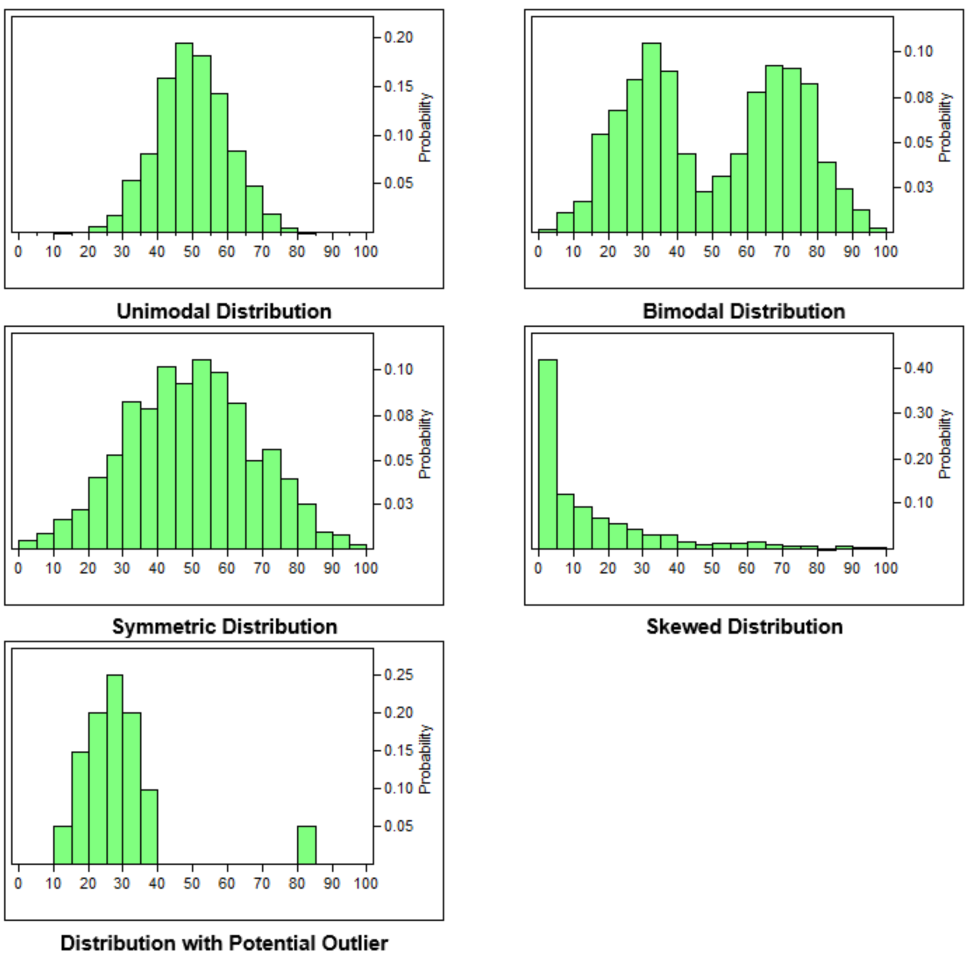


Figure D 2-4 Examples of Different Types of Distribution

D2.2.2.2 Box Plots

In the JMP software package, this type of graphical plot is identified as the Outlier Box Plot. It is a schematic that displays key aspects of the data distribution and identifies extreme points that could be potential outliers. Another common name for a box plot is a box-and-whisker plot.

A box plot displays key percentiles of the data distribution. A *p*-th percentile of a data distribution defines a data value such that *p* percent of all data would be equal to or below that value. Thus, for example, a 75th percentile would be such that 75 percent of all data would be less than or equal to it. The 75th percentile is also called the *upper quartile*. Correspondingly, the 25th percentile is called the *lower quartile.* The difference between the upper and lower quartiles is called the *interquartile range.* The 50th percentile is commonly referred to as the *median*. Thus, 50 percent of the data would be below the median and 50 percent would be above it.

The JMP software (and other similar software packages) provides the option of plotting the *mean diamond*. This diamond shows the mean and the 95th percentile confidence interval around the mean. Note that for Caltrans studies, the 90 percent confidence interval should be presented.

Figure D 2-5 shows a box plot (as well annotated features) for Example D2-1.

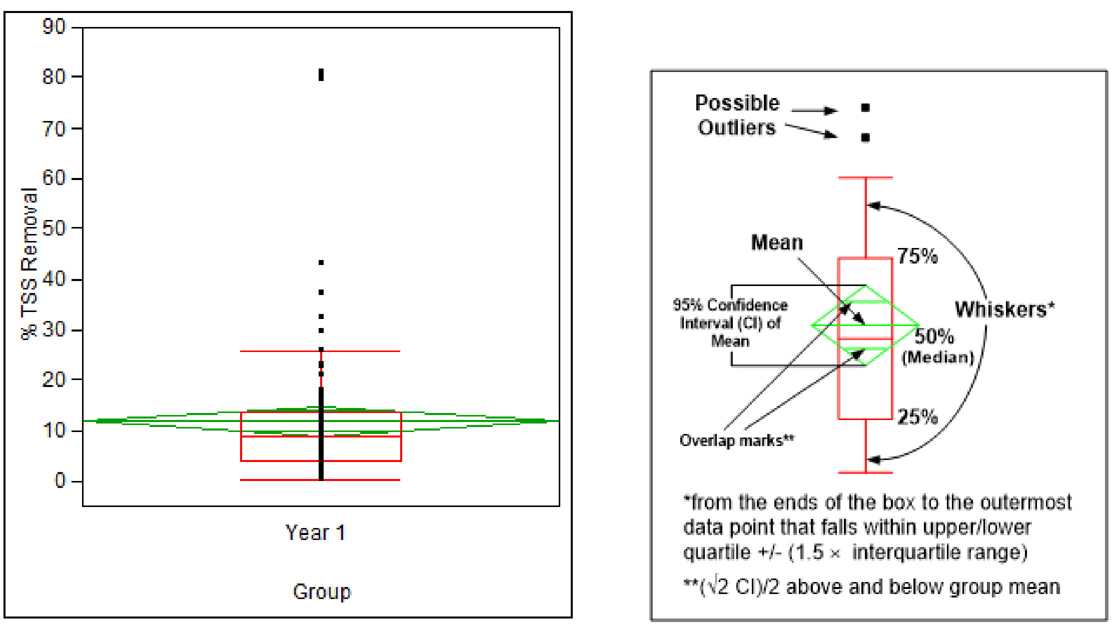


Figure D 2-5 Box Plot for Example D2-1

A box plot has a box and lines (whiskers) drawn on either side of the box. The upper edge of the box is the 75th percentile and the lower edge of the box is the 25th percentile.

The line inside the box is the median (i.e., the 50th percentile). On the side of the 75th percentile, the whisker extends from the end of the box to the highest data point that is still lower than the value of (75th percentile + 1.5 × (interquartile range)). Similarly, on the side of the 25th percentile, the whisker extends from the end of the box to the lowest data point that is still higher than the value of (25th percentile + 1.5 × (interquartile range)). The basic idea is to draw the whiskers such that most of the data would be inside the end points of the whiskers. If any data values do plot outside the end points of the whiskers, they may be considered to be anomalous and hence potential outliers. In the JMP outlier box plot, such potential outliers are shown as dots.

A box plot is useful in answering the following questions:

1. Is the data distribution symmetric or skewed?

If the upper box (above the median) and whisker are approximately the same length as the lower box and whisker, the data are distributed symmetrically. If the upper box and whisker are longer than the lower box and whisker, the data are right-skewed. Conversely, if the upper box and whisker are shorter than the lower box and whisker, the data are left-skewed. A conclusion about data symmetry drawn from a box plot should be checked against that drawn from a histogram.

If the mean diamond is plotted, then the comparison of the mean and median provides an additional check on data symmetry. If the mean and median are about the same, the data are distributed symmetrically. If the mean is greater than the median, the data are right-skewed. If the mean is less than the median, the data are left-skewed.

2. Are there potential outliers?

If any data points plot outside the whiskers (appearing as dots in the plot outside the whiskers), these may be considered to be potential outliers and require further investigation. Again, information from a histogram of the same data should be used to confirm the presence of potential outliers.

Since the box plot is compact (essentially one-dimensional), several plots for different data sets could be placed on a single graph. This provides a simple, yet informative tool to compare different groups of data. For example, box plots for different years of BMP performance could be placed on a single graph. Such a graph will provide a quick visual comparison of BMP effectiveness over time. Figure D 2-6 provides such a multi-group box plot for Example D2-1. For this example, Year 2 appears to have a higher TSS percent removal.

As noted previously, the mean diamond for each data set (shown in green in Figure 2-6) shows the sample mean and the 95 percent confidence interval around the mean. In addition, horizontal lines are shown inside the mean diamond above and below the sample mean. They provide a simple visual comparison between the means of different data sets. For data sets with equal sample sizes, if the intervals defined by these horizontal lines overlap, the means of the two data sets are not significantly different at the 95 percent confidence level. Note that for Caltrans studies, the 90 percent confidence level should be used.

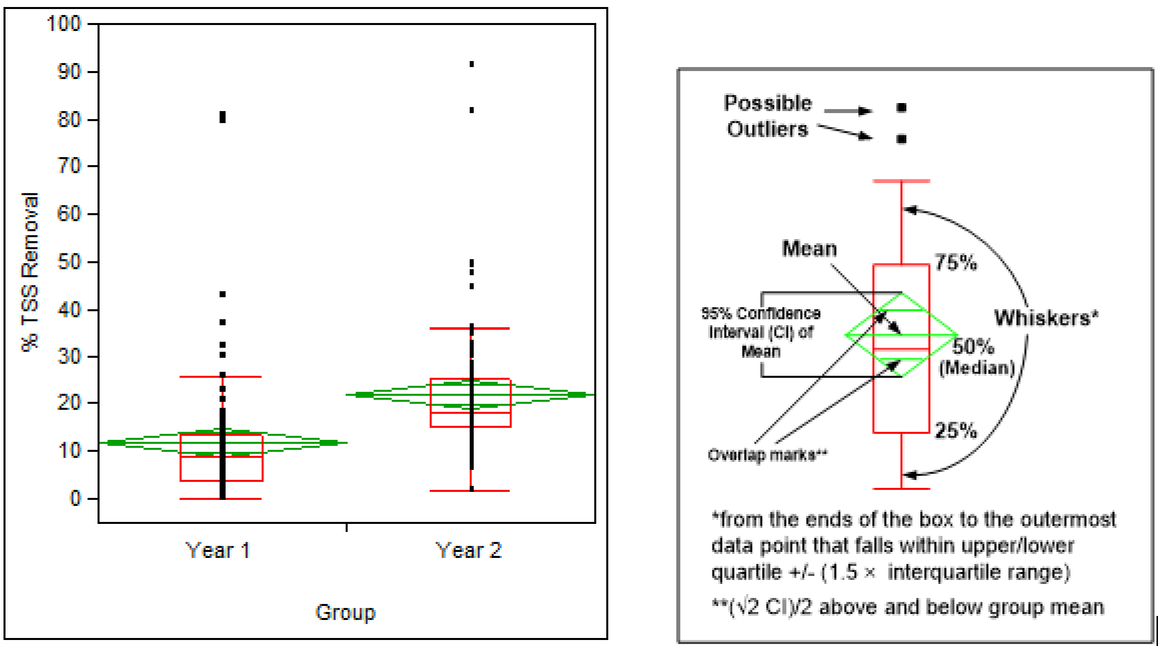


Figure D 2-6 Example of Multi-Group Graph Box Plots

<u>Interpretation of the Box Plot for Example D2-1</u>

The box plot shown in Figure D 2-6 for Example D2-1 confirms the two observations made from the histogram; namely, the data distribution is nonsymmetric (right-skewed) and there are some extreme values with percent removal around 80 percent.

D2.2.2.3 Normal Probability Plot

Many standard statistical tests assume that the data are normally distributed. It is important to get an early indication of whether this assumption is reasonable. A normal probability plot provides a visual check on the normal distribution assumption.

A normal probability plot is a graph that plots different percentiles of the actual data against the corresponding percentiles of a standard normal distribution. If the graph is approximately linear, this is an indication that the data are normally distributed. This finding should be confirmed with a formal test (formal tests for verifying the normal distribution assumption are described in Appendix D4). Figure D 2-7 shows the normal probability plot for Example D2-1.

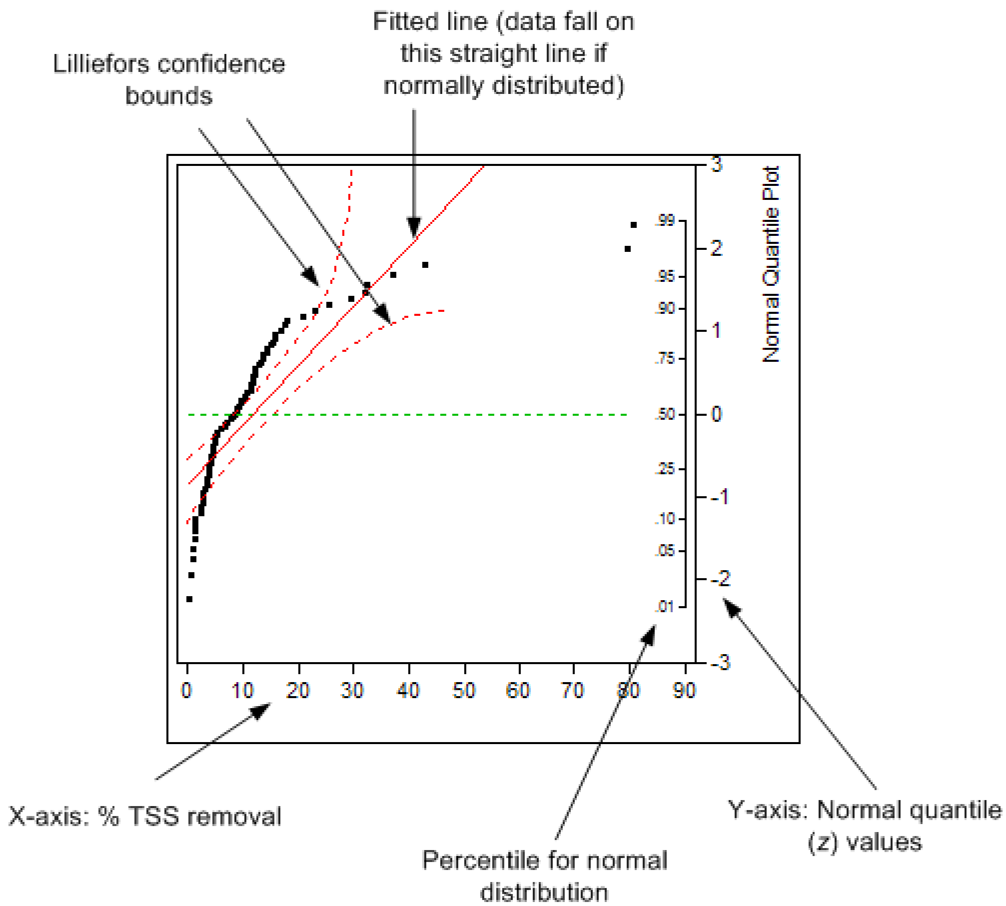


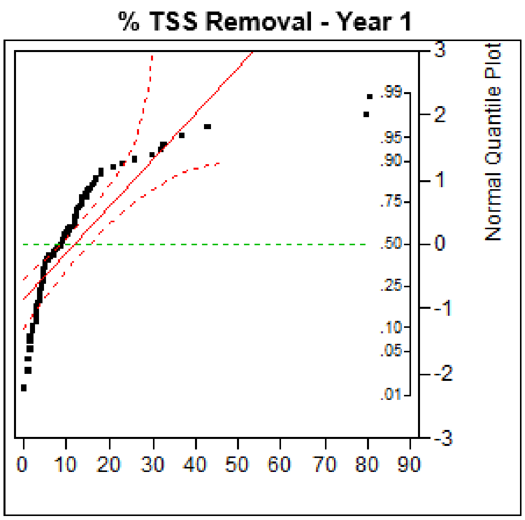
Figure D 2-7 Normal Probability Plot for Example D2-1

If the graph is not linear, data may be transformed and a new plot prepared using the transformed data. A common data transformation for right-skewed data is the log transformation. Environmental data (such as pollutant concentrations) are often right-skewed. For such data, the normal probability plot would not be linear when the raw data are used, but may become linear when log-transformed data are used. Figure D 2-8 illustrates this behavior for Example D2-1.

<u>Interpretation of the Normal Probability Plot for Example D2-1</u>

The normal probability plots shown in Figure D 2-7 and Figure D 2-8 reveal the following features:

- The plot in Figure D 2-7 is strongly nonlinear, suggesting that the data are not normally distributed.
- Log-transformed values in Figure D 2-8 result in a more linear normal probability plot, suggesting that a lognormal distribution would be more appropriate for these data.



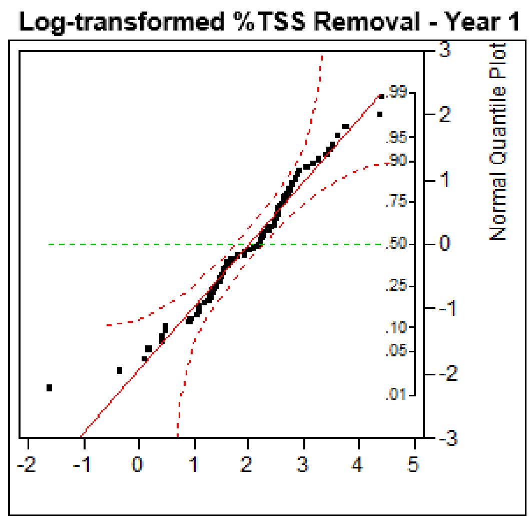


Figure D 2-8 Normal Probability Plot for Log-transformed Data

D2.2.2.4 Scatter Plot

This data plot is of interest when one is trying to explore the relationship between two variables. Typical examples include influent-effluent measurements on a particular type of BMP, measurements on a BMP at several sites over two years, or the influence of a particular storm characteristic (such as rainfall intensity) on BMP performance. This data plot helps to answer such questions as: Do the effluent concentrations increase with increasing influent concentrations? Are Year 2 concentrations consistently lower than Year 1 concentrations? Does BMP performance depend on certain storm characteristics?

A scatter plot is a useful graphical tool to visually inspect the relationship or association between two variables that are commonly denoted as Y and X. The variable Y is generally a response variable (such as percent pollutant removal in BMP studies) and X is the explanatory variable that may help explain how Y varies as a function of X. For each value of X, there is a corresponding value of Y. With the common graphing convention, Y values are plotted on the vertical axis and the corresponding X values are plotted on the horizontal axis. The nature of this Y versus X plot reveals if there is a linear or nonlinear relationship between the two variables.

Figure D 2-9 shows a scatter plot for Year 2 percent removal versus Year 1 percent removal. Each data point represents information about one specific filter.

If the pairs of (Y, X) points generally fall on a straight line, this would suggest a linear relationship between the two variables. If the slope of the line is negative, this suggests a negative correlation between the two variables (i.e., Y decreases as X increases). Conversely, a positive slope of the line suggests a positive correlation between the two variables.

If the pairs of (Y, X) points do not plot on a straight line, but do fall on a curve, this would suggest a nonlinear relationship between the two variables. In such a case, it is useful to try different data transformations to see whether the relationship could become linear with some transform. A common transform is the log-transform on Y, X, or both.

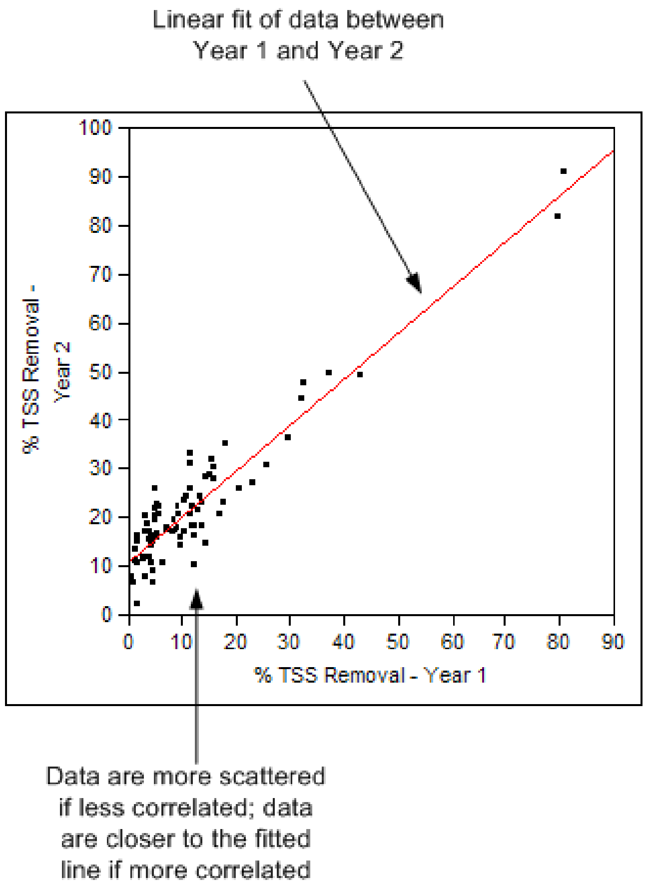


Figure D 2-9 Scatter Plot for Year 2 Percent Removal Versus Year 1 Percent Removal

Interpretation of the Scatter Plot for Example D2-1

The scatter plot shown in Figure D 2-9 reveals the following features:

- The Year 2 percent removal shows a strong increasing linear relationship with Year 1 percent removal. A sand filter that is performing well in Year 1 seems to be performing equally well or even better in Year 2. This suggests that the design modifications made at the end of Year 1 seem to be working.
- Two of the data points are well separated from the rest of the data. These two points correspond to percent TSS removal around 80 percent in both Year 1 and Year 2, while the maximum percent removal for other points is less than 50 percent. These outlying observations might well be legitimate observations representing conditions in which sand filters are especially effective, but they might also be different because they were taken under different experimental conditions not shared by the other 78 filters. One would want to address the issue of the legitimacy of these points before proceeding with a formal statistical analysis.

D2.2.3 Step 3. Prepare and Interpret Numerical Summary

In addition to the graphical summaries described above, it is useful to prepare a table of statistical measures of the data. Measures of central tendency, data variability, and relative standing provide information about key aspects of data distribution. Definitions of the various measures are provided below. JMP's standard output includes these measures. For equations, the reader is referred to the USEPA guidance document on data quality assessment (USEPA 2006a).

Measure of Central Tendency

Group	n	Mean	Median	Mode
% TSS Removal Year 1	80	11.85	8.75	n/a
% TSS Removal Year 2	80	21.91	18.30	n/a

Measure of Data Variability

Group	n	Min	Max	Variance	Std Dev	COV	IQR
% TSS Removal Year 1	80	0.20	80.70	193.36	13.91	117.38	9.63
% TSS Removal Year 2	80	1.90	91.00	195.43	13.98	63.79	10.08

Measure of Relative Standing

Group	n	5%	25%	50%	75%	95%
% TSS Removal Year 1	80	1.22	4.03	8.75	13.65	36.77
% TSS Removal Year 2	80	7.60	15.10	18.30	25.18	49.22

Figure D 2-10 shows a summary table of various statistical measures for Example D2-1.

Measure of Central Tendency

Group	n	Mean	Median	Mode
% TSS Removal Year 1	80	11.85	8.75	n/a
% TSS Removal Year 2	80	21.91	18.30	n/a

Measure of Data Variability

Group	n	Min	Max	Variance	Std Dev	COV	IQR
% TSS Removal Year 1	80	0.20	80.70	193.36	13.91	117.38	9.63
% TSS Removal Year 2	80	1.90	91.00	195.43	13.98	63.79	10.08

Measure of Relative Standing

Group	n	5%	25%	50%	75%	95%
% TSS Removal Year 1	80	1.22	4.03	8.75	13.65	36.77
% TSS Removal Year 2	80	7.60	15.10	18.30	25.18	49.22

Figure D 2-10 Numerical Summary for Example D2-1

D2.2.3.1 Measures of Central Tendency

Measures of central tendency characterize the center of a data set. The three most common measures are the mean, median, and mode. Since these and other measures are calculated using available sample data, they are referred to as sample measures.

The sample mean is the arithmetic average of the data. It is sensitive to extreme values (large or small) and presence of non-detects (see Appendix D3 for methods to deal with non-detects).

The sample median (also called the 50th percentile; see Section D2.2.3.3, Measures of Relative Standing) is the middle value in an ordered data set. Thus, half the data would be larger than the sample median and half would be smaller. The median is not influenced by extreme values and can be easily computed even if non-detects are present.

The sample mode is the value that occurs with the highest frequency. Since the sample mode may not exist or be unique, this measure is not commonly reported for quantitative data; however, it is useful for qualitative data.

If the sample mean is substantially different from the median, this would suggest a skewed distribution. For environmental data, the mean tends to be higher than the median, indicating right-skewed data.

D2.2.3.2 Measures of Data Variability

Measures of data variability provide information about the spread of values around the center of the data. Common measures include the minimum and maximum sample values (which define the range of the data), sample variance, sample standard deviation, sample coefficient of variation, and sample interquartile range.

The minimum and maximum sample values define the range of the data. The range is not very useful in drawing reliable conclusions from the sample data, because it is greatly affected by extreme values.

The sample variance is the averaged squared distance of the data points from the sample mean. A large sample variance implies that the data are not clustered close to the mean. The sample variance is affected by extreme values and by a large number of non-detects.

The sample standard deviation is the square root of the sample variance and has the same unit of measure as the data. Because it has the same units as the data, the standard deviation might be more easily interpreted and hence is more useful than variance.

The sample coefficient of variation is the relative standard deviation; that is, the sample standard deviation divided by the sample mean. It is unitless and allows the comparison of data variability across different data sets. A sample coefficient of variation of greater than 1 is generally considered to be an indication of non-normally distributed data.

The sample interquartile range is the difference between the 75th percentile and 25th percentile data values. (Recall that a percentile is the data value that is greater than or equal to a specified percentage of the data values.) The interquartile range is not much affected by extreme values. When extreme values are present, the interquartile range may be more representative of the data variability than the standard deviation.

D2.2.3.3 Measures of Relative Standing

Measures of relative standing are different percentiles of the data. Commonly reported are the 25th, 50th (also called the median), and 75th percentile values. (Recall these percentiles are displayed in a box plot). For environmental data, higher percentiles (such as 90th and 95th percentiles) may be of interest when only a small percentage of data may be allowed to exceed some standard. Also, the comparison of the maximum (minimum) data value to a high (low) percentile, such as the 95th (5th) percentile is useful in identifying a potential outlier. If the maximum (minimum) value is substantially higher (lower) than the 95th (5th) percentile, the maximum (minimum) value may be a potential outlier.

<u>Interpretation of Numerical Summary for Example D2-1</u>

The numerical summary for Example D2-1 shown in Figure D 2-10reveals the following features:

- The mean is higher than the median for both data sets, suggesting that the data distribution is non-symmetric and right-skewed. This observation confirms the conclusion that was reached previously based on a review of the graphical displays.
- The percent removal is consistently higher for Year 2 for most filters. This is reflected in higher values in Year 2 for mean, median, and most percentiles. In fact, a review of the data for all 80 filters shows that the percent removal is higher in Year 2 for all but one filter.
- The standard deviation is about the same for both years. Since the mean for Year 2 is substantially higher, the coefficient of variation (i.e., the standard deviation divided by the mean) is lower for Year 2.

D2.2.4 Step 4. Assess Whether Results of Step 2 or 3 Suggest the Presence of Potential Outliers

If the graphical and numerical methods of data review suggest that some of the high or low data points may be anomalous or inconsistent with most of the data, such points may be considered potential outliers and checked further. If no anomalous or inconsistent data values are identified, all data points should be used in the subsequent analysis.

D2.2.5 Step 5. Check Potential Outliers Identified in Step 4 for Specific Errors or for Not Being Representative

A data point should not be deleted simply based on the data review without providing specific evidence of a sampling or analytical error or a condition that would make the point non-representative of the population one is trying to study. Examples of specific errors include transcription errors, data-coding errors, sampling equipment malfunction, improper sample collection method, and laboratory errors. A data point may be considered to be non-representative of the BMP being studied if, for example, it was collected for a different BMP.

D2.2.6 Step 6. Assess Whether Any Potential Outliers Could Be Confirmed as Actual Outliers

If the checking in Step 5 reveals a specific, valid reason for concluding that a data point is an outlier, such a data point may be considered to be an actual outlier and appropriate action may be taken as discussed in the next step. If a specific, valid reason is not identified in Step 5, no data point should be removed from further analysis.

D2.2.7 Step 7. Correct or Exclude Confirmed Outliers from Further Analysis

If a specific error is identified and the error could be corrected, correct the error and include the corrected data point in further analysis. For example, if a data point is recorded in units different from other data, the data would be recorded in the correct units and used in the subsequent analysis. If a data point is confirmed as being an outlier by physical evidence, field notes, laboratory records, etc., and the analysis methods presented herein, but no correctible error is identified, one may exclude this data point from further analysis. Again, discarding an outlier from a data set should not be done without reason, particularly for environmental (such as pollutant concentration) data, which are often skewed and may naturally contain extreme values.

Appendix D3 How to Examine Data Quality in the Presence of Non-Detected Values

D3.1 Purpose and Organization

Pollutant concentration data collected for typical BMP studies often contain a mixture of data – those that are below a detection threshold and those that are above this threshold. The former values are reported as ND (for non-detect) or <MDL, which stands for method detection limit. The latter values are reported as concentrations. Data sets that contain both detect and non-detect values present special difficulties in estimating summary statistics (e.g., mean and standard deviation) and performing statistical tests of group comparison. In the statistical literature, data containing non-detects are called censored data.

This appendix (i.e., Appendix D3) describes the difficulties encountered when a data set contains non-detect values and presents methods to deal with these difficulties. An informative reference for this subject is the book Non-detects and Data Analysis. Statistics for Censored Environmental Data (Helsel 2005) that presents several methods, ranging from simple to advanced, for incorporating non-detects into the statistical analysis. Helsel's methods will be referenced throughout this appendix, which includes methods specifically related to data quality review in the presence of non-detects. The use of these methods is illustrated with examples relevant to BMP studies. Appendices D5 through D10 include methods for performing specific statistical tests in the presence of non-detects.

D3.2 Difficulties Introduced by the Presence of Non-detects

When non-detects are present in a data set, common statistical quantities (e.g., sample mean and standard deviation) cannot be computed and common statistical methods of group comparisons cannot be applied without making assumptions about the non-detects. A common method to deal with this issue is the substitution method where the non-detects are replaced with a constant value such as half the detection limit. The fundamental problem with this approach is that it assumes something is known (the values for non-detects) that is really not known. The choice for the substitution value is arbitrary, whether it is half the detection limit, equal to the detection limit, or zero, and the conclusions can be quite sensitive to that choice. Furthermore, the effect of using a particular substitution value on the conclusion is unpredictable. For example, the sample mean could overestimate or underestimate the true mean, depending on the nature of the data for the same substitution value.

D3.3 Methods of Data Quality Review in the Presence of Non-Detects

This section discusses the methods used to prepare the graphical and numerical summaries described in Appendix D2 for data containing non-detects. The data for Example D3-1, shown in Appendix D3, will be used to illustrate the methods. The data are arsenic concentrations (micrograms/liter) in the BMP effluent.

Table D 3-1 BMP Effluent Arsenic Concentrations for Example D3-1

As (µg/L)	As (µg/L)	As (µg/L)
<1.8	2.1	4.2
<1.8	2.2	4.2
<1.8	2.2	5.5
<1.8	2.4	5.8
<1.8	2.6	6.0
<1.8	3.1	6.4
<1.8	3.6	6.8
<1.8	3.6	7.1
<1.8	3.9	7.9
<1.8	4.0	8.8

D3.3.1 Recommended Graphical Methods

A histogram, box plot, and normal probability plot are the recommended graphical methods. Although the three data plots are somewhat redundant, all three should be prepared because they help reinforce the conclusions drawn from any single plot.

D3.3.1.1 Histogram

For purposes of a histogram, all non-detects are grouped in a single bin and highlighted in the histogram. For example, one could set all non-detects equal to the detection limit and define the first bin for the histogram to be 0 to the detection limit. Keep in mind that this bin may also contain any detect values that are at the detection limit. However, for purposes of understanding the key features of the data distribution, this should not be a problem.

Figure D 3-1 shows a histogram for Example D3-1. The data distribution is non-symmetric (right-skewed) and does not appear to fit a normal distribution.

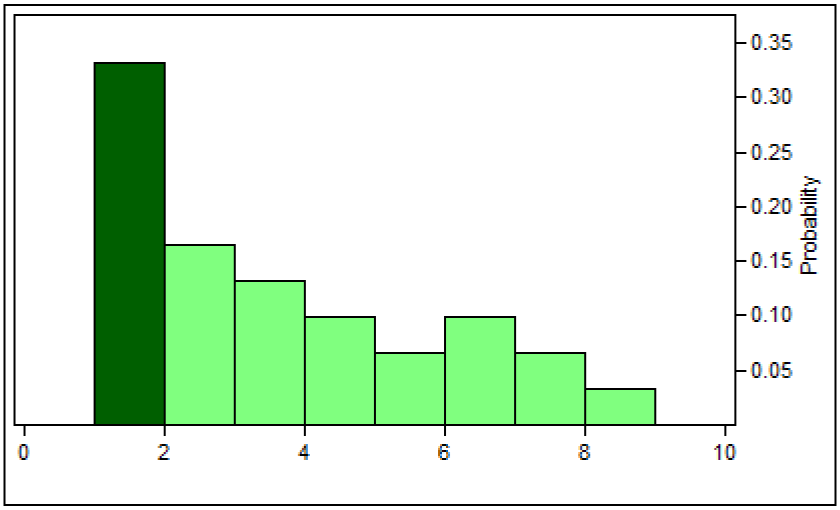


Figure D 3-1 Histogram of Arsenic Concentrations in µg/L for Example D3-1 (non-detects shown in the first bin with darker color)

D3.3.1.2 Box Plot

Similar to the histogram, data for a box plot are prepared by setting all non-detects equal to the detection limit. If the data contain *p* percent non-detects, then one cannot estimate a percentile lower than the *p*th percentile. However, percentiles higher than the *p*th percentile can be estimated without making any assumption about specific values for non-detects. In the box plot, the portion dealing with percentiles that can be estimated (i.e., percentiles higher than the *p*th percentile) is shown, while the portion dealing with percentiles lower than *p*th percentile is not shown. Such a display helps focus on the valid percentile estimates at the higher end of the data, while not displaying any information about the lower end that is not supported by the data.

Figure D 3-2 shows a box plot for Example D3-1. The plot shows that the percentage of data below approximately 2 micrograms/liter cannot be estimated with any reliability. However, the percentage of data above this level can be estimated without making any assumptions about the non-detects.

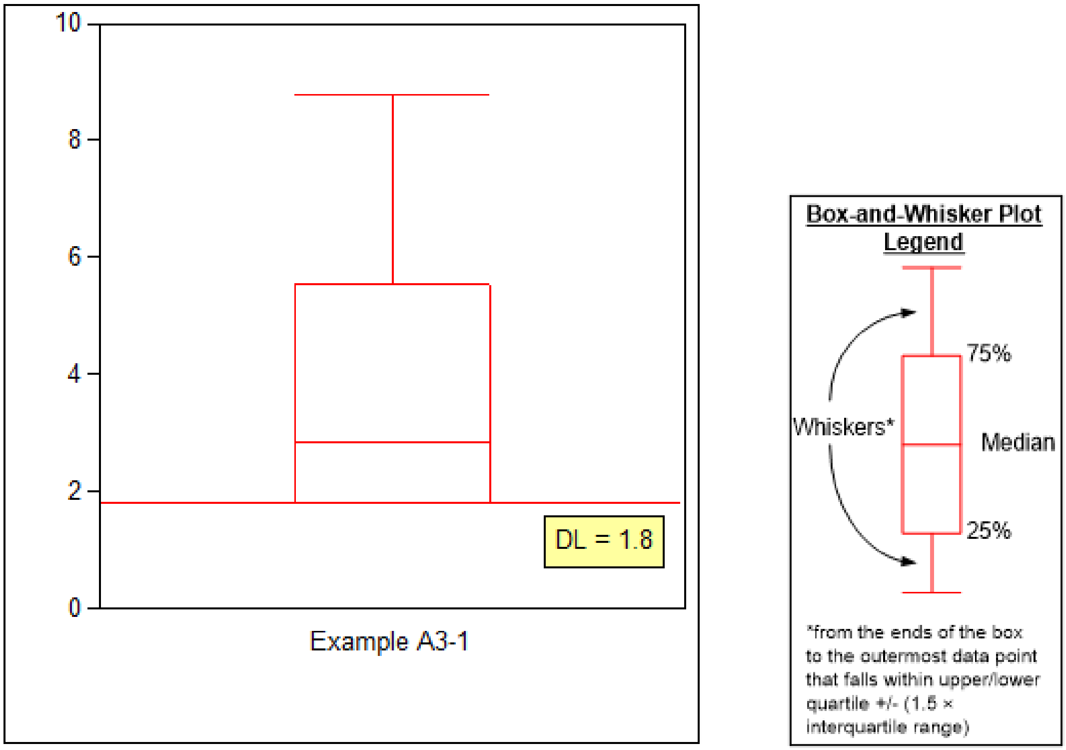


Figure D 3-2 Plot of Arsenic Concentrations in µg/L for Example D3-1

D3.3.1.3 Normal Probability Plot

As described in Appendix D2, a normal probability plot is useful to assess whether the data (original or log-transformed) follow a normal distribution. When the data contain non-detects, the preparation of such a plot presents a problem. Because the actual value of a non-detect is unknown, only its upper-bound is known, one cannot assign the position of a non-detect on the X-axis. Consequently, a point corresponding to a non-detect cannot be plotted.

In his book, Helsel describes a graphical method to prepare a normal probability plot for censored data (Helsel 2005). Figure D 3-3 shows normal probability plots for original and log-transformed data for Example D3-1 using the Helsel method. Note that only the detect values are plotted, but the plotting positions of those values depend on the percentage of non-detects. The normal probability plot for log-transformed data shows a better linear fit than the original data. Therefore, lognormal distribution is a more reasonable distributional assumption for the data.

When the data contain multiple detection limits, the probability plotting method becomes more complicated. The method is described in greater detail in Helsel's book. Consult with a statistician if assistance is needed to apply this method.

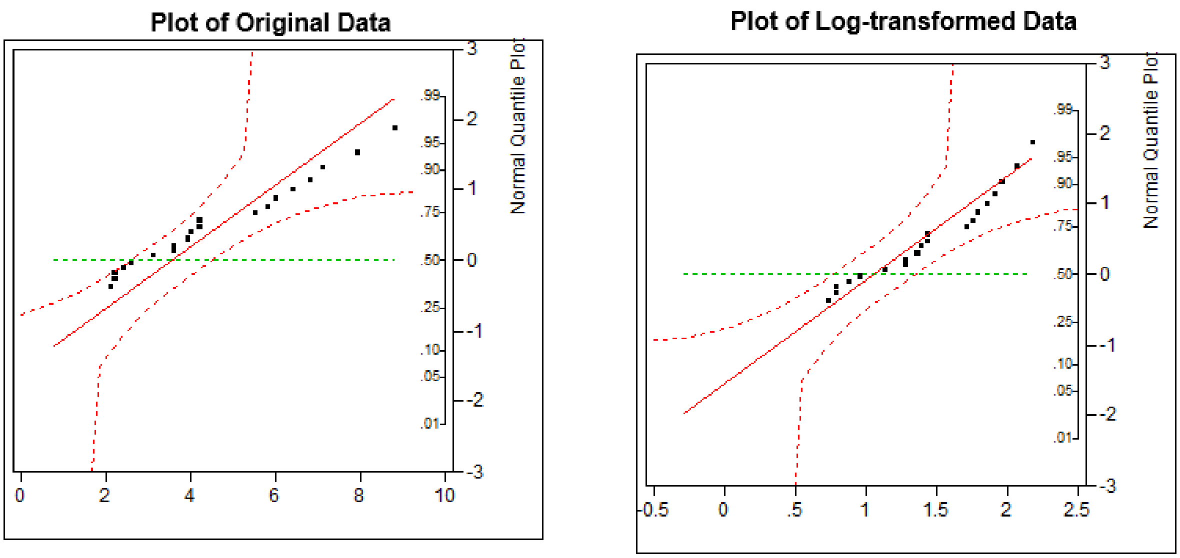


Figure D 3-3 Normal Probability Plot of Arsenic Concentrations in µg/L for Example D3-1 (only the detect values are plotted)

D3.3.2 <u>Recommended Method for Preparing Numerical Summaries</u>

The recommended method for preparing numerical summaries depends on the percentage of non-detects, the number of detects, and whether the censored data (original or log-transformed) follow a normal distribution.

If the percentage of non-detects is greater than 80 percent, no method would generate reliable numerical summaries. For this case, one should only report the percentage of data above a meaningful threshold (such as a legal standard).

If the percentage of non-detects does not exceed 80 percent, there are at least three detects, and the censored data (original or log-transformed) follow a normal distribution, the use of the robust Regression on Order Statistics (ROS) is recommended. Caltrans has developed a Data Analysis Tool (DAT) for the implementation of the ROS method. However, it does not check on the normality of the censored data. The method described in the Helsel book (pages 45-47) should be used to prepare a normal probability plot. If the plot shows that the detect values reasonably fit a straight line, the censored data may be assumed to follow a normal distribution and the robust ROS method should be applied. The numerical summaries may be obtained using the Caltrans DAT. A useful enhancement to the Caltrans DAT will be the preparation of a normal probability plot for censored data.

If the percentage of non-detects does not exceed 80 percent and there are fewer than three detects or the censored data do not follow a normal distribution, the use of the Kaplan-Meier (K-M) method is recommended. The K-M method is available in commercial statistics software packages including JMP and Minitab. However, the use of the method requires a rearrangement of the data to fit the format of the particular software package. The user should follow Helsel's instructions (page 63 of his book) to transform the data, use JMP or Minitab to generate results, and then back-transform the results to the original scale.

D3.3.3 Method to Avoid for Preparing Numerical Summaries

The USEPA guidance document (USEPA 2006a) suggests substituting non-detects with half the detection limit when the percentage of non-detects is relatively small (less than 15 percent). With this approach, the calculation of numerical summaries is no different for censored data than for non-censored data. Appendix D2 discusses the calculation of numerical summaries for non-censored data. The same methods would apply for a censored data set after the non-detects are replaced with half the detection limit.

However, Helsel performed extensive testing of several alternative methods for the calculation of summary statistics, including the substitution method. Based on this research, Helsel strongly recommends that the substitution method should be avoided in all cases, because better methods are available.

Results for Example D3-1

Figure D 3-4 shows the common statistical quantities computed for Example D3-1 using the robust ROS method. For contrast, the figure also shows the same quantities computed using the substitution method in which the non-detects are replaced with half the detection limit. As shown in the figure, the estimated mean and standard deviation, and 10th percentile and 25th percentile values for the two methods are different. Conversely, the median, 75th percentile, and 90th percentile are the same. This reflects the fact that non-detects constitute 33 percent of the data.

Summary Statistics Using the Robust ROS Method

No. of Data Points	Percent NDs	Mean	Std Dev	10%	25%	Median	75%	90%
30	33.3%	3.56	2.26	1.127	1.725	2.85	5.575	7.07

Summary Statistics Using Substitution Method

No. of Data Points	Percent NDs	Mean	Std Dev	10%	25%	Median	75%	90%
30	33.3%	3.38	2.43	0.9	0.9	2.85	5.575	7.07

Figure D 3-4 Statistical Quantities Computed for Example D3-1 Using the Robust ROS Method and the Substitution Method

D3.3.4 Other Methods for Preparing Numerical Summaries

If one is interested only in estimating certain percentiles of a dataset, a method simpler than either the robust ROS or the K-M method is available. The simpler method is to include in the numerical summaries only the percentiles that could be estimated directly from the data. Statistical quantities such as the sample mean and standard deviation would not be reported. This percentile method will not provide as much information as either the robust ROS or K-M methods. However, the results generated would not assume any information that is not present in the data, and the results would still provide useful insights into higher values in the data distribution. For example, the percentile method will be valid to use to address the following question: What is the concentration that would be exceeded no more than 10 percent of the time?

If the percentage of non-detects is *p* percent, then one only estimates percentiles higher than *p* percent. For example, if a data set contains 40 percent non-detects, one estimates percentiles higher than 40 percent. Thus, for this example, one would only show the 50th and 75th percentiles and the maximum value in the numerical summary table.

The data set for Example D3-1 contains 33 percent non-detects. Therefore, only percentiles above 33 percent could be directly estimated without making any assumption about the non-detects. For this calculation, all non-detects are set equal to a value less than the smallest detected value. The values of 50 percent, 75 percent, and 90 percent are shown in Figure D 3 4.

Appendix D4 How to Verify Common Assumptions for the Selection of an Appropriate Statistical Test

D4.1 Purpose and Organization

The purpose of this appendix is to present methods to verify common assumptions made in statistical tests so that an appropriate test can be selected. For example, a standard category of statistical tests called parametric tests assumes that the data are normally distributed. If the data strongly deviate from this assumption, a parametric test should not be used. Instead, one should use a test drawn from the alternative category of nonparametric tests.

Methods to verify the following assumptions are presented in this appendix:

- Data are independent (i.e., random) in time and space.
- Data (original or transformed) follow a normal distribution.
- Data (original or transformed) from different groups have the same variance.

Examples are presented to illustrate the application of the methods. As with other appendices, the JMP statistical software package is used to display the results of applying different methods.

D4.2 Verification of Data Independence

An important assumption in standard statistical tests is that the sample data are independent in time and space. The assumption of independence means that each data point is drawn randomly from some distribution and the data value does not depend on any other data value(s). Independent data do not show any correlation in time or space. It is important to verify that the data are independent. If the data are not independent and the statistical test assumes that they are, the confidence of reaching the correct conclusion could be substantially lower than what is implied by the results of the statistical test.

For BMP studies, the number of sampling locations is generally small and hence the sample data may not be sufficient to verify the spatial independence assumption. The monitoring stations should be selected carefully, making sure that they are sufficiently apart from each other so that they can be assumed to be spatially independent. If data were available, one could assess the correlation coefficient between the measurements at adjacent sampling locations. If the correlation coefficient is relatively small (say, less than 0.5), the data from these locations could be assumed to independent.

At each monitoring location, data would be collected for each storm event that meets certain criteria over a period of, say, 2 to 3 years. Therefore, sufficient data should be available to verify the time independence assumption. A recommended method is to prepare a time series plot that shows a graph of measurements versus time interval (e.g., the number of days from the study start date). Such a plot could be prepared using Excel graphing functions. Independent data should show a plot with no structure (i.e., a random white noise pattern). A visual examination of such a plot can reveal whether the data show any strong time dependencies. For example, whether there is a trend over time (trend analysis methods are presented in Appendix D10) or whether there is a clustering of high and low values (e.g., one year shows successive high values and the next year shows successive low values). Figure D 4-1 shows a schematic of different structures of a time series plot. Independent data should show a plot similar to the one illustrated in Figure D 4-1(a).

Data that do show time dependencies should not be combined and treated as one single data group (i.e., one statistical population). For example, assume that the effectiveness of a particular BMP improves with time. In such a case, a time series plot of the percent pollutant removal versus time would likely show an increasing trend. For statistical analysis of this data set, one may need to separately analyze the data for each monitoring year.

Example D4-1

Assume that for an upstream-downstream monitoring program, the upstream dissolved aluminum concentrations (shown in Table D 4-1) were measured over a period of 2 years. Figure D 4-2 shows a time series plot of these data using Excel. The plot shows no evidence of any particular structure or pattern, or time trend. Data values appear to be randomly distributed above and below an average value. Given these observations, the assumption of time independence for this data set would be reasonable.

Table D 4-1 Example D4-1 - Upstream Dissolved Aluminum Concentrations Measured Over a Period of 2 Years

Date	Al (mg/L)	Date	Al (mg/L)	Date	Al (mg/L)
1/5/2005	12.5	9/2/2005	10.04	5/12/2006	7.93
2/8/2005	10.34	10/10/2005	6.4	6/7/2006	12.78
3/1/2005	3.09	11/11/2005	3.82	7/3/2006	2.25
4/15/2005	13.66	12/2/2005	6.67	8/4/2006	9.73
5/10/2005	1.8	1/2/2006	15.74	9/11/2006	6.39
6/4/2005	12.64	2/12/2006	5.23	10/7/2006	14.67
7/6/2005	10.79	3/5/2006	8.36	11/15/2006	5.74
8/8/2005	5.5	4/10/2006	8.5	12/10/2006	8.25

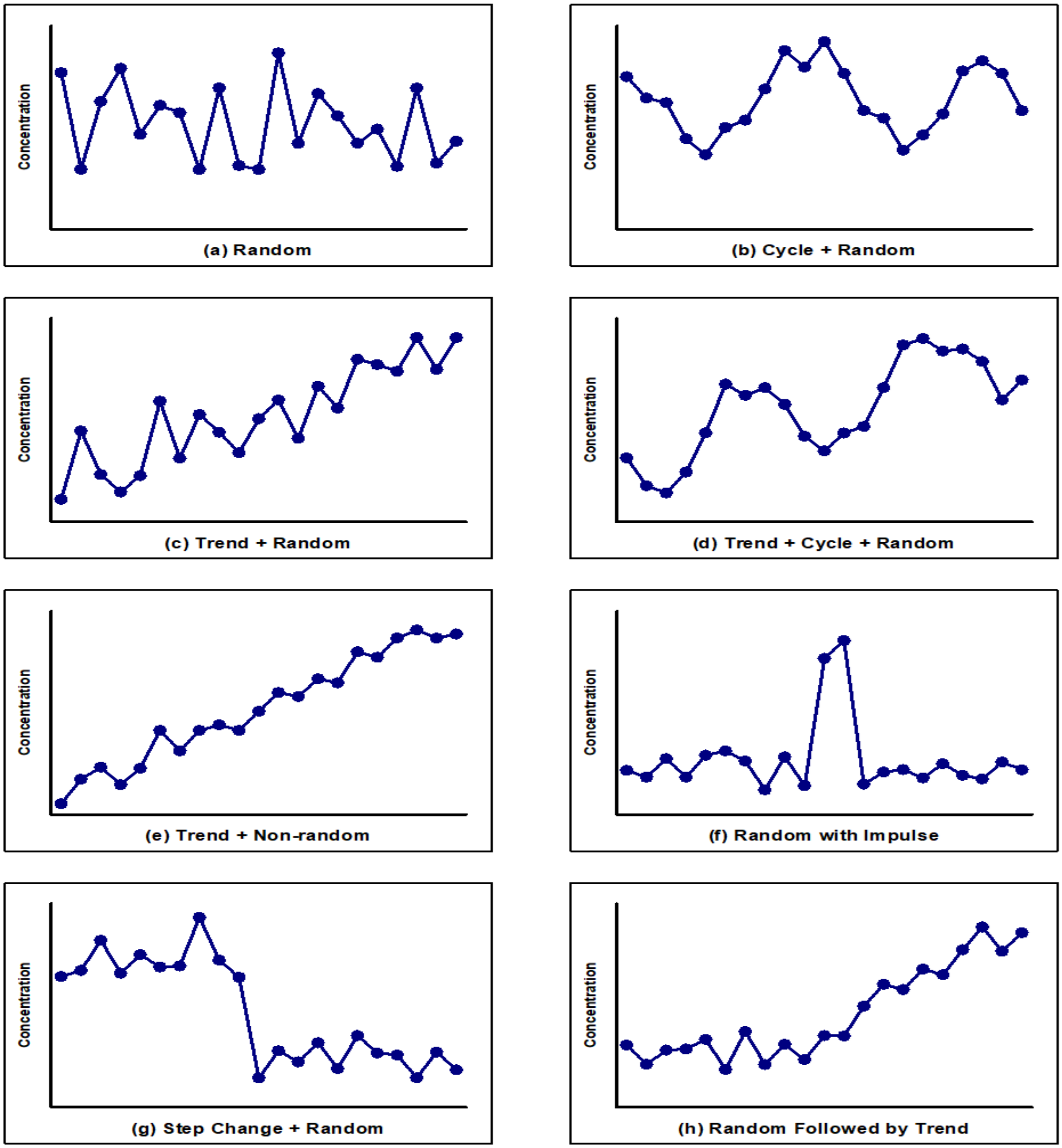


Figure D 4-1 Different Structures of Time Series Plot (Gilbert 1987)

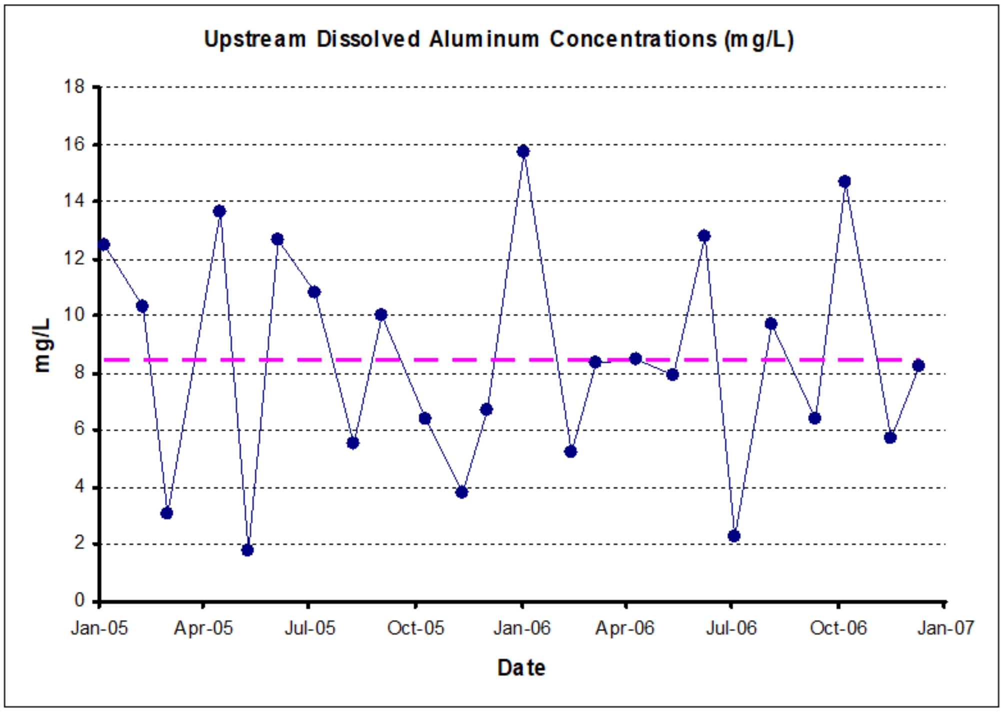


Figure D 4-2 Time Series Plot of Example D4-1

D4.3 Verification of Normal Distribution

Many standard statistical tests come from the category of parametric tests. Appendices D5 through D8 present several parametric tests that are used to answer specific study questions. A critical assumption for parametric tests is that the data follow a specific probability distribution. The most common probability distribution used in these tests is the normal distribution that takes the familiar shape of a bell-shaped symmetric distribution. The normal distribution may be assumed for either the original or transformed data. Environmental data are often nonsymmetric (typically right-skewed or long tails towards higher values). For such data, log transformation is commonly used to produce a data distribution that is symmetric and more likely to follow the normal distribution. If the log-transformed data follow the normal distribution, the original data are said to follow the lognormal distribution.

The verification of the normal distribution should begin with the use of the graphical methods described in Appendix D2. These methods include the histogram, box plot, and the probability plot. As described in Appendix D2, these plots provide a good understanding of whether the data distribution is symmetric and whether there are any potential outliers that need further investigation. If the graphical plots show that the data distribution is nonsymmetric, one should prepare the same plots using log-transformed data.

Next, preliminary impressions from a visual review of the data plots is confirmed with a formal statistical test. The Shapiro-Wilk W test is recommended in several EPA guidance documents (USEPA 2006a) and many statistical texts (Gibbons 1994). Such tests are called "Goodness-of-Fit" tests because they test how well a particular distribution fits the data. The test uses the linearity of the normal probability plot to produce a test statistic called the *W* statistic. The more linear the normal probability plot, the higher the *W* statistic, and the greater the confidence that the data distribution is normal. Statistical packages are used to determine if the probability of getting a *W* statistic in repetitive sampling is lower than that calculated. If this probability is small enough (e.g., less than 0.05), one would conclude that the assumption of a normal distribution is not reasonable. This significance probability is denoted by the letter *p*. Thus, if the *p* value for the test is less than 0.05, one may conclude that the normal distribution is not a reasonable choice for the data being analyzed. The smaller the value of *p*, the greater is the confidence that the data distribution is non-normal. On the other hand, if the *p* value is higher than 0.05, one would conclude that the assumption of a normal distribution cannot be rejected.

<u>Example D4-2</u>

The example from Appendix D2 (Example D2-1) will be used here to build on the results of the Appendix D2 graphical methods. For ease of reference, the example and the data plots (histograms, box plots, and normal probability plots) are copied below in Figure D 4-3. As discussed in Appendix D2, these plots suggest that the data distribution is non-symmetric (specifically, right-skewed) and hence, the assumption of a normal distribution may not be valid. However, the normal probability plot of log-transformed values showed a reasonable linear fit suggesting that the log-transformed values may follow a normal distribution (which also means that the original values may follow a lognormal distribution).

The Shapiro-Wilk W test was run using JMP software to confirm preliminary impressions. Figure D 4-3 shows the relevant JMP output, including the *W* statistic and the significance probability, *p* (shown under the label Prob<W). If this probability is less than 0.05, one would conclude that the assumption of normal distribution is not reasonable. For the example, the probability is less than 0.0001, which suggests that the data are highly non-normal. This finding reinforces the impressions from visual review of the data plots. Note that for Caltrans studies, the 90 percent confidence level should be used.

Since the histogram and box plot suggest that the data are right-skewed, it is appropriate to examine whether a logarithmic transformation of the data may yield a symmetric distribution that could be assumed to be normal. In Figure D 4-4, the results of the Shapiro-Wilk W test are also shown for log-transformed values. The *W* statistic now is 0.97 (in contrast to being 0.64 for the original data values; higher values suggest a better fit to a straight line and greater justification for the assumed probability distribution) and the significance level, *p*, is 0.11, suggesting that the assumption of a normal distribution for the log-transformed data cannot be rejected.

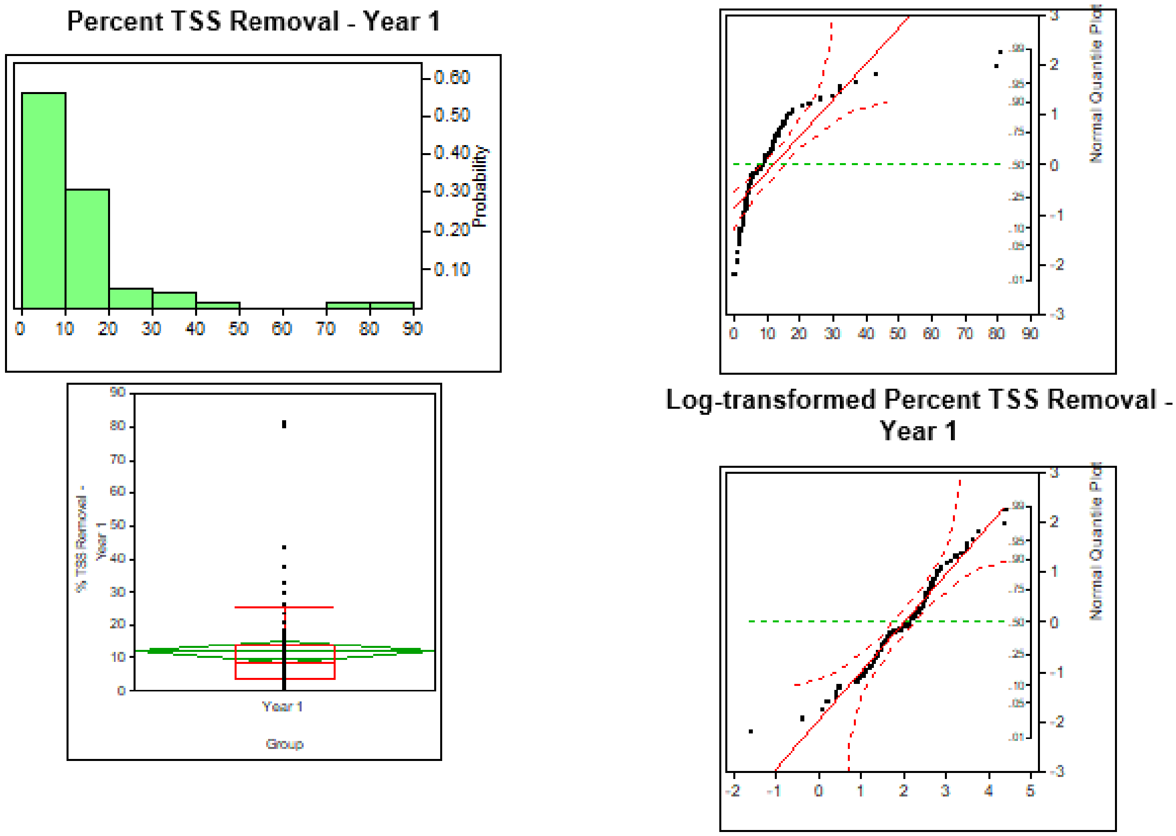


Figure D 4-3 Example D4-2 Histograms, Box Plots, and Normal Probability Plots

Percent TSS Removal - Year 1
Fitted Normal
Parameter Estimates

Type	*Parameter*	*Estimate*	*Lower 95%*	*Upper 95%*
Location	μ	11.84625	8.7517622	14.940738
Dispersion	σ	13.905367	12.034439	16.470526

Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.642457	<.0001

Log-transformed Percent TSS Removal - Year 1
Fitted Normal
Parameter Estimates

Type	*Parameter*	*Estimate*	*Lower 95%*	*Upper 95%*
Location	μ	1.9990029	1.7706293	2.2273764
Dispersion	σ	1.0262177	0.888143	1.2155268

Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.974228	0.1057

Figure D 4-4 Shapiro-Wilk W Test Results for Example D4-2

D4.4 Verification of Equal Variance Between Groups

In Appendix D2, it was noted that some of the study questions would be related to evaluating the difference among different groups. Common statistical methods for answering these questions are drawn from the category of parametric tests. As noted above, one important assumption made in parametric tests is that each data group is normally distributed. An additional assumption made in these tests is that the data variability (expressed in terms of the variance) of the different data groups is the same.

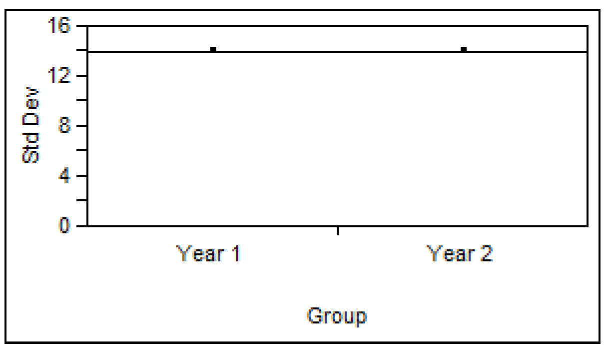
Tests for verifying this assumption fall under the category Homogeneity of Variance tests. A common test in this category is the Levene's test. The JMP software package includes five different tests for checking homogeneity of variance, one of which is the Levene's test. Although the results of the five tests often are consistent (i.e., they lead to the same conclusion), that is not always the case. To simplify the analysis, it is recommended that only the results of the Levene's test be used to draw conclusions regarding homogeneity of variance.

The JMP software summarizes the results of the Levene's test (and other tests) in terms of the significance probability, *p*. As in the case of the Shapiro-Wilk W test, this probability is compared to some threshold value (typically 0.05). If *p* is less than 0.05, the variances are considered to be unequal. If p is equal to or greater than 0.05, the assumption of equal variances cannot be rejected. If one concludes that the variances are unequal, alternative methods of statistical analysis should be used. Appendix D6 describes a method called Welch ANOVA that could be used if the data sets can be assumed to be normally distributed, but have unequal variances.

Example D4-3

Example D2-1, previously used in Appendix D2, will be used to compare the BMP performance for Years 1 and 2. The data plots and summary statistics for the two years were shown previously in Appendix D2. The JMP software package was used to apply Levene's test to check whether the variances of the data for the two years are the same. The results are shown in D 4-5. The Levene's test results are highlighted. The significance probability is listed as approximately 0.78. Based on this probability, one would conclude that the assumption of equal variances cannot be rejected.

Tests that the Variances are Equal



Level	*Count*	*Std Dev*	*MeanAbsDif to Mean*	*MeanAbsDif to Median*
Year 1	80	13.90537	8.240125	7.798750
Year 2	80	13.97970	8.735031	8.253750

Test	*F Ratio*	*DFNum*	*DFDen*	*p-Value*
O'Brien[.5]	0.0003	1	158	0.9854
Brown-Forsythe	0.0589	1	158	0.8085
Levene	**0.0807**	**1**	**158**	**0.7767**
Bartlett	0.0022	1	.	0.9623
F Test 2-sided	1.0107	79	79	0.9623

Figure D 4-5 Levene's Test Results for Example D4-3

Appendix D5 How to Estimate Probabilities Using Data for a Single Variable

D5.1 Purpose and Organization

The purpose of this appendix is to present methods to analyze data for a single variable (e.g., effluent concentration of a specified constituent of concern) to address the following types of study questions:

- How often would the effluent concentration exceed a specified legal limit?
- What is the estimated effluent concentration that would be exceeded no more than a specified percentage of time?
- What is an estimate of the mean concentration with a specified confidence level?

The methods presented are organized by assumptions that could be made about the probability distribution of the sample data for the variable of interest. Methods presented in Appendix D4 can be used to assess whether it is reasonable to assume a normal distribution for original or log-transformed data.

To address the first two study questions posed above, one needs to work with a probability distribution for the data. Two commonly used probability distributions – normal and lognormal – will be considered. The key features of each distribution are described and displayed graphically, and the use of each distribution to estimate probabilities of interest is illustrated with an example. Methods are also presented if neither normal nor lognormal distribution can be assumed for given data or the data set contains non-detects.

To address the third study question posed above, the basic concepts are presented and the use of a software program developed by the USEPA is described.

D5.2 Use of Normal Distribution to Estimate Probabilities

The graphical methods described in Appendix D2 can be used to develop preliminary conclusions regarding whether the data distribution is symmetric or skewed and whether the normal probability plot fits close to a straight line. If these preliminary conclusions suggest that the data distribution appears to be symmetric and the normal probability plot shows a reasonable fit to a straight line, one should use the formal goodness-of-fit test (i.e., the Shapiro-Wilk W test) to check whether the assumption of a normal distribution is reasonable. If the Shapiro-Wilk W test verifies the reasonableness of the normality assumption, one can use the normal distribution to assess various probabilities of interest for the variable being studied.

Key features of a normal distribution and an example to illustrate its use to address study questions of interest are provided below.

The normal distribution is a bell-shaped, symmetric distribution characterized by two parameters – the mean (μ) and variance (σ2) (or standard deviation, which is the square root of variance). These two parameters can be estimated by the sample mean (commonly denoted by) and sample variance (commonly denoted by s2). The use of the statistical software package JMP to develop numerical data summaries was described in Appendix D2. For simple statistical tasks (such as estimating sample mean and sample variance), one can also use the statistical functions built in Microsoft Excel. The Excel functions for calculating the sample mean and sample variance are AVERAGE and VAR.S, respectively. The Excel function STDEV.S calculates the sample standard deviation, which is the square root of the sample variance.

The normal distribution of a variable, X, (e.g., zinc concentration in a BMP effluent) is mathematically characterized in terms of its probability density function, f(x), which is a plot of the relative frequency against different values of the variable. This function covers the range from minus infinity to plus infinity and is fully defined by the two parameters, namely, the mean and variance. The probability density function may be thought of as a continuous curve fitted to a histogram. Figure D 5-1 shows typical normal probability density functions for two different combinations of mean and variance. The height (ordinate) of the curve, f(x) at a value x, is the probability that the variable X would be between x and x+ dx. Extending this concept to a range of x values, the probability that the variable X would be between x1 and x2 is the area under the curve between the limits of x1 to x2. Figure D 5-1 illustrates this concept by shading the area under the curve that would correspond to the probability of being between x1 and x2. The area under the entire curve, from minus infinity to plus infinity, is one.

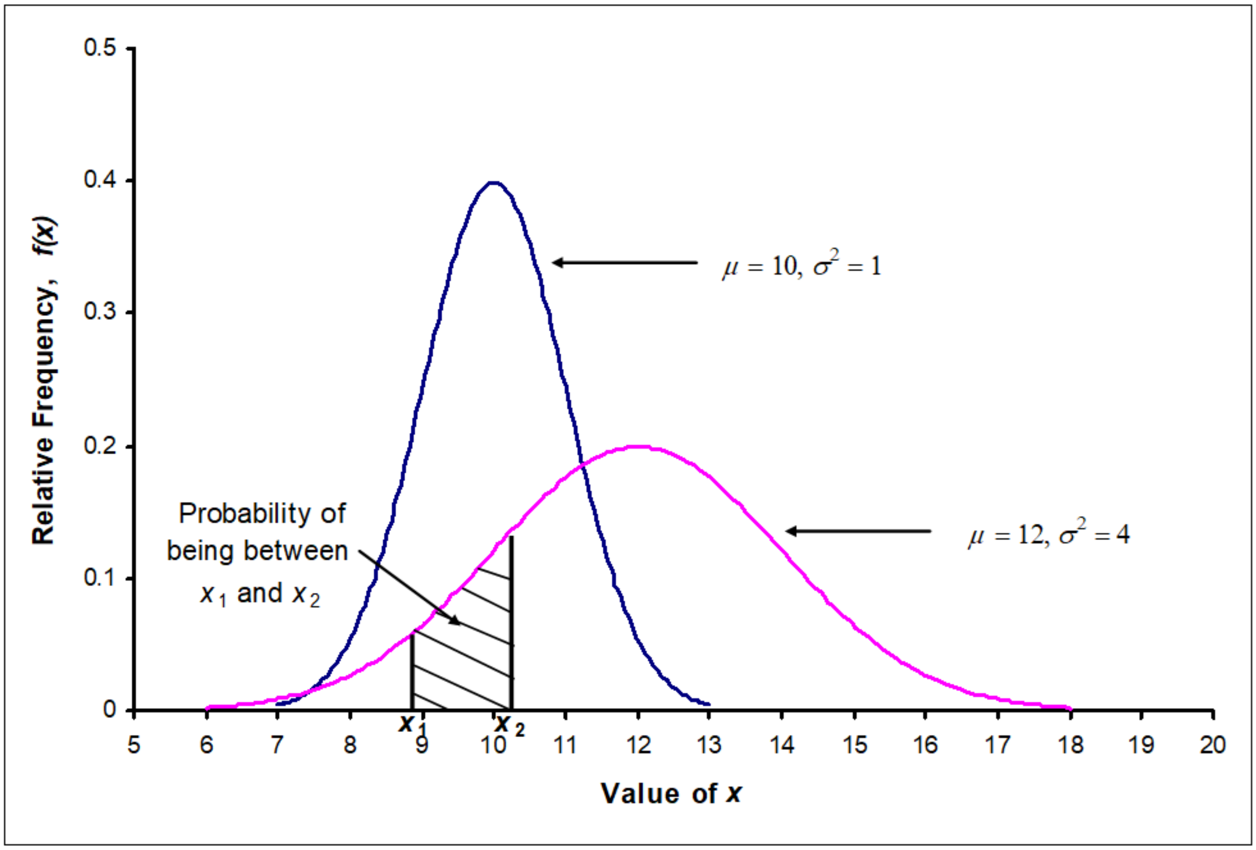


Figure D 5-1 Two Normal Distributions: N(10, 1) and N(12, 4)

The normal probability density function is centered on, and symmetric around, its mean value. It has a single mode (peak) also at the mean value, and the median (50 percent probability) is also the same as the mean.

Although the complete range of a normal distribution includes negative values, a variable that only takes on positive values can still be assumed to be normally distributed if the area (probability) of the normal probability density function below 0 is negligible.

Example D5-1

Assume that zinc concentrations were measured in the BMP effluent over a period of two years. Table D 5-1 shows the data.

One wants to: (1) assess the probability that the zinc effluent concentration exceeds 18 mg/L, (2) assess the probability that zinc effluent concentration is between 10 and 15 mg/L, and (3) find the zinc effluent concentration that would be exceeded no more than 10 percent of the time. JMP will be used to prepare certain graphical plots and Excel will be used to estimate the probabilities of interest (JMP could also be used to estimate these probabilities).

Table D 5-2 normal probability plot for the example data. These plots suggest that the data distribution is symmetric and no outliers appear to be present. The normal probability plot shows that the data points fit reasonably well on a straight line. Based on these observations, a preliminary conclusion can be drawn that the assumption of a normal distribution for these data may be reasonable and should be formally checked. The Shapiro-Wilk W test was applied using JMP and the results are shown in the figure. The significance probability, p, is 1.0, well above the threshold of 0.05, verifying the assumption of a normal distribution is not unreasonable for this example data set.

Next, one estimates the two parameters of the normal distribution, namely, the mean and standard deviation. Using Excel functions, the estimated values of the mean and standard deviation are 10.5 and 4.36, respectively.

One should then verify that the probability of getting negative values is negligible. In Excel, the statistical function NORM.DIST(x, mean, standard deviation, true) returns the (cumulative) probability of being less than or equal to x from a normal distribution with the specified mean and standard deviation. For the illustrative example, this probability is found from NORM.DIST(0, 10.5, 4.36, true). This value is 0.0080, which is negligible. Hence, the use of a normal distribution is reasonable for the example even though zinc concentration can only assume positive values.

Example D5-1, Continued

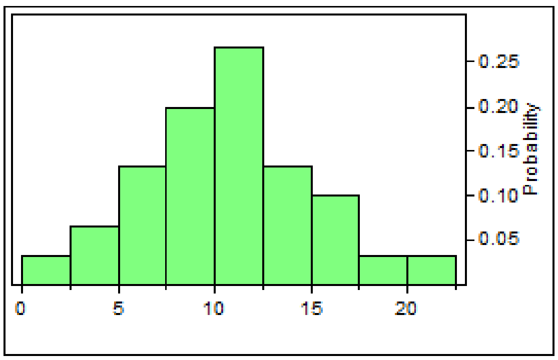
Next, one assesses the probability that zinc concentration exceeds 18 mg/L. Note that the Excel function NORM.DIST returns the (cumulative) probability of being less than or equal to a specified value. The sum of the (cumulative) probability of being less than or equal to a specified value and the probability of exceeding the same value would be one. Hence, the probability of exceeding the specified value can be calculated by subtracting the cumulative probability from 1. For this example, the probability of interest is 1 – NORM.DIST(18, 10.5, 4.36, true). This probability is found to be 0.043.

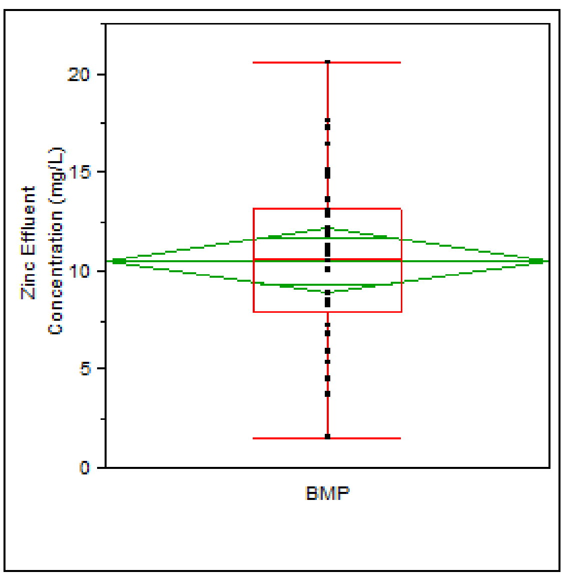
To estimate the probability that zinc concentration is between the limits of 10 and 15 mg/L, one uses the NORM.DIST function twice – once at the value of 10 and once at the value of 15. The difference between these two cumulative probabilities is the probability of being within the range of 10 to 15 mg/L. Thus, the probability of interest is [NORM.DIST(15, 10.5, 4.36, true) – NORM.DIST(10, 10.5, 4.36, true)] = 0.40.

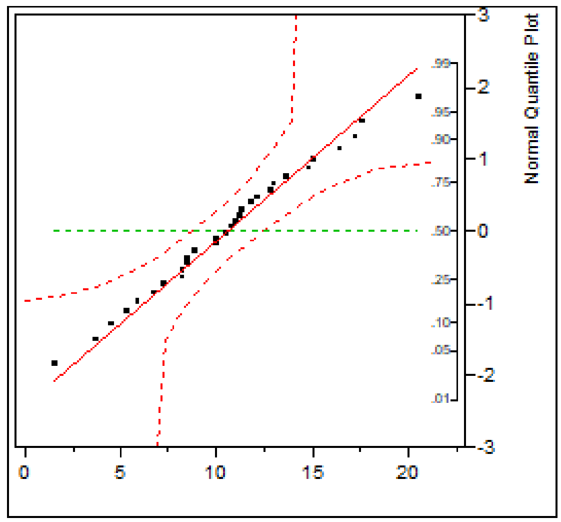
Finally, to find the zinc effluent concentration that would be exceeded no more than 10 percent of the time, one uses the Excel statistical function NORM.INV(cumulative probability, mean, standard deviation). This function is the inverse of the function NORM.DIST; it returns the value of the variable that has the specified cumulative probability given the specified mean and standard deviation of the normal variable. To find the value that would be exceeded no more than a certain proportion of the time, one must first find the corresponding cumulative probability, which would be (1 – the specified proportion of exceeding) and then use the function NORM.INV. For this example, one wants to find the zinc effluent concentration that would be exceeded no more than 10 percent of the time. The corresponding cumulative probability is (1 – 0.1) = 0.9. The Excel function NORM.INV(0.9, 10.5, 4.36) returns the value of 16.1. Thus, the zinc concentration of 16.1 mg/L would be exceeded no more than 10 percent of the time.

Table D 5-1 Example D5-1 Data, Zinc Effluent Measured over 2 Years

Zn (mg/L)	Zn (mg/L)	Zn (mg/L)
1.52	8.47	12.09
3.71	8.86	12.81
4.49	9.99	12.98
5.35	10.01	13.59
5.89	10.47	14.72
6.78	10.81	15.07
7.18	10.94	16.37
8.21	11.14	17.22
8.23	11.24	17.54
8.43	11.82	20.48







Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob< W*
0.992706	0.9987

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho

Figure D 5-2 Histogram, Box Plot, Normal Probability Plot, and Shapiro Wilk W Test Results of Example D5-1, Zinc Effluent Concentration (mg/L)

D5.3 Use of Lognormal Distribution to Estimate Probabilities

If the histogram and box plot show that the data distribution is not symmetric, but is right-skewed, this would suggest that the assumption of a normal distribution might not be reasonable. This would be further reflected in the normal probability plot that shows that the data points do not fit well on a straight line. In this case, a lognormal distribution may provide a better fit to the data. For many environmental variables, the data distribution is right-skewed, and hence the assumption of a lognormal distribution, rather than the symmetric normal distribution, might be more appropriate.

If the variable X is lognormally distributed, the variable Y, which is the natural logarithm of X (Y = ln(X)), would be normally distributed. Therefore, the methods described in the preceding section on the normal distribution can be directly used on Y. Thus, the mean (μy) and variance (σy2) of Y would fully define the normal distribution of Y, which, in turn, will fully define the lognormal distribution of X. Note that the mean of Y is also the median of Y, which can be back-transformed into the median of X (i.e., median of X = exponentiate of (mean of Y)). An alternative way of characterizing a lognormal distribution is to specify the median of X and the variance of Y.

Figure D 5 3 shows typical lognormal distributions of X for different combinations of mean and variance of Y. Note that when the variance of Y is relatively small (less than 0.25), the lognormal distribution becomes more symmetric and resembles a normal distribution.

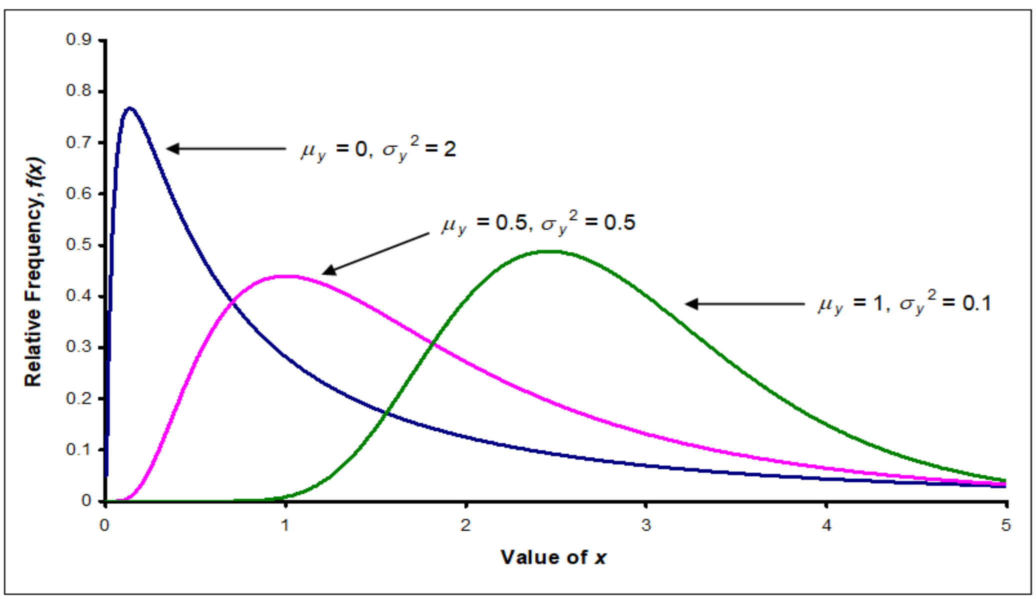


Figure D 5-3 Three Lognormal Distributions: LN(0, 2), LN(0.5, 0.5), and LN(1, 0.1)

Example D5-2

Consider the zinc effluent concentrations shown in Table D 5-2. Figure D 5-3 shows the histogram, box plot, and normal probability plot. The histogram and box plot suggest a nonsymmetric (right-skewed) distribution and the normal probability plot shows a nonlinear pattern. Based on these observations, one would conclude that the data may not be normally distributed. This is confirmed through the Shapiro-Wilk W test results also shown in Figure D 5-4. The significance probability, *p*, for the Shapiro-Wilk W test is 0.0002, well below the threshold of 0.05, which leads to the conclusion that the data do not fit a normal distribution.

Given the right-skewed data distribution, it would be appropriate to explore the possibility of fitting a lognormal distribution. Figure D 5-5 shows the histogram, box plot, normal probability plot, and results of the Shapiro-Wilk W test using the log-transformed data. All results suggest that the assumption of a lognormal distribution would be appropriate for these data. Note that, unlike a normal distribution, a lognormal distribution can only take on positive values.

The study questions are similar to those previously posed. For this example, one wants to: (1) assess the probability that the zinc effluent concentration exceeds a limit of 5 mg/L, (2) assess the probability that zinc effluent concentration is between 3 and 4 mg/L, and (3) find the zinc effluent concentration that would be exceeded no more than 10 percent of the time.

The normal distribution of log-transformed concentrations will be used to address these questions. The mean and standard deviation of *ln*(zinc concentration) are calculated to be 0.138 and 1.06, respectively. The probability that the zinc concentration exceeds the limit of 5 mg/L is the same as the probability that *ln*(zinc concentration) exceeds *ln* of 5. Thus, the probability of interest is given by (1-NORM.DIST(LN(5), 0.138, 1.06, true)), which is equal to 0.083.

Next, the probability that zinc concentration is between 3 and 4 mg/L is given by (NORM.DIST(LN(4), 0.138, 1.06, true) – NORM.DIST(LN(3),0.138, 1.06, true)), which is equal to 0.063.

Finally, the zinc concentration that would be exceeded no more than 10 percent of the time is found as EXP(NORM.INV((1-0.1), 0.138, 1.06)), which is equal to 4.46 mg/L.

Table D 5-2 Example D5-2 Data, Zinc Effluent Measured over a Period of 2 Years

Zn (mg/L)	Zn (mg/L)	Zn (mg/L)	Zn (mg/L)	Zn (mg/L)	Zn (mg/L)
1.27	3.12	1.75	0.37	1.34	1.21
1.36	4.14	0.12	0.6	1.61	1.27
5.05	0.12	7.34	4.18	0.88	0.63
0.41	3.68	1.48	0.61	0.29	2.66
1.1	1.06	0.48	1.73	0.36	4.84

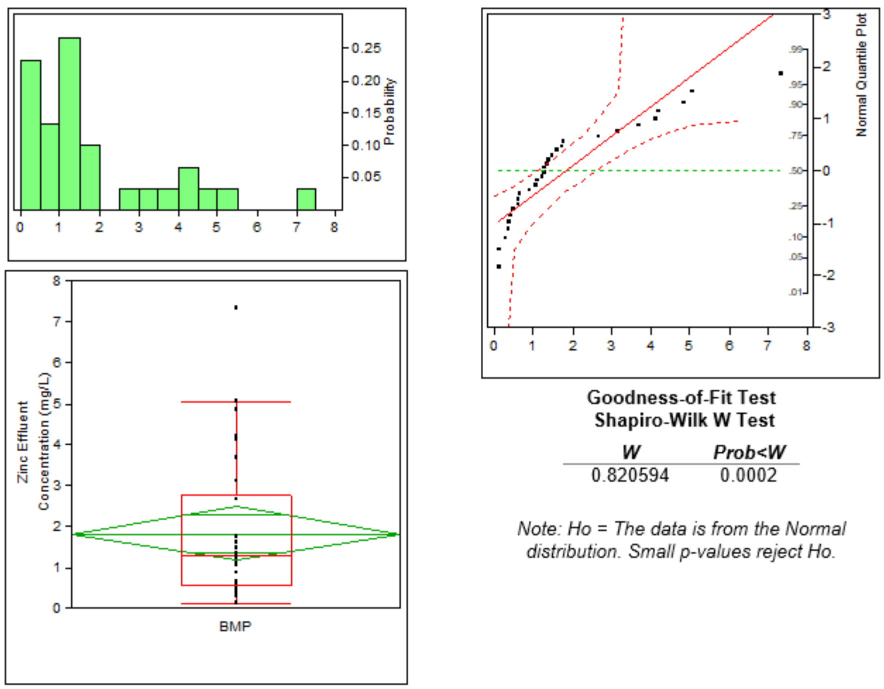


Figure D 5-4 Histogram, Box Plot, Normal Probability Plot, and Shapiro Wilk W Test Results of Example D5-2, Zinc Effluent Concentration (mg/L)

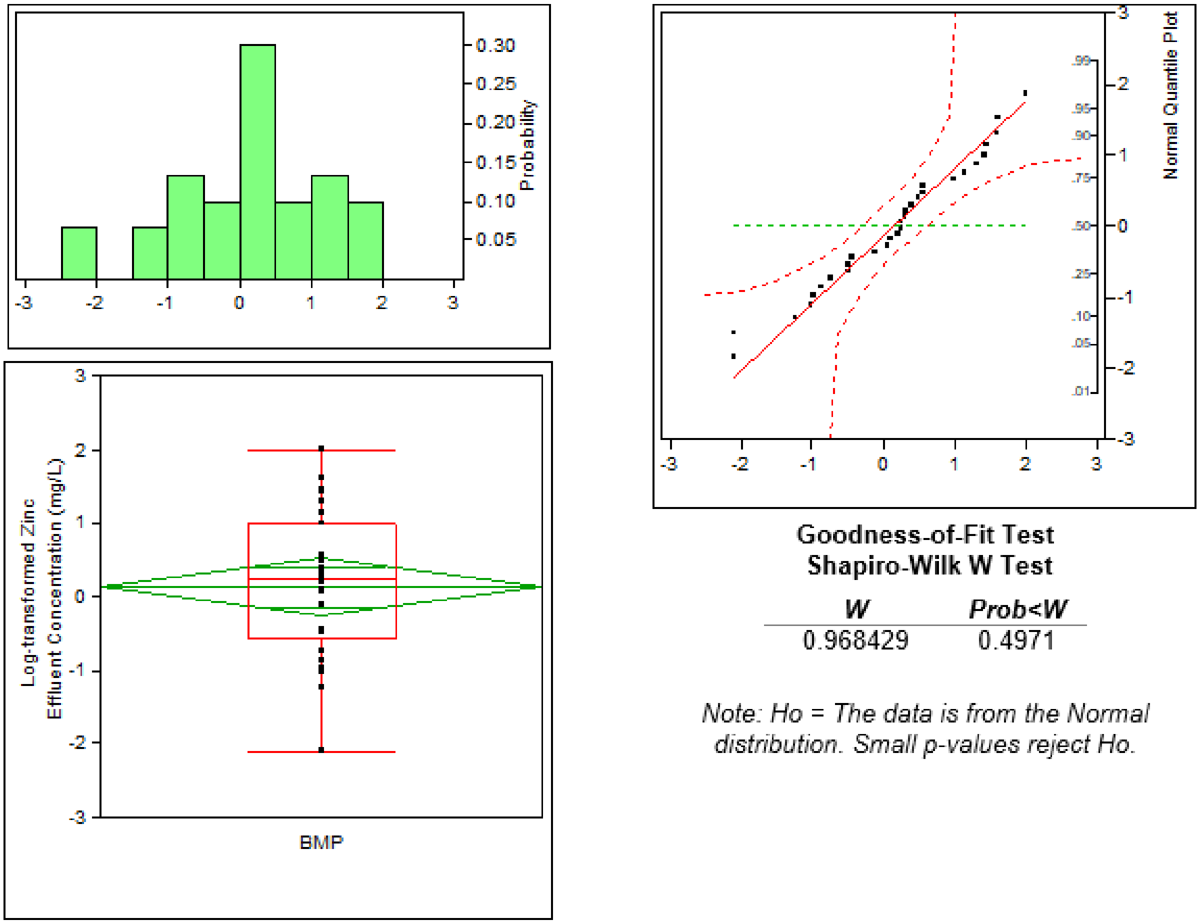


Figure D 5-5 Histogram, Box Plot, Normal Probability Plot, and Shapiro Wilk W Test Results of Example D5-2, Log-transformed Zinc Effluent (mg/L)

D5.4 If Neither Normal nor Lognormal Distribution Can Be Assumed

If neither a normal nor a lognormal distribution can be assumed for a given data set, different methods of analysis would be needed. The recommended method is described below.

The recommended method is to use empirical estimates of the proportion of data below (or above) a specified value. These estimates do not assume a specific data distribution. In Excel, the statistical functions PERCENTRANK.INC(data array, *x*) and PERCENTRANK.EXC(data array, *x*) can be used to find an empirical estimate of the proportion of data values that would be less than or equal to the specified value, *x*. This proportion then can be interpreted as the cumulative probability of being less than or equal to *x*. One limitation of this approach is that *x* cannot be greater than the maximum data value. The probability of exceeding the maximum data value is calculated to be 0.

For a small sample size, the estimated probability of exceeding a relatively high value may therefore not be reliable.

To obtain an empirical estimate of the value *x* that is exceeded no more than a specified proportion of the time, one can use the Excel function PERCENTILE.INC(data array, cumulative proportion). For example, to find the *x* value that is exceeded no more than 10 percent of the time, one would set the cumulative proportion to (1 – 0.1) = 0.9.

Example D5-3

Example D5-2 from the preceding section will be used again. However, this time, a lognormal distribution will not be assumed. Instead, the empirical estimation approach will be used.

To estimate the probability that zinc effluent concentration does not exceed the limit of 5 mg/L, one can use the Excel function PERCENTRANK.INC (data array, 5). PRECENTRANK.INC provides an empirical estimate of the cumulative probability distribution. The result is 0.957, whereas the same estimate from lognormal distribution is 0.918.

To find the probability that the zinc concentration is between 3 and 4 mg/L, one can use the Excel functions PERCENTRANK.INC (data array, 3) and PERCENTRANK.INC (data array, 4) and find the difference between the two proportions. The result is 0.067, whereas the same estimate from lognormal distribution is 0.063.

To find the zinc effluent concentration that would be exceeded no more than 10 percent of the time, one can use the Excel function PERCENTILE.INC(data array, 0.9). The result is 4.25 mg/L, whereas the same estimate from lognormal distribution is 4.46 mg/L.

D5.5 If Data Contain Non-detects

If the data set contains non-detects, the recommended methods described in Appendix D3 may be used. If the robust Regression on Order Statistics (ROS) method is used, it will generate replacement values for non-detects. These replacement values can be combined with the detect values and the combined data set can be used as if no censoring occurred. Then, the methods described in the preceding sections will be applicable to the combined data set. If the percentile method is used and the percentage of detects is greater than the percentile of interest, one can estimate this percentile without making any assumptions about the non-detects. For example, if one is interested in estimating the 80th percentile and the percentage of detects is 90 percent, one can estimate the 80th percentile directly from the data without making any assumptions about the non-detects.

D5.6 Estimation of Mean with Specified Confidence Level

For health risk assessment, an estimate of the mean with 95 percent upper confidence limit (UCL) is often required. This UCL estimate is commonly denoted as UCL95. However, for Caltrans studies, the 90 percent confidence level should be used when comparing differences in BMP treatment. The USEPA has developed a software program called ProUCL that could be downloaded from the following site: https://www.epa.gov/land-research/proucl-software. Extensive research has gone into the development of ProUCL and it is currently the best option to estimate UCL95. There is detailed documentation on the methodology used in ProUCL, which can also be downloaded from the same site.

ProUCL automatically evaluates a large number of methods and distributions to estimate UCL95 and then recommends one preferred method and the resulting UCL95. ProUCL does address the issue of non-detects.

Some of the newer methods that are evaluated in ProUCL include nonparametric bootstrap methods. Bootstrap tests belong to a group of tests called resampling tests. In resampling tests, repeated samples are drawn from the original data to create new sampling distributions for a statistic of interest. The difference between a bootstrap test and a permutation test, which is another resampling test, is that in bootstrapping samples are drawn with replacement while in a permutation test samples are drawn without replacement. Bootstrapping is a distribution free test, i.e., it does not require the distribution to be known. An advantage of bootstrap tests over other tests is that they can be used to estimate confidence intervals for a variety of different statistics such as medians, variances, percentiles, etc., and not just means. JMP provides bootstrapping in most of its statistical platforms. When selected, bootstrap reruns the entire analysis that appears in the platform report from which bootstrap is invoked.

<u>Example D5-4</u>

Example D5-2 will be used again to estimate UCL95 using the ProUCL software. The ProUCL output is displayed in Figure D 5-6. The recommended method for estimating UCL95, an Adjusted Gamma UCL, assumes a gamma distribution. The UCL95 using this method is 2.51 mg/L. In this example, the mean zinc concentration is 1.84 mg/L and the 95 percent upper confidence concentration is 2.51 mg/L. Note that for Caltrans studies, the 90 percent confidence level should be used when comparing differences in BMP treatment.

Zinc (mg/L)			
General Statistics			
Total Number of Observations	30	Number of Distinct Observations	28
		Number of Missing Observations	0
Minimum	0.12	Mean	1.835
Maximum	7.34	Median	1.27
SD	1.77	Std. Error of Mean	0.323
Coefficient of Variation	0.965	Skewness	1.515
Normal GOF Test			
Shapiro Wilk Test Statistic	0.821	**Shapiro Wilk GOF Test**	
5% Shapiro Wilk Critical Value	0.927	Data Not Normal at 5% Significance Level	
Lilliefors Test Statistic	0.253	**Lilliefors GOF Test**	
5% Lilliefors Critical Value	0.159	Data Not Normal at 5% Significance Level	
Data Not Normal at 5% Significance Level			
Assuming Normal Distribution			
95% Normal UCL		**95% UCLs (Adjusted for Skewness)**	
95% Student's-t UCL	2.385	95% Adjusted-CLT UCL (Chen-1995)	2.463
		95% Modified-t UCL (Johnson-1978)	2.399
Gamma GOF Test			
A-D Test Statistic	0.393	**Anderson-Darling Gamma GOF Test**	
5% A-D Critical Value	0.77	Detected data appear Gamma Distributed at 5% Significance Level	
K-S Test Statistic	0.132	**Kolmogorov-Smirnov Gamma GOF Test**	
5% K-S Critical Value	0.164	Detected data appear Gamma Distributed at 5% Significance Level	
Detected data appear Gamma Distributed at 5% Significance Level			
Gamma Statistics			
k hat (MLE)	1.205	k star (bias corrected MLE)	1.107
Theta hat (MLE)	1.523	Theta star (bias corrected MLE)	1.658
nu hat (MLE)	72.31	nu star (bias corrected)	66.41
MLE Mean (bias corrected)	1.835	MLE Sd (bias corrected)	1.744
		Approximate Chi Square Value (0.05)	48.66
Adjusted Level of Significance	0.041	Adjusted Chi Square Value	47.78
Assuming Gamma Distribution			
95% Approximate Gamma UCL (use when n>=50)	2.505	95% Adjusted Gamma UCL (use when n<50)	2.551
Lognormal GOF Test			
Shapiro Wilk Test Statistic	0.965	**Shapiro Wilk Lognormal GOF Test**	
5% Shapiro Wilk Critical Value	0.927	Data appear Lognormal at 5% Significance Level	
Lilliefors Test Statistic	0.103	**Lilliefors Lognormal GOF Test**	
5% Lilliefors Critical Value	0.159	Data appear Lognormal at 5% Significance Level	
Data appear Lognormal at 5% Significance Level			
Lognormal Statistics			
Minimum of Logged Data	-2.12	Mean of logged Data	0.138
Maximum of Logged Data	1.993	SD of logged Data	1.058
Assuming Lognormal Distribution			
95% H-UCL	3.319	90% Chebyshev (MVUE) UCL	3.273
95% Chebyshev (MVUE) UCL	3.87	97.5% Chebyshev (MVUE) UCL	4.698
99% Chebyshev (MVUE) UCL	6.326		
Nonparametric Distribution Free UCL Statistics			
Data appear to follow a Discernible Distribution at 5% Significance Level			
Nonparametric Distribution Free UCLs			
95% CLT UCL	2.367	95% Jackknife UCL	2.385
95% Standard Bootstrap UCL	2.361	95% Bootstrap-t UCL	2.52
95% Hall's Bootstrap UCL	2.466	95% Percentile Bootstrap UCL	2.399
95% BCA Bootstrap UCL	2.501		
90% Chebyshev(Mean, Sd) UCL	2.805	95% Chebyshev(Mean, Sd) UCL	3.244
97.5% Chebyshev(Mean, Sd) UCL	3.854	99% Chebyshev(Mean, Sd) UCL	5.052
Suggested UCL to Use			
95% Adjusted Gamma UCL	2.551		
Note: Suggestions regarding the selection of a 95% UCL are provided to help the user to select the most appropriate 95% UCL.			
Recommendations are based upon data size, data distribution, and skewness.			
These recommendations are based upon the results of the simulation studies summarized in Singh, Maichle, and Lee (2006).			
However, simulations results will not cover all Real World data sets; for additional insight the user may want to consult a statistician.			

Figure D 5-6 ProUCL (Version 5.1) Output for Data of Example D5-2

Appendix D6 How to Compare Two Independent Data Sets

D6.1 Purpose and Organization

The purpose of this appendix is to present statistical methods to compare two independent data sets and draw conclusions regarding whether observed differences in the two data sets are statistically significant. For example, BMP monitoring approaches, such as paired watershed monitoring, generate data that are representative of two independent populations of interest (e.g., water quality in control (undisturbed) and treatment (disturbed) watersheds). Similarly, studies to evaluate the effectiveness of two different BMPs generate data on two independent populations representing the effects of the two BMPs.

Typical study questions of interest in the analysis of two independent data sets are:

- In a paired watershed approach, how does one decide whether there is a significant increase in pollutant concentration in the disturbed watershed?
- How does one compare the effectiveness of two pilot BMPs at a given geographic location?

Since the statistical methods to address these and other similar study questions fall under the general category of hypothesis testing, the basic concepts of hypothesis testing will be discussed first. A decision flowchart is presented to guide the user in the selection of an appropriate statistical method, depending on the characteristics of each data set. The final sections in this appendix provide an overview of the most common methods and illustrate their application with examples related to BMP studies.

D6.2 Basic Concepts of Hypothesis Testing

The statistical methods described in this appendix (as well as in Appendices D7 through D10) belong to a general category of statistical methods called hypothesis testing. Figure D 6-1 provides a flowchart of the key steps in hypothesis testing. An overview of each step is provided below.

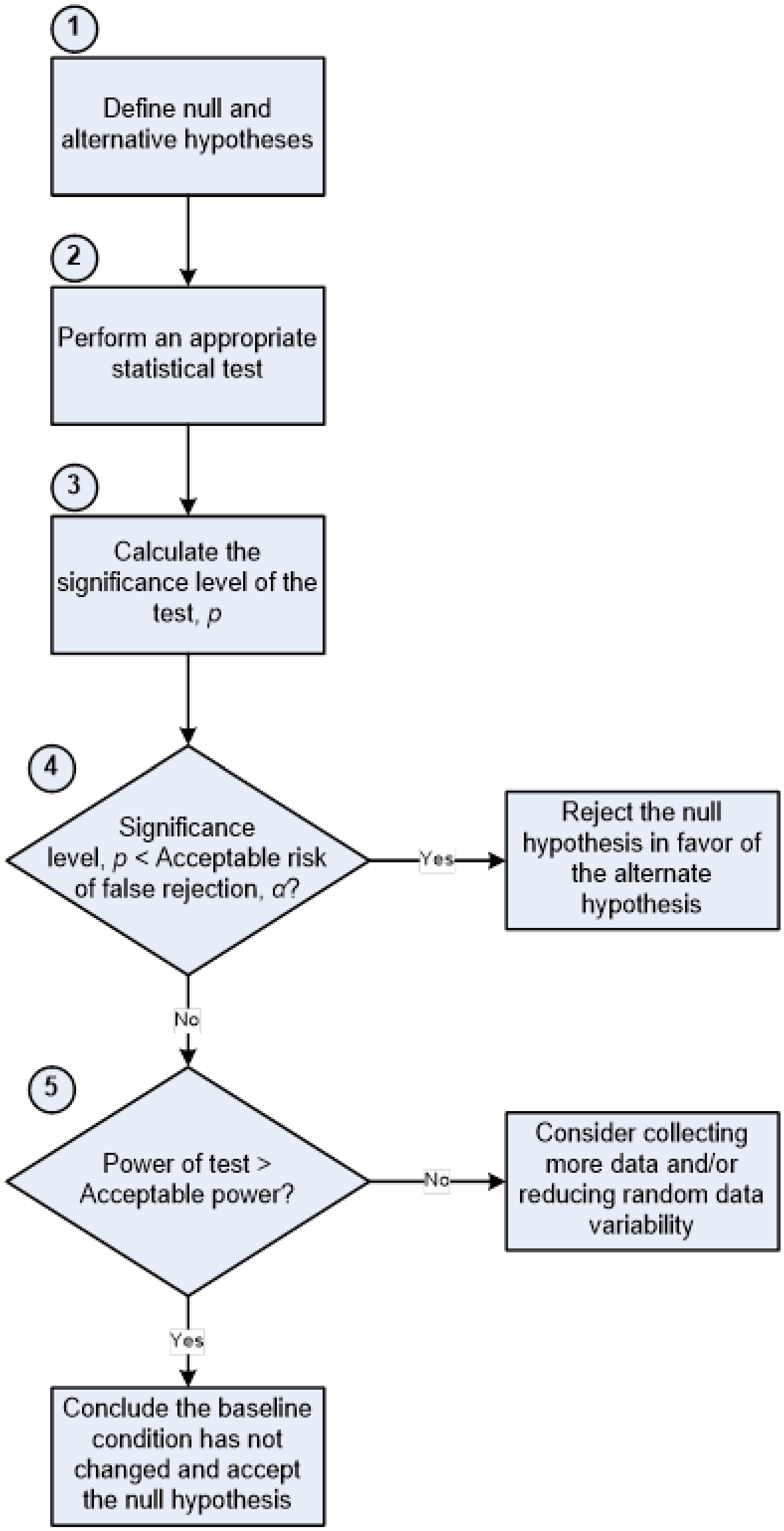


Figure D 6-1 Flowchart of Key Steps in Hypothesis Testing

D6.2.1 Define Null and Alternative Hypotheses

Hypothesis testing involves deciding whether a change from an assumed baseline condition has occurred. The baseline condition is the condition that is assumed to exist unless there is contrary evidence provided by the sample data. In the statistical literature, the baseline condition is called the *null hypothesis* and is denoted by H_0. The changed condition is called the *alternative hypothesis* and is denoted by H_A.

The following example will be used to illustrate the definition of null and alternative hypotheses. Assume that a paired watershed approach will be used to evaluate whether the mean pollutant concentration in a disturbed watershed is significantly higher than that in a matched undisturbed watershed. The assumed baseline condition will be that the average pollutant concentration in the disturbed watershed is no higher than that in the undisturbed watershed. If the sample data show strong contrary evidence, one can reject the assumed baseline condition and conclude that the average pollutant concentration in the disturbed watershed is significantly higher than that in the undisturbed watershed.

Let μ_1 and μ_2 denote the true average pollutant concentrations in the disturbed and undisturbed watersheds, respectively. Then, the null and alternative hypotheses for this example can be defined as follows:

- Null Hypothesis, H_0: $\mu_1 \leq \mu_2$
- Alternative Hypothesis, H_A: $\mu_1 > \mu_2$

D6.2.2 Step 2. Perform an Appropriate Statistical Test

A flowchart of appropriate statistical tests for the comparison of two independent populations is presented in Section D6.3.

D6.2.3 Step 3. Calculate the Significance Level, *p*, of the Test

The significance level of a statistical test, denoted by *p*, is the probability of observing a change from the baseline condition at least as extreme as that found in the sample data by chance alone when the baseline condition, in fact, has not changed. The smaller the *p*, the greater the confidence that the observed change in the sample data is “real”; that is, the observed change reflects a true change from the baseline condition and is not simply because of random chance.

When a change from the baseline condition only in one direction is of interest, the significance level is the probability of observing a change from the baseline condition in the direction of interest that is greater than that found in the sample data by chance alone. This is called a one-tailed statistical test. For example, in a paired watershed study, one is interested in evaluating whether a pollutant concentration in the disturbed watershed is higher than that in the undisturbed watershed. This evaluation should be based on a one-tailed statistical test. That is, the significance level in this case is the probability of observing an increase in the pollutant concentration greater than that found in the sample data by chance alone. Conversely, if a change from the baseline condition in either direction is of interest, the significance level is the probability of observing a change from the baseline condition that is either higher or lower than that found in the sample data by chance alone. Equivalently, it is the probability of observing a change that is greater than the absolute value of the change found in the sample data by chance alone. This is called a two-tailed statistical test. For example, in evaluating changes in pH, both high and low levels of pH would be of concern. This evaluation should be based on a two-tailed test.

D6.2.4 Step 4. Compare the Significance Level to an Acceptable Risk of False Rejection

The acceptable risk of a false rejection (also called the probability of Type I error) is denoted by α. It is the probability of incorrectly concluding that the baseline condition has changed when, in fact, it has not changed. This type of an error could occur when, by chance alone, one gets an extreme sample that is not representative of the population as a whole. Typical values of α are 0.01, 0.05, and 0.1. These correspond to confidence levels of 0.99, 0.95, and 0.9, respectively, in reaching the correct conclusion of no change from the baseline condition. Note that for Caltrans studies, the 90 percent confidence level should be used when evaluating the treatment performance of BMPs.

If the significance level, *p*, is equal to or smaller than the acceptable risk of false rejection, α, one would conclude that the risk of a false rejection is sufficiently small. Therefore, one would reject the null hypothesis (H_0) in favor of the alternative hypothesis (H_A); that is, one would conclude that the baseline condition has changed. On the other hand, if *p* is higher than α, one would conclude that the risk of a false rejection is too high. Therefore, one would not reject the null hypothesis; that is, one would conclude that the sample data do not provide strong enough evidence to conclude that the baseline condition has changed.

Continuing with the watershed example, let $\overline{X_1}$ and $\overline{X_2}$ denote the sample mean concentrations in the disturbed and undisturbed watersheds, respectively. If the water quality in the disturbed watershed has deteriorated, one would expect that $\overline{X_1}$ would be greater than $\overline{X_2}$. Let *d* denote the difference between the two means; that is, $d = \overline{X_1} - \overline{X_2}$. Positive values of the difference, *d*, would suggest that the water quality in the disturbed watershed might have deteriorated. Since *d* is estimated based on sample means, which vary from one set of samples to another, one could, by chance alone, get a relatively high value of *d* for the actual sample set that has been collected. However, the higher the value of *d*, the greater would be the confidence that the observed difference between the two sample means is real (i.e., due to disturbance) and not by chance alone.

Since one is interested only in a potential increase in pollutant concentration in the disturbed watershed, this evaluation should be based on a one-tailed test. The significance level in this case is the probability of observing an increase in the pollutant concentration that is higher than that found in the sample data. If this significance level is smaller than the specified α, one would conclude that the average pollutant concentration in the disturbed watershed is higher than that in the undisturbed watershed. Otherwise, one would conclude that the disturbance does not seem to have a significant impact on water quality.

D6.2.5 Step 5. Evaluate the Power of the Test if the Null Hypothesis is Not Rejected

If the null hypothesis is not rejected, one must be confident that, had the baseline condition changed by some magnitude of concern, the statistical test would have detected this change. This confidence is called the power of the test. Specifically, it is the probability of detecting a change, Δ, from the baseline condition. The power achieved for a given sample size is compared to the acceptable level of the power, which is denoted by (1 - β). The term β is called the probability of false acceptance, or the probability of Type II error. Typical values of the power (1 - β) are 0.8, 0.9, and 0.95. Note that for Caltrans studies, a value of 0.8 should be used.

Procedures to calculate the power of specific statistical tests are described in a subsequent section of this appendix. If the null hypothesis is not rejected (i.e., $p > \alpha$) and the power of the test to detect a specified change Δ is equal to or greater than (1 - β), one will have sufficient confidence to conclude that the baseline condition has not changed. On the other hand, if the null hypothesis is not rejected and the power of detecting a change of Δ is less than (1 - β), one only will be able to say that the baseline condition does not appear to have changed; however, one would not have sufficiently high confidence in making that statement. To increase the power of the test, one should consider collecting more data.

D6.3 Decision Flowchart

The choice of an appropriate statistical method for comparing two independent data sets depends on whether any N) values are present and whether the assumptions of normality and equal variance are satisfied. Figure D 6-2 Select an Appropriate Statistical Method for Comparing Two Independent Data Sets shows a flowchart for the selection of an appropriate method for different situations regarding these assumptions.

D6.3.1 No Non-detects in Both Data Sets

If there are no NDs in both data sets, the possibility of assuming a normal or lognormal distribution can be explored (Appendix D4). If either the original or log-transformed data fit a normal distribution, one needs to further check whether the two data sets have equal variances (Appendix D4). If both the assumptions of a normal distribution (for the original or log-transformed data) and equality of variances are satisfied, the Student's *t*-test with equal variance is used to compare the means of the two data sets. If the assumption of a normal distribution (for original or log-transformed data) is satisfied, but the assumption of equal variances is not satisfied, the Student's *t*-test with unequal variances is used. If the underlying distributions cannot be identified as normal or normalized via a transformation or if neither assumption is satisfied, a nonparametric method that does not assume any particular distribution is used. Recommended nonparametric tests in these situations are the Wilcoxon Rank Sum (WRS) test and permutation test.

The WRS test, also called the Mann-Whitney U test, is the nonparametric equivalent to the two-sample Student's *t*-test. The WRS test is a special case of permutation test based on the ranks of the data while the permutation test is based on the actual data. The effect of ranking the data is to simplify calculations and minimize the impact of the outliers, but this can be at the expense of reduced power compared to a permutation test. Note that since rank-based nonparametric tests do not use the actual data, they provide comparisons of differences in medians and not means. Permutation tests belong to a group of tests called resampling tests. In resampling tests, repeated samples are drawn from the original data to create new sampling distributions for a statistic of interest. Permutation tests are discussed further in Section D6.4.4.

D6.3.2 One or More Non-detects (but at Least One Detect) in Either Data Set

If there is only one reporting limit (RL) for non-detect data and there is at least one detect in one or both data sets, a WRS test or a Generalized Wilcoxon test is recommended. The Generalized Wilcoxon test, also called the Peto-Prentice test, is a weighted log rank test that compares the difference between cumulative distribution functions of the two data sets to determine whether their percentiles differ. The test is able to accommodate unequal sample sizes and multiple RLs. Consultation with a statistician is recommended.

One simple, although approximate, method is to set all data points below the highest RL to a common lowest value (such as half the RL or zero) and then use the procedures shown for the case of a single RL. The substitution method is not recommended for general use because substitution results in underestimated variances, and hypothesis testing of differences between groups is likely to produce lower p-values than would be the case in the absence of censoring.

D6.3.3 No Detects in Both Data Sets

If there are no detects in both data sets, no statistical analysis is possible. Professional judgment informed by qualitative factors may be used to compare the two data sets and draw conclusions.

The remainder of this appendix provides an overview of the commonly used statistical tests noted in D6.2 and examples to illustrate their use.

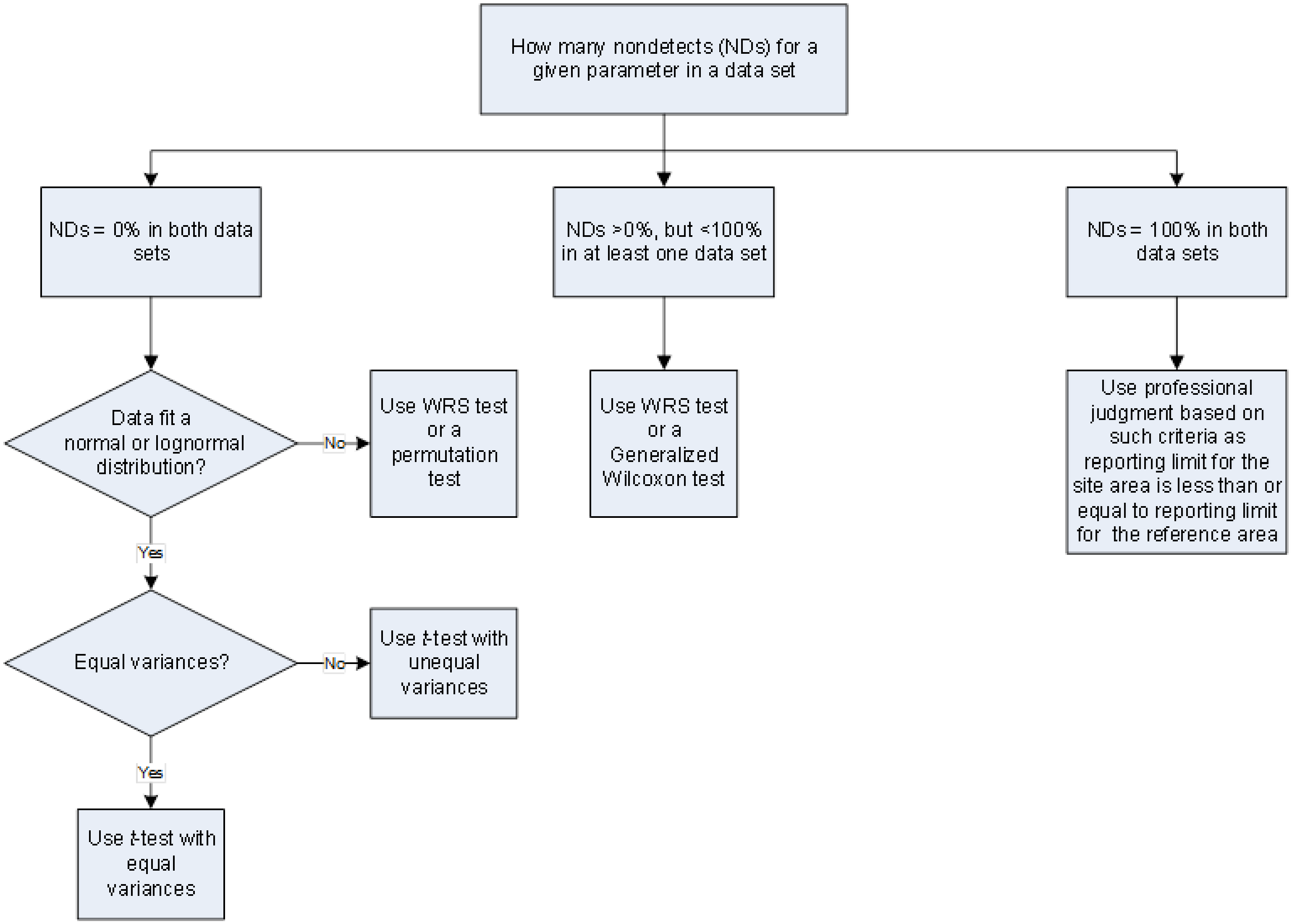


Figure D 6-2 Select an Appropriate Statistical Method for Comparing Two Independent Data Sets

D6.4 Description of Statistical Tests for Comparison of Two Independent Data Sets

The flowchart in Figure D 6-2 identifies the following statistical tests:

- Student's t-test with equal variances
- Student's t-test with unequal variances
- Wilcoxon Rank Sum (WRS) test
- Permutation test
- Generalized Wilcoxon test

The first two tests are parametric (i.e., they assume a particular probability distribution), while the remaining tests are nonparametric (i.e., they assume no specific distribution). The Student's t-test with equal and unequal variances, the WRS test, and the permutation test are discussed in the following sections.

D6.4.1 Student's t-Test with Equal Variances

This test can be used to evaluate the difference between the means of two independent data sets when the sample data contain no non-detect measurements. The key assumptions made in this test are:

Each data set follows a normal distribution
The variances of the two data sets are equal

The test is robust to moderate violations of the normality assumption, but not to large inequalities of variances. It is also not robust against outliers because sample means and standard deviations are sensitive to outliers. The normality assumption can be checked with graphical methods (histogram and normal probability plot, see Appendix D2) and a formal goodness-of-fit test (Shapiro-Wilk W test, see Appendix D4). Levene's test (Appendix D4) can be used to test the equality of variances. Methods described in Appendix D2 can be used to check whether the data contains potential outliers.

As noted previously, if the null hypothesis is not rejected, the power of the test to detect a specified change from the baseline condition should be evaluated. For the Student's t-test with equal variance, the relatively simple procedure described in USEPA Guidance Document (USEAP 2006a) is recommended. The procedure consists of finding the minimum sample size in each of the two data sets to satisfy the requirements of specified α, β, and Δ. The equation or the graphs provided in Appendix D1 can be used to find the minimum sample size. If the actual sample size in each data set is equal to or greater than the calculated minimum sample size, the test is considered to have adequate power.

Example D6-1

In a paired watershed study, lead concentration data (Table D 6-1) was collected. The study question was: Is the average lead concentration in the disturbed watershed higher than that in the matched undisturbed watershed? Hypothesis testing (Table D 6-1) will be followed to answer this question. The null and alternative hypotheses were defined as follows:

- H_0: Mean of lead concentration in the disturbed watershed, $\mu_1 \leq$ Mean of lead concentration in the undisturbed watershed, μ_2 (i.e., the disturbance has not impacted the average lead concentration
- H_A: $\mu_1 > \mu_2$ (i.e., the disturbance has increased the average lead concentration)

Application of the methods from Appendix D2 revealed no outliers. JMP software was used to perform the Shapiro-Wilk W test to verify the assumption of normality of each data set. Levene's test was used to verify the assumption of equal variances. The results of the two tests are shown in Table D 6-3. The relevant parts of the results are highlighted in the figure. The significance levels for the Shapiro-Wilk W and Levene's test are 0.9673 and 0.3397, respectively. Both are much higher than a typical threshold, α of 0.05. These results show that the assumptions of normality of each data set and equality of variances between the two data sets are reasonable. Therefore, the Student's t-test with equal variances is appropriate to use. The JMP software was used to perform this test. Figure D 6-4 shows the results.

As described in Appendix D2, the horizontal lines inside the mean diamonds provide a simple visual method to assess whether the means of the two data sets are significantly different at the 95 percent confidence level. Note that for Caltrans studies, the 90 percent confidence level should be used. Because the intervals defined by these horizontal lines in Figure D 6-4 do not overlap, the means of the two data sets are significantly different at the 95 percent confidence level. This finding is next confirmed with the results of the formal t-test. The significance level, *p*, of the test is <0.0001, as highlighted in Figure D 6-4. Because only increased lead concentrations in the disturbed watershed are of concern, this is a one-tailed test and the one-tailed *p* associated with this setup is shown under the label "Prob > t." Assuming that the acceptable risk of false rejection, α, is set to 0.05, and since $p < \alpha$, one will reject the null hypothesis and accept the alternative hypothesis that the average lead concentration is higher in the disturbed watershed than in the undisturbed watershed.

Table D 6-1 Lead Concentration Data for Example D6-1

Disturbed Watershed

Lead (mg)	Lead (mg)	Lead (mg)	Lead (mg)	Lead (mg)	Lead (mg)
4.81	13.71	14.67	18.26	21.29	22.71
10.07	14.07	16.02	19.9	21.57	23.25
11.97	14.14	17.29	20.04	21.61	25.46
12.85	14.15	17.47	20.25	22.19	26.48
13.51	14.18	18.05	21.16	22.58	29.2

Undisturbed Watershed

Lead (mg)	Lead (mg)	Lead (mg)	Lead (mg)	Lead (mg)	Lead (mg)
0.46	5.91	8.19	9.76	11.62	14.2
0.8	5.98	8.55	9.99	11.95	14.74
3.91	6.44	9.02	10.91	12.65	15.61
4.23	6.49	9.37	11.39	13.49	18.28
5.57	7.02	9.45	11.44	13.94	19.01

Distributions Group=Disturbed
Lead Concentration (mg/L)
Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.979079	0.8006

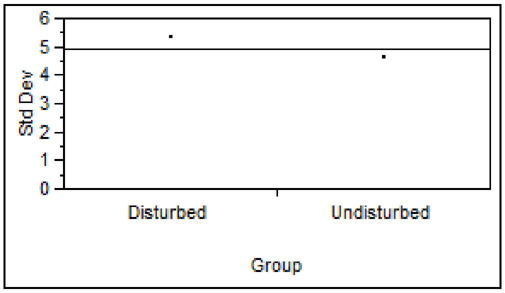
Note: Ho = The data is from the normal distribution. Small p-values reject Ho.

Distributions Group=Undisturbed
Lead Concentration (mg/L)
Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.987097	0.9673

Note: Ho = The data is from the normal distribution. Small p-values reject Ho.

Tests that the Variances are Equal

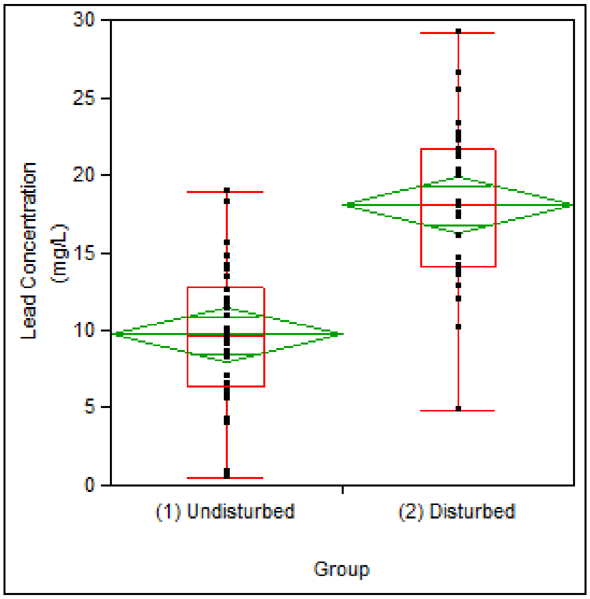


Level	*Count*	*Std Dev*	*MeanAbsDif to Mean*	*MeanAbsDif to Median*
Disturbed	30	5.299161	4.299667	4.299667
Undisturbed	30	4.562705	3.586333	3.586333

Test	*F Ratio*	*DFNum*	*DFDen*	*p-Value*
O'Brien[.5]	0.6564	1	58	0.4211
Brown-Forsythe	0.9264	1	58	0.3398
Levene	0.9269	1	58	0.3397
Bartlett	0.6359	1	.	0.4252
F Test 2-sided	1.3489	29	29	0.4252

Figure D 6-3 Shapiro-Wilk W Test and Levene's Test Results for Example D6-1

Oneway Analysis of Lead Concentration (mg/L) By Group



t-Test

(2) Disturbed-(1) Undisturbed
Assuming equal variances

Parameter	*Value*	*Parameter*	*Value*
Difference	8.418	t Ratio	6.594
Std Err Dif	1.277	DF	58
Upper CL Dif	10.974	Prob > |t|	<.0001
Lower CL Dif	5.862	Prob > t	<.0001
Confidence	0.95	Prob < t	1.0000

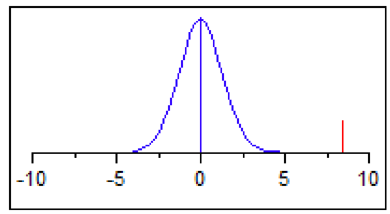


Figure D 6-4 Student's t-Test with Equal Variances Result for Example D6-1

D6.4.2 <u>Student's t-Test with Unequal Variances</u>

This test can be used to evaluate the difference between the means of two independent populations when each population is normally distributed and the variances of the two populations are unequal. As with the previous test, the sample data should contain no non-detects measurements. JMP labels the results of this test as "t Test Assuming unequal variances." Example D6-2 considers this situation. There is no simple method available to evaluate the power of the Student's t-test with unequal variances. Consultation with a statistician is recommended.

For the lead concentration data shown in Table D 6-2for undisturbed and disturbed watersheds, the application of the methods described in Appendix D2 revealed no outliers. Figure D 6-5 shows the results of applying the Shapiro-Wilk W test and the Levene's test. The relevant parts of the results are highlighted in green. Shapiro-Wilk W test shows that the significance level for this test is 0.3870 and hence each data set may be assumed to be normally distributed. The Levene's test shows that the significance level for this test is 0.0284 and hence variances of the two populations are considered to be significantly different. Therefore, the t-test assuming unequal variances should be used. The significance level, *p*, of the test is 0.7143, as highlighted in Figure D 6-6. Because only increases in the lead concentration in the disturbed watershed are of concern, this is a one-tailed test and the one-tailed *p* associated with this setup is shown under the label "Prob > t." Assuming that the acceptable risk of false rejection, α, is set to 0.05, and since $p > \alpha$, one cannot reject the null hypothesis. In other words, the disturbance does not appear to have a significant adverse impact on water quality.

Table D 6-2 Lead Concentration Data for Example D6-2

Disturbed Watershed

Lead (mg)	Lead (mg)	Lead (mg)
4.85	7.48	11.25
5	8	11.44
5.53	8.4	12.74
5.68	9.12	12.94
6.22	9.44	13.1
6.23	9.52	13.13
6.28	9.89	13.88
6.61	10.64	13.97
7.41	10.64	16.23
7.47	10.97	17

Undisturbed Watershed

Lead (mg)	Lead (mg)	Lead (mg)
1.9	7.67	12.37
2.14	7.75	12.4
2.61	8.36	13.71
3.28	9.81	13.99
3.92	10.76	15.16
5.72	10.81	15.83
6.35	11.11	15.84
6.43	11.32	18.14
6.78	11.67	21.62
7.16	12.15	24.6

Distributions Group=Disturbed
Lead Concentration (mg/L)
Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.954969	0.2292

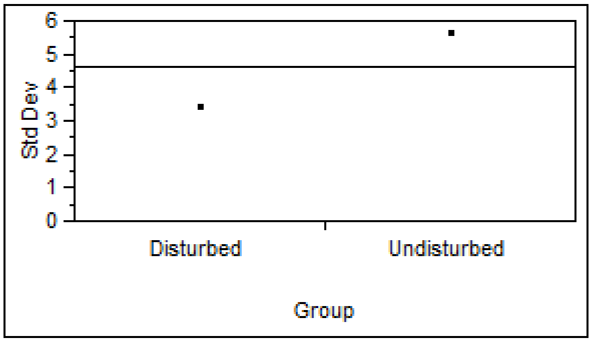
Note: Ho = The data is from the Normal distribution. Small p-values reject Ho

Distributions Group=Undisturbed
Lead Concentration (mg/L)
Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.963851	**0.3870**

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho

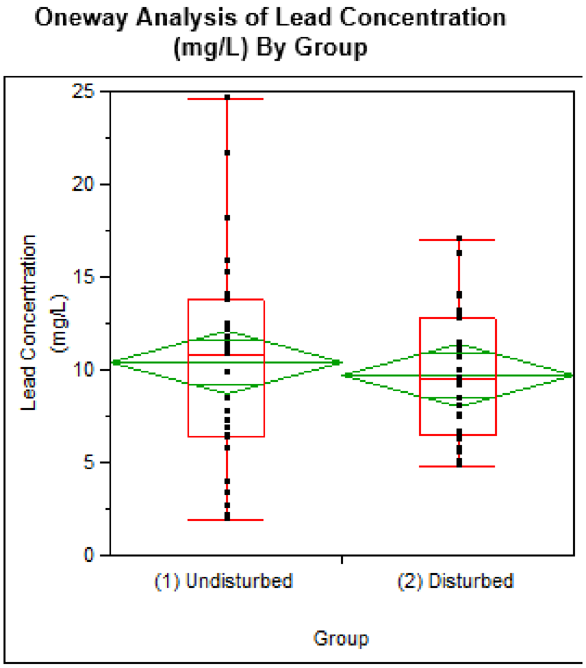
Tests that the Variances are Equal



Level	*Count*	*Std Dev*	*MeanAbsDif to Mean*	*MeanAbsDif to Median*
Disturbed	30	3.364448	2.799467	2.787333
Undisturbed	30	5.563758	4.361422	4.336000

Test	*F Ratio*	*DFNum*	*DFDen*	*p-Value*
O'Brien[.5]	5.0645	1	58	0.0282
Brown-Forsythe	4.7934	1	58	0.0326
Levene	5.0512	1	58	**0.0284**
Bartlett	6.9281	1	.	0.0085
F Test 2-sided	2.7347	29	29	0.0085

Figure D 6-5 Shapiro-Wilk W Test and Levene's Test Results for Example D6-2



t Test

(2) Disturbed-(1) Undisturbed
Assuming unequal variances

Parameter	Value	Parameter	Value
Difference	-0.677	t Ratio	-0.570
Std Err Dif	1.187	DF	47.707
Upper CL Dif	1.711	Prob > |t|	0.571
Lower CL Dif	-3.064	Prob > t	0.714
Confidence	0.95	Prob < t	0.286

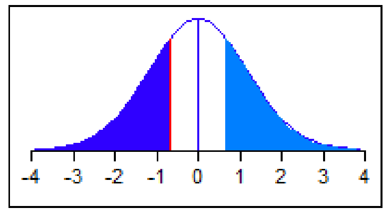


Figure D 6-6 Student's t-Test with Unequal Variances Result for Example D6-2

D6.4.3 Wilcoxon Rank Sum (WRS) Test

This test can be used to evaluate the difference between the medians of two independent populations under the following conditions:

- Neither data set contains a non-detect, and at least one data set is not normally distributed; or
- At least one data set contains one or more non-detects (but not 100 percent NDs in both data sets).

Although the normal distribution is not assumed, the WRS test does assume that the two data distributions have approximately the same shape and the only difference between them is a shift in the median. When no non-detects are present, the assumption of similar shape can be verified qualitatively by comparing the histograms or box plots of the two data sets (Appendix D2). Example D6-3 considers this situation. The WRS test is not recommended for very small sample sizes (< 5).

There is no simple method available to evaluate the power of the WRS test. The method described in the USEPA Guidance Document (USEPA 1994) could be used in consultation with a statistician. If the power of the WRS test is low, consider using a permutation test. Permutation tests are described in Section D6.4.4. In JMP, permutation tests can be conducted in the Simulate platform (available only in JMP Pro). Permutation tests are not recommended when sample sizes are less than about 20.

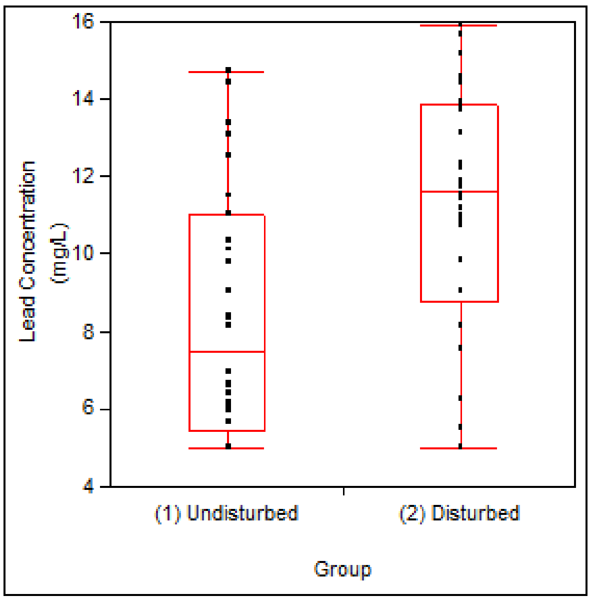
Example D6-3

Consider the lead concentration data shown in Table D 6-3 for disturbed and undisturbed watersheds. The application of the methods described in Appendix D2 revealed no outliers. Both data sets contain some non-detects. The WRS test is appropriate to use in this case to evaluate the differences between the medians of the two populations.

The JMP software provides for the application of WRS test under the "Wilcoxon Test" option in the nonparametric ANOVA menu. The results of the WRS test for this example are shown in Figure D 6-7. The significance level, *p*, is listed as "Prob>$|Z|$" under the normal approximation. Note that this is a two-tailed probability, appropriate when the alternative hypothesis is defined for the condition that the two population medians are different without regard to which median is higher or lower. For the paired watershed monitoring approach, the interest is only on the one-tailed condition; that is, when the disturbed watershed median is higher than the undisturbed watershed median. For one-tailed comparison, the significance level, *p*, is half of "Prob>|Z|" listed in the JMP results if the disturbed watershed has a higher "Score Mean." Conversely, if undisturbed watershed has a higher "Score Mean", the one-tailed significance level, *p*, is (1 - half of "Prob>|Z|"). For this example, the one-tailed *p*-value is half of 0.0031 (since disturbed watershed has a higher "Score Mean"), and equal to 0.0016. Assuming that the acceptable risk of false rejection, α, is set to 0.05, and since $p < \alpha$, one will reject the null hypothesis and accept the alternative hypothesis. In other words, the disturbance appears to have a significant adverse impact on water quality.

Table D 6-3 Lead Concentration Data for Example D6-3

Disturbed Watershed					Undisturbed Watershed				
Lead (mg/L)	Lead (mg/L)	Lead (mg/L)	Lead (mg/L)	Lead (mg/L)	Lead (mg/L)	Lead (mg/L)	Lead (mg/L)	Lead (mg/L)	Lead (mg/L)
<5	8.12	11.17	12.32	14.41	<5	5.01	6.59	9.02	11.48
<5	9.04	11.41	13.11	14.54	<5	5.61	6.65	9.78	12.52
<5	9.8	11.52	13.12	15.15	<5	5.97	6.92	10.11	13.06
5.5	10.73	11.75	13.68	15.62	<5	6.01	8.12	10.32	13.38
6.25	10.83	11.86	13.8	15.68	<5	6.18	8.3	11.02	14.43
7.56	10.94	12.19	13.89	15.92	<5	6.42	8.39	11.03	14.71



2-Sample Test, Normal Approximation

S	*Z*	*Prob>|Z|*
1115.5	2.96187	0.0031

1-way Test, ChiSquare Approximation

ChiSquare	*DF*	*Prob>ChiSq*
8.8166	1	0.0030

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	*n*	*Score Sum*	*Score Mean*	*(Mean-Mean0)/Std0*
(1) Undis-turbed	30	714.5	23.817	-2.962
(2) Dis-turbed	30	1115.5	37.183	2.962

Figure D 6-7 WRS Test Result for Example D6-3

D6.4.4 Permutation Test

The permutation test belongs to a group of tests called resampling tests. In resampling tests, repeated samples are drawn from the original data to create new sampling distributions for a statistic of interest such as the mean or median. The difference between a permutation test and a bootstrap test (another resampling test) is that in a permutation test samples are drawn without replacement while in bootstrapping samples are drawn with replacement. Permutation tests are described in detail in USGS (2020).

An advantage of permutation tests over parametric tests such as the *t*-test is that they do not require that data be normal. However, permutation tests should be used with caution when variances for the data sets are not equal because of loss of power. Unlike rank-based tests like the WRS test described above, permutation tests can be used to test for differences in means, as well as a variety of different statistics such as medians, variances, and percentiles.

The key assumptions made in this test are:

- Sample sizes are sufficiently large. Permutation tests are not recommended when sample sizes are less than about 20.
- The variances of the two data sets are equal.

Levene's test (Appendix D4) evaluates the equality of variances. Methods described in Appendix D2 can be used to check whether the data contain potential outliers.

To perform a permutation test, the observed data from all groups are randomly assigned to the individual groups being compared according to their sample sizes. A statistic of interest (e.g., the difference in group means or medians) is calculated after each rearrangement, which represents the distribution of differences to be expected when the null hypothesis is true. The process is repeated multiple times (e.g., 10,000) to generate a distribution of observed values for the statistic of interest. The p-value is then the proportion of outcomes that are equal to, or more extreme than, the one in the observed data.

<u>Example D6-4</u>

Consider the paired watershed monitoring study described in Example D6-1. The lead concentration data are shown again in Table D 6-4. The study question is: Is the average lead concentration in the disturbed watershed higher than that in the matched undisturbed watershed? The hypothesis testing steps (Figure D 6-1) will be followed to answer this question.

The application of the methods from Appendix D2 revealed no outliers and the assumption of equality of variances between the two data sets is found to be reasonable (see Example D6-1). JMP software was used to perform the premutation test. In JMP, permutation tests are conducted in the Simulate platform (available only in JMP Pro). The initial means for the data sets are first calculated and Simulate is then run by right clicking on the appropriate statistic of interest, such as the group mean. The simulated means are analyzed to the determine the p value.

Figure D 6-8 shows the output from JMP for the permutation test. The output gives the histogram estimating the permutation distribution for the difference in means between the two watersheds and the associated p-value using 10,000 random permutations. The p-value is the proportion of the 10,000 resamples that yield values for differences in means that are at least as extreme as the observed difference of 8.418. For this example, the one-tailed *p*-value is smaller than 0.0001. Assuming that the acceptable risk of false rejection, α, is set to 0.05, and since $p < \alpha$, one will reject the null hypothesis and accept the alternative hypothesis. In other words, as was concluded using the *t*-test, the disturbance appears to have a significant adverse impact on water quality.

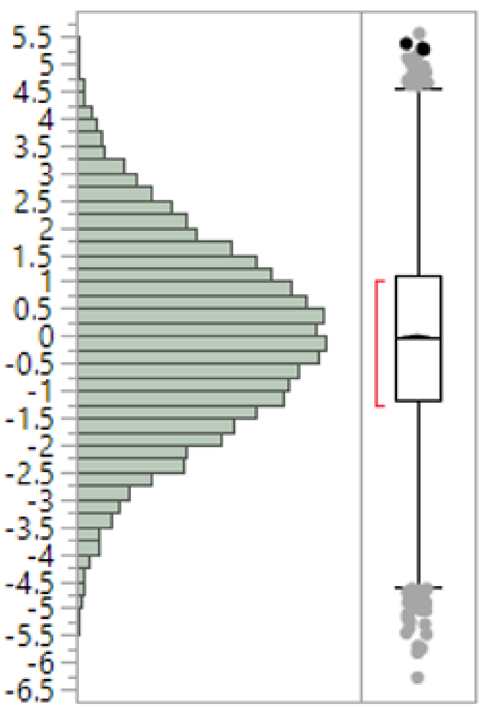
Table D 6-4 Lead Concentration Data for Example D6-4

Disturbed Watershed

Lead (mg)	Lead (mg)	Lead (mg)
4.81	14.67	21.29
10.07	16.02	21.57
11.97	17.29	21.61
12.85	17.47	22.19
13.51	18.05	22.58
13.71	18.26	22.71
14.07	19.9	23.25
14.14	20.04	25.46
14.15	20.25	26.48
14.18	21.16	29.2

Undisturbed Watershed

Lead (mg)	Lead (mg)	Lead (mg)
0.46	8.19	11.62
0.8	8.55	11.95
3.91	9.02	12.65
4.23	9.37	13.49
5.57	9.45	13.94
5.91	9.76	14.2
5.98	9.99	14.74
6.44	10.91	15.61
6.49	11.39	18.28
7.02	11.44	19.01



Summary Statistics

Statistics	Value
Mean	-0.020569
Std Dev	1.688855
Std Err Mean	0.0168886
Upper 95% Mean	0.0125361
Lower 95% Mean	-0.053674
N	10000

Simulation Results

Yo = 8.418 (original Estimate)

Confidence Levels and Empirical p Values

Alpha	Lower CI	Upper CI	Test	p-Value
0.05	-3.35	3.288	$Y \geq \lvert Y_0 \rvert$	<.0001*
0.10	-2.761	2.777	$Y \leq Y_0$	1.0000
0.20	-2.191	2.163	$Y \geq Y_0$	<.0001*
0.50	-1.171	1.127		

Figure D 6-8 Permutation Test Result for Example D6-4

Appendix D7 How to Compare Two Paired Data Sets

D7.1 Purpose and Organization

The purpose of this appendix is to present statistical methods to compare two paired data sets and draw conclusions regarding whether observed differences in the two data sets are statistically significant. In BMP studies, paired data sets arise when monitoring approaches such as influent-effluent monitoring and before-after monitoring are used. For example, in an influent-effluent monitoring approach, BMP structures may be placed at several locations. Water samples are collected from the flow entering into, and exiting from, each BMP structure. Such paired sampling designs are useful to block out sources of background noise and focus on the effect being studied. For example, in influent-effluent monitoring, the percent reduction in the effluent concentration for a given storm event would be mainly due to the BMP effectiveness and may not be affected by other factors such as highway traffic and prior maintenance at the time of the storm event. Consequently, for paired sampling designs, the data variability is reduced and a greater power of detecting differences is achieved.

Typical study questions of interest in the analysis of two paired data sets are:

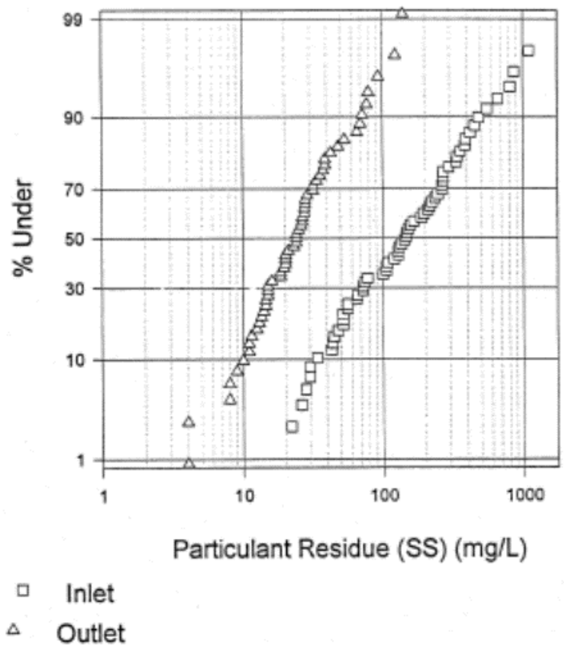
- Is the BMP effective in reducing pollutant concentration?
- Are average pollutant concentrations measured before and after a particular maintenance practice different?

In this appendix, a decision flowchart is presented to guide the user in the selection of an appropriate statistical method depending on the characteristics of each data set. An overview of each method is presented and examples related to BMP studies are provided to illustrate the use of each method.

Upon completion of the comparison tests described in this appendix, the Effluent Probability Method should be employed. A discussion of this method (taken from the ASCE Urban Stormwater BMP Performance Monitoring, Section 2.9.3 [ASCE 2002]) is described herein. This can be a useful approach to quantifying BMP efficiency by examining either a cumulative distribution function of influent and effluent quality or a standard parallel probability plot. Before any efficiency plots are generated, appropriate parametric (or if applicable nonparametric) statistical tests should be conducted to indicate if any perceived differences in influent and effluent mean (or if applicable median) event mean concentrations are statistically significant (the level of significance should be provided, instead of just noting if the result was significant, assume a 90% confidence level).

The Effluent Probability Method is a straightforward method for assessing the ultimate measure of BMP effectiveness, effluent water quality. The most useful approach for examining these curves is to plot the results on a standard parallel probability plot. A normal probability plot should be generated showing the log transform of both inflow and outflow EMCs for all storms for the BMP. If the log-transformed data deviates significantly from normality, other transformations can be explored to determine if a better distributional fit exists. Figure D 7-1 shows two types of results that can be observed when plotting pollutant reduction observations on probability plots. The plot for particulant residue shows that suspended solids (SS) are generally removed over influent concentrations ranging from 20 to over 1,000 mg/L. A simple calculation of "percent removal" would not show this removal over the full range of observations. In contrast, the filtered residue plot generally shows poor removal of total dissolved solids TDS. The "percent removal" for TDS would be close to zero and no additional surprises are indicated on this plot. Note that effluent probability plots ignore pairing of influent and effluent data, so it is possible that effluent concentrations may be higher than corresponding influent concentrations for some storm events. A closer examination of the data is required to draw firm conclusions on the consistency of treatment.

Water quality observations do not generally form a straight line on normal probability paper, but do (at least from about the 10th to 90th percentile level) on log-normal probability plots. This indicates that the samples generally have a log-normal distribution as described previously in this document and many parametric statistical tests can often be used (e.g., analysis of variance), but only after the data are log-transformed. These plots indicate the central tendency (median) of the data, along with their possible distribution type and variance (the steeper the plot, the smaller the COV and the flatter the slope of the plot, the larger the COV for the data). Multiple data sets can also be plotted on the same plot (such as for different sites, different seasons, different habitats, etc.) to indicate obvious similarities (or differences) in the data sets. Most statistical methods used to compare different data sets require that the sets have the same variances, and many require normal distributions. Similar variances are indicated by generally parallel plots of the data on the probability paper, while normal distributions would be reflected by the data plotted in a straight line on normal probability paper.



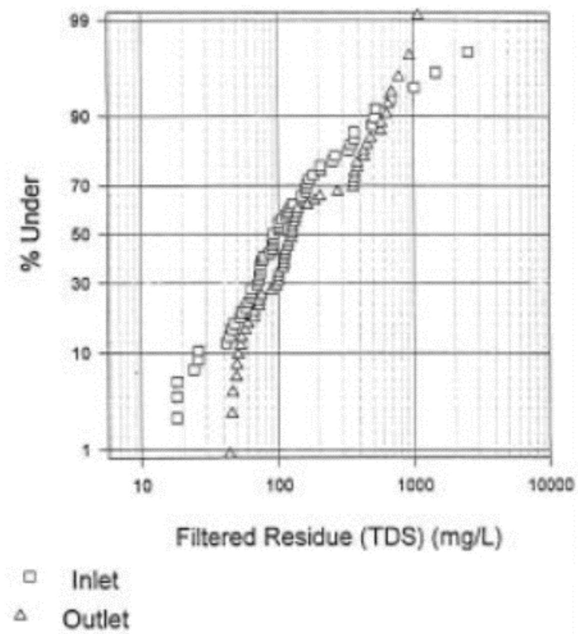


Figure D 7-1 Sample Probability Plots for Suspended Solids (SS) and Total Dissolved Solids (TDS)

D7.2 Decision Flowchart

Data for two paired data sets can be reduced to a single data set by taking differences between each pair of observations. The choice of an appropriate statistical test then depends on whether non-detects are present and the distribution of the differences. Figure D 7-2 shows a flowchart to select an appropriate statistical test for different situations. The main pathways of the flowchart are discussed below.

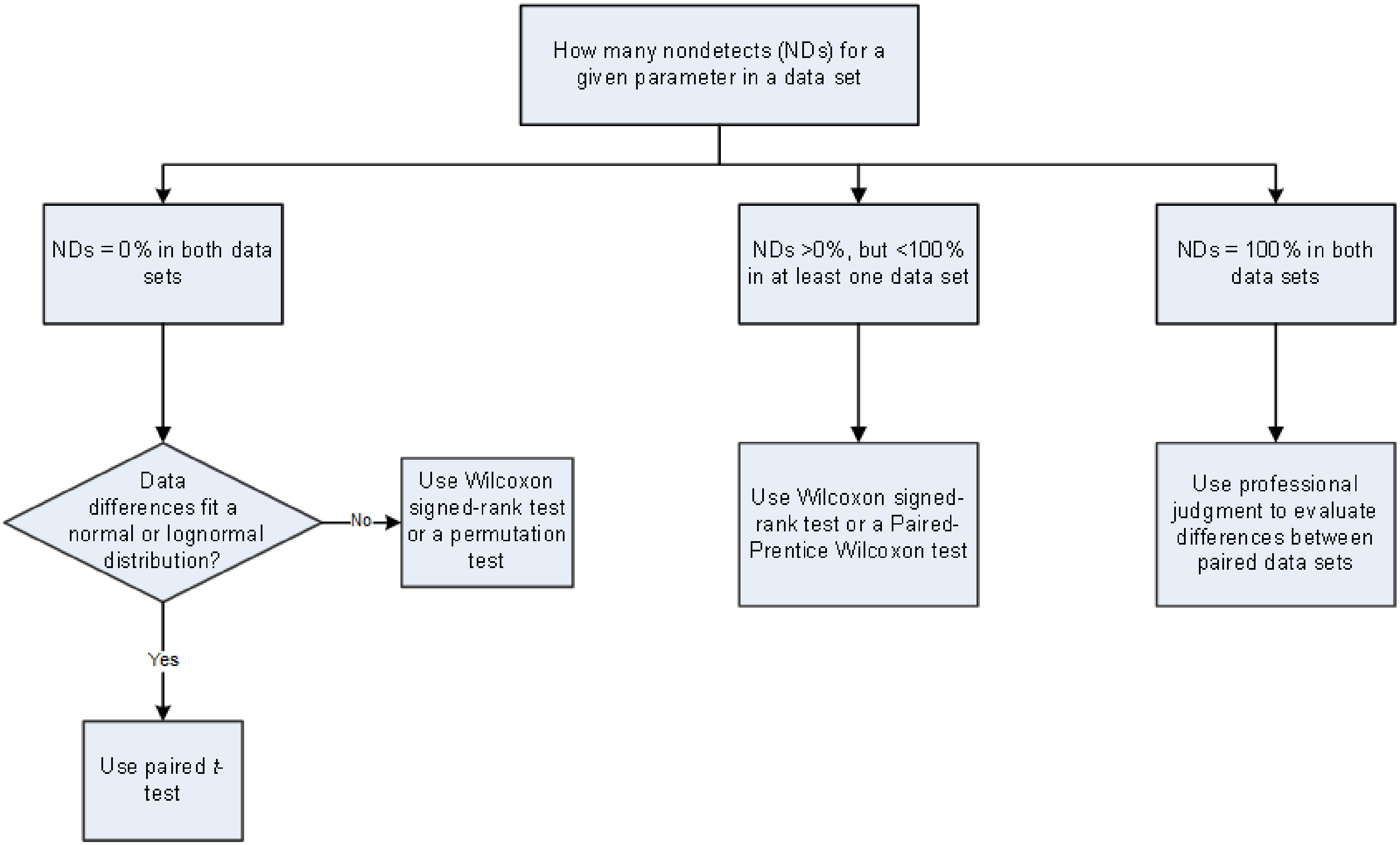


Figure D 7-2 Select an Appropriate Statistical Method for Comparing Two Paired Data Sets

D7.3 Neither of the Two Paired Data Sets Contains Any Non-detects

If the original or log-transformed differences can be assumed to follow a normal distribution, the paired t-test can be used on the differences between the two paired data sets. The assumption of a normal distribution can be verified through graphical methods described in Appendix D2 (histogram and normal probability plot) and the formal Shapiro-Wilk W test described in Appendix D4. If neither the original nor the log-transformed data can be assumed to follow a normal distribution, a nonparametric method that does not assume a specific distribution should be used. One such nonparametric test is called the Wilcoxon signed-rank test. This test assumes that the distribution of the differences is symmetric, but not necessarily normal. The assumption of a symmetric distribution can be verified through a histogram. If the power of the Wilcoxon signed-rank test is low, consider using a permutation test. Permutation tests are described in Section D6.4.4. In JMP, permutation tests can be conducted in the Simulate platform (available only in JMP Pro). Permutation tests are not recommended when sample sizes are less than about 20.

D7.4 At Least One of the Two Paired Data Sets Contains More Than 0 Percent, But Less Than 100 Percent Non-detects

In this case, a nonparametric test should be used. Again, the Wilcoxon signed-rank test, available in statistical software like JMP, may be used. If the power of the Wilcoxon signed-rank test is low, consider using the alternative Paired-Prentice Wilcoxon (PPW) test. The PPW test evaluates whether there is a difference in the distributions of the two sample groups. This test is currently unavailable in JMP. Consultation with a statistician is recommended.

D7.5 No Detects in Both Data Sets

No statistical analysis is possible for this case. Professional judgment informed by qualitative factors may be used to compare the two data sets and draw conclusions. For BMP studies, the question of whether a BMP is effective may be moot if all data are non-detects.

D7.6 Description of Statistical Tests for Comparison of Two Paired Data Sets

The flowchart in Figure D 7-2 identifies the following statistical tests:

- Paired *t*-test

- Wilcoxon signed-rank test
- Permutation test
- Paired Prentice-Wilcoxon test

The first test is parametric (i.e., it assumes a particular probability distribution), while the remaining tests are nonparametric (i.e., they assume no specific distribution). A brief description of the Paired t-test and the Wilcoxon signed-rank test follows.

D7.6.1 Paired t-Test

This test can be used to evaluate the differences between two paired data sets when each data set contains only detects. The null and alternative hypotheses are defined as follows:

- Null Hypothesis, H_0: Mean difference ≤ 0 (i.e., the BMP is ineffective in reducing pollutant concentration)
- Alternative Hypothesis, H_A: Mean difference > 0 (i.e., the BMP does reduce the pollutant concentration)

The key assumption for this test is that the differences between the paired observations are normally distributed. For BMP data, the percent reduction in pollutant concentration, rather than the actual difference in pollutant concentrations, is more likely to be centered on a fixed mean with a constant variance, and hence more likely to be normally distributed. The test is robust to moderate violations of the normality assumption, but not to outliers. The normality assumption can be checked with graphical methods (histogram and normal probability plot described in Appendix D2) and a formal goodness-of-fit test (Shapiro-Wilk W test described in Appendix D4). Methods described in Appendix D2 can be used to check whether the data contain potential outliers.

Example D7-1

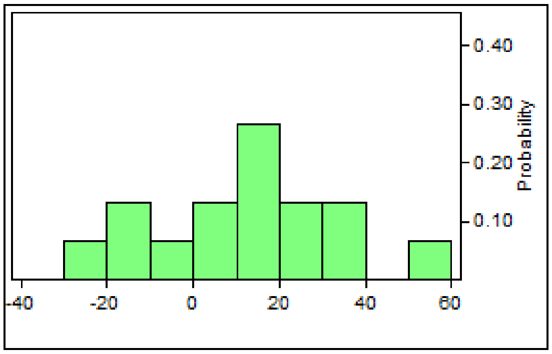
Table D 7-1 shows lead concentration data collected in an influent-effluent BMP monitoring study. The influent and effluent data are paired since both measurements are taken for the same BMP. If the BMP is effective, the average percent reduction, rather than the average actual difference, in the lead concentration is likely to be constant over a range of pollutant concentrations. Table D 7-1 also shows the percent reduction in lead concentration.

Table D 7-1 Influent-Effluent Lead Concentration Data for Example D7-1

Station	Influent Lead Concentration (mg/L)	Effluent Lead Concentration (mg/L)	Percent Reduction
1	9.45	8.22	13
2	11.73	10.05	14.3
3	6.34	4.56	28.1
4	2.25	2.5	-11.1
5	8.53	3.84	55
6	7.68	7.53	2
7	3.87	4.5	-16.3
8	8.85	5.66	36
9	6.61	5.15	22.1
10	1.71	2.1	-22.8
11	2.68	2.35	12.3
12	0.9	0.63	30
13	2.6	2.7	-3.8
14	9.07	7.88	13.1
15	3.7	3.5	5.4

Figure D 7-3 shows a histogram, a normal probability plot, and the results of the Shapiro-Wilk W test for the percent reduction data. The methods described in Appendix D2 were used to check for outliers; no outliers were identified. The assumption of a normal distribution appears to be reasonable for this data. Therefore, the paired t-test is appropriate to use. The JMP software was used to apply this test to the percent reduction data in Table D 7-1. The results for both one- and two-tailed tests are shown in Figure D 7-4 (see Appendix D6 for a discussion of one- and two-tailed tests). Note that the probability distribution of the percent reduction shown in Figure D 7-4 highlights both tails. For this example, the interest is in assessing whether the BMP reduces the lead concentration and hence the one-tailed test is applicable. The significance level, p, of the test is 0.0225, as highlighted in the figure (because this is a one-tailed test and the paired t-test was performed because of interest in evaluating reduction in effluent concentrations, the one-tailed p associated with this comparison is "Prob > t"). Assuming that the acceptable risk of false rejection, α, is set to 0.05, and since $p < \alpha$, one will reject the null hypothesis and accept the alternative hypothesis that the average percent reduction in lead concentration is greater than 0 percent. In other words, the BMP is working effectively to reduce lead concentrations. Note that for Caltrans studies, the 90 percent confidence level should be used when evaluating the performance of BMPs.

Figure D 7-3 Histogram Normal Probability Plot, and Shapiro-Wilk W Test Results for Example D7-1, Percent Reduction of Zinc Concentration



Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.982582	0.9841

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

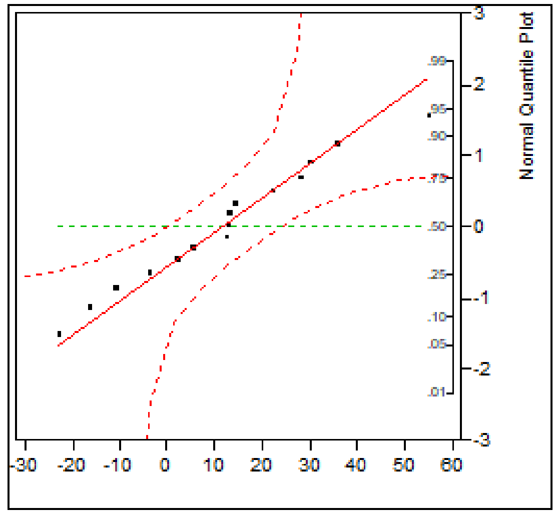


Figure D 7-4 Histogram Normal Probability Plot, and Shapiro-Wilk W Test Results for Example D7-1, Percent Reduction of Zinc Concentration

Test Mean=value

Hypothesized Value	*0*
Actual Estimate	11.82
Df	14
Std Dev	20.7985

	t Test
Test Statistic	2.2011
Prob > |t|	0.0450
Prob > t	**0.0225**
Prob < t	0.9775

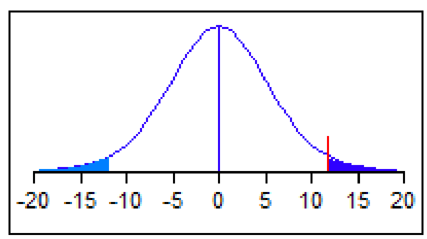


Figure D 7-5 Paired t-Test Results for Example D7-1, Percent Reduction of Zinc Concentration

D7.6.2 Wilcoxon Signed-Rank Test

This test can be used to evaluate differences in the medians of two paired data sets when: (1) the data differences do not fit a normal or lognormal distribution, or (2) at least one data set contains non-detects (but not 100 percent NDs in both data sets). For BMP studies, the null and alternative hypotheses may be defined as follows:

- Null Hypothesis, H_0: Median difference ≤ 0 (i.e., the BMP is ineffective in reducing pollutant concentration)
- Alternative Hypothesis, H_A: Median difference > 0 (i.e., the BMP does reduce the pollutant concentration)

The test calculates and ranks the absolute differences between the paired observations. The sum of the positive signed ranks is used to calculate the test statistic. The Wilcoxon signed-rank test is not recommended for very small sample sizes (< 5).

Since the test does not use the actual magnitude of the differences between paired observations, and instead uses only the rank and sign of the differences, one can use a simple, although approximate, procedure to handle non-detects. The procedure will be to set non-detects to half the reporting limit. This way, all pairs of data could be used for the Wilcoxon signed-rank test, thus improving the power of the test to detect differences. If the data contain multiple reporting limits, one option would be to censor the data at the highest reporting limit, but a substantial amount of useful information could be lost. It may be necessary to consult with a statistician to use a more rigorous method.

Example D7-2

For this example, the data for Example D7-1 will be modified to include some non-detects. The modified data are shown in Table D 7-2. The non-detects will be replaced with half of the reporting limit.

The JMP results of the Wilcoxon signed-rank test for this example are shown in Figure D 7-5. The significance level, *p*, for the test is 0.8853, as highlighted in the figure (because this is a one-tailed test and the Wilcoxon signed-rank test was performed because of interest in evaluating reduction in effluent concentrations, the one-tailed *p* associated with this comparison is “Prob > t” in the JMP output). Assuming that the acceptable risk of false rejection, α, is set to 0.05, and since $p > \alpha$, one cannot reject the null hypothesis. In other words, the BMP does not appear to be working effectively to reduce lead concentrations. Note that for Caltrans studies, the 90 percent confidence level should be used when evaluating the performance of BMPs.

Table D 7-2 Influent-Effluent Lead Concentration Data for Example D7-2

Station	Influent Lead Concentration (mg/L)	Effluent Lead Concentration (mg/L)	Percent Reduction
1	9.45	8.22	13
2	11.73	10.05	14.3
3	6.34	4.56	28.1
4	2.25	2.5	-11.1
5	8.53	3.84	55
6	7.68	7.53	2
7	3.87	4.5	-16.3
8	8.85	5.66	36
9	6.61	5.15	22.1
10	<0.5	2.1	-740
11	<0.5	2.35	-840
12	<0.5	0.63	-152
13	<0.5	2.7	-980
14	<0.5	7.88	-3052
15	<0.5	3.5	-1300

Test Mean=value

Hypothesized Value	*0*
Actual Estimate	-461.39
Df	14
Std Dev	847.455

	t Test	*Signed-Rank*
Test Statistic	-2.1086	-22.0000
Prob > |t|	0.0535	0.2293
Prob > t	0.9733	0.8853
Prob < t	0.0267	0.1147

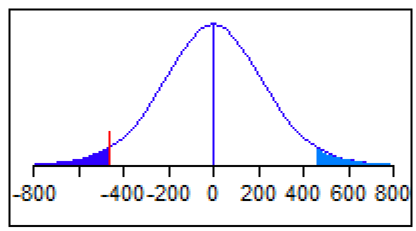


Figure D 7-6 Wilcoxon Signed-Rank Test Results for Example D7-2, Percent Reduction of Zinc Concentration

Appendix D8 How to Compare Three or More Independent Data Sets

D8.1 Purpose and Organization

The purpose of this appendix is to present statistical methods to compare three or more independent data sets and draw conclusions regarding whether observed differences in the data sets are statistically significant. Typical study questions of interest in the analysis of three or more independent data sets are:

- Four pilot BMPs are close to each other. Are the influent water quality characteristics different from each other or essentially the same? If the BMPs behave differently, which specific BMPs are different from others?
- Given influent and effluent data for three pilot BMPs of the same type, how can one tell whether they are all operating in an equivalent manner?
- Three different downstream watersheds each with a BMP are being monitored for certain pollutants. Are the average pollutant concentrations in the three watersheds similar or significantly different?

In this appendix, a decision flowchart is presented to guide the user in the selection of an appropriate statistical method, depending on the characteristics of each data set. An overview of each method is presented and examples related to BMP studies are provided to illustrate the use of each method.

D8.2 Decision Flowchart

The statistical methods to compare three or more independent data sets fall under the category of Analysis of Variance (ANOVA). The Student's *t*-test, described in Appendix D6 for comparison of two independent data sets, is a special case of ANOVA methods. As in Appendix D6, the choice of an appropriate statistical method depends on the percentage of non-detects in the different data sets and the probability distribution that can be assumed for each data set. Figure D 8-1 provides a flowchart to guide the user in the selection of an appropriate statistical method, depending on the characteristics of each data set. The main pathways of the flowchart are discussed below. The details of each method and illustrative examples are provided in the next section.

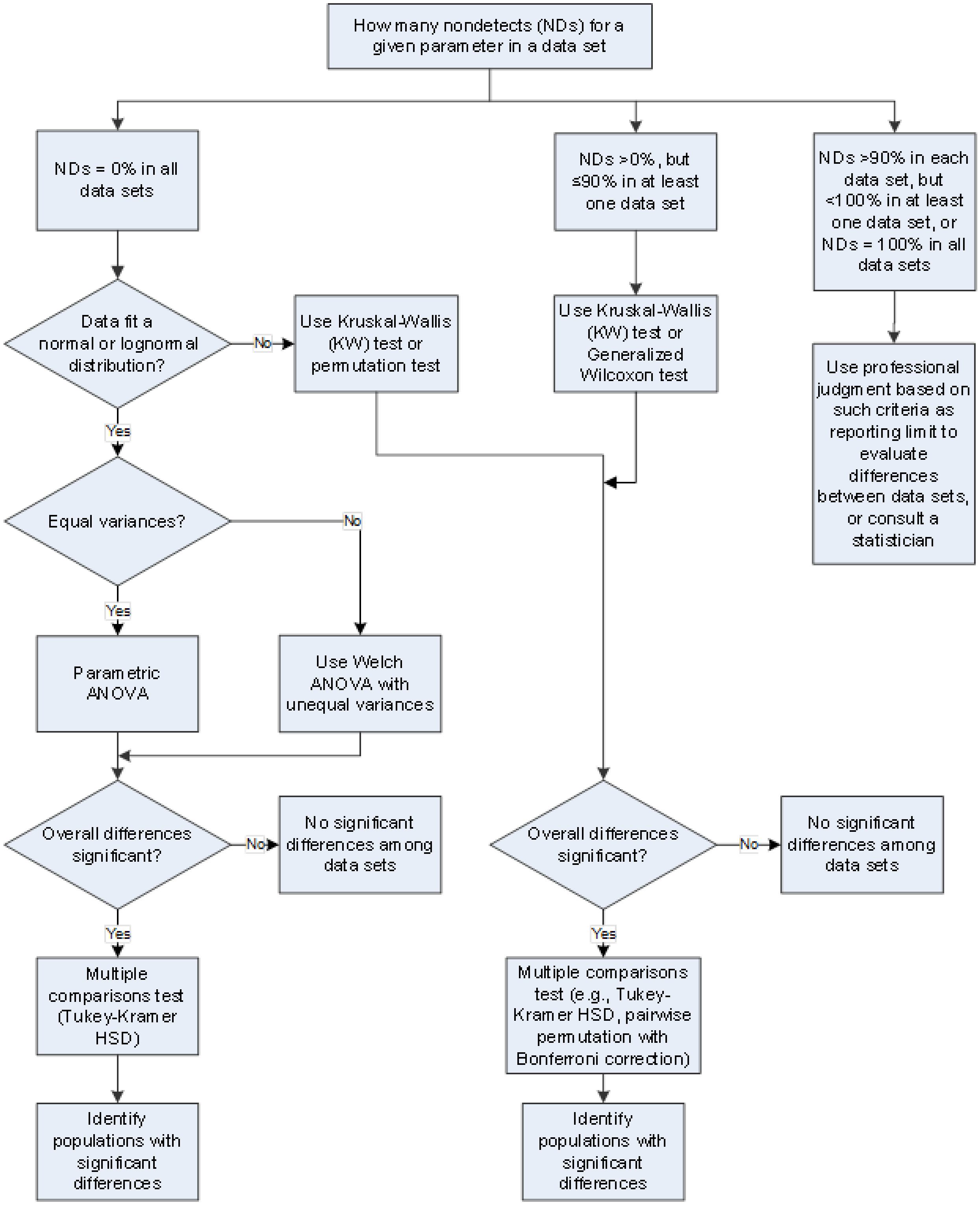


Figure D 8-1 Select an Appropriate Statistical Method for Comparing Three or More Independent Data Sets

D8.3 No Non-detects in Any Data Set

If there are no non-detects in any of the data sets, the possibility of using a parametric method (i.e., one that assumes a specific probability distribution) can be explored. If the original or log-transformed data in each set can be assumed to follow a normal distribution, then further check whether the variances of the data sets are equal. If they are equal, the parametric ANOVA method can be used. If they are not equal, one can use the Welch ANOVA method, which still assumes that each data set is normally distributed, but allows for unequal variances.

If the parametric or Welch ANOVA methods show the differences among the data sets are statistically significant, the next step is to identify the specific data sets that are different from other data sets. This part of the ANOVA is called the multiple comparisons test. One such test is the Tukey-Kramer Honestly Significant Difference (HSD) test for making multiple comparisons. This test will be described in the next section.

If a normal distribution cannot be assumed for one or more data sets, a nonparametric method should be used. One common method in this situation is the Kruskal-Wallis (KW) test. For the multiple comparisons test in this case, the rank-transformed data should be used for the Tukey-Kramer HSD test. If the power of the KW test is low, consider using a permutation test. Permutation tests are described in Section D6.4.4. Permutation tests are not recommended when sample sizes are less than about 20. In JMP, permutation tests can be conducted in the Simulate platform (available only in JMP Pro). If the permutation test shows the differences among data sets are statistically significant, one has to perform a multiple comparisons test to identify the specific data sets that are different from other data sets. One simple approach is to perform pairwise permutation tests (permutation tests for every possible pairs of data sets) and correct for Type I error using the Bonferroni procedure. The correction to the Type I error is required because multiple pairwise comparisons increase the risk of incorrectly concluding data sets are different simply due to chance. Empirical methods that provide better control over the proportion of Type I errors are also available.

D8.3.1 <u>Non-detects are Greater than 0 Percent, but Less Than 90 Percent in at Least One Data Set</u>

In this case, the nonparametric methods KW test and Tukey-Kramer HSD test with rank-transformed data are recommended. If the power of the KW test is low, consider using a Generalized Wilcoxon test. The Generalized Wilcoxon test, also called the Peto-Prentice test, is a weighted log rank test that compares the difference between cumulative distribution functions of the two data sets to determine whether their percentiles differ.

The test is able to accommodate unequal sample sizes and multiple reporting limits. Consultation with a statistician is recommended.

D8.3.2 Non-detects are Greater than 90 Percent in Each Data Set, but Less Than 100 Percent in at Least One Data Set

When the percentage of non-detects is high (>90 percent), the location (mean or median) of a data set is not well defined. Hence, a method that tests for a shift in location would not be reliable, even when using a nonparametric method. In this case, no simple method is available and consultation with a statistician may be necessary to investigate the use of an advanced method such as logistic regression for nominal response.

D8.3.3 No Detects in Any Data Set

No statistical analysis is possible for this case. Professional judgment informed by qualitative factors may be used to compare the data sets and draw conclusions.

D8.4 Description of Statistical Tests for Comparison of Three or More Independent Data Sets

The flowchart in Figure D 8-1 identifies the following statistical tests:

- Parametric ANOVA (equal variances)
- Welch ANOVA (unequal variances)
- Tukey-Kramer HSD test for multiple comparisons
- Kruskal-Wallis (KW) test
- Permutation test
- Bonferroni correction for multiple comparisons
- Generalized Wilcoxon test

Descriptions of the parametric ANOVA (equal variances), Welch ANOVA (unequal variances), Tukey-Kramer HSD test for multiple comparisons, and Kruskal-Wallis (KW) tests are provided below.

D8.4.1 Parametric ANOVA (Equal Variances)

This test can be used to evaluate the difference between the means of three or more independent data sets when the sample data contain no non-detect measurements. The key assumptions made in this test are:

- Each data set (in the original or log-transformed scale) follows a normal distribution; and
- The variances of the data sets are equal.

The test is robust to moderate violations of the normality assumption, but not to large inequalities of variances. It is also not robust against outliers because sample means and standard deviations are sensitive to outliers. The normality assumption can be checked with graphical methods (histogram and normal probability plot described in Appendix D2) and a formal goodness-of-fit test (Shapiro-Wilk W test described in Appendix D4). Levene's test described in Appendix D4 can be used to test the equality of variances. Methods described in Appendix D2 can be used to check whether the data contain potential outliers.

The ANOVA method belongs to the class of hypothesis testing methods. As such, the basic concepts of hypothesis testing and the flowchart of the main steps in performing hypothesis testing, which were described in Appendix D6, apply to ANOVA. Specifically, the significance level, *p*, of the ANOVA is the key statistic to interpret the results. The smaller the *p*, the greater are the differences between the population means and the smaller is the risk of false rejection; that is, incorrectly concluding that the population means are different when, in fact, they are the same.

Example D8-1

The effectiveness of three different BMPs in a similar environmental setting is to be evaluated. The influent-effluent monitoring is used to obtain data on percent reduction in lead concentrations, which is shown in Table D 8-1. Is the average percent reduction about the same for the three BMPs?

Let μ_1, μ_2, and μ_3 denote the population mean percent reduction in lead concentration for the three BMPs. Then, the null and alternative hypotheses are defined as follows:

- Null hypothesis, H_0: $\mu_1 = \mu_2 = \mu_3$ (i.e., the three BMPs perform similarly)
- Alternative hypothesis, H_A: One or more population means are different from others

Figure D 8-2 shows a histogram, a normal probability plot, and the results of the Shapiro-Wilk W test for each data set. The assumption of a normal distribution appears to be reasonable for each data set. Figure D 8-3 shows a box plot for the three data sets and the results of Levene's test for equality of variance. The assumption of equal variance for the three data sets appears to be reasonable. Also, the methods described in Appendix D2 were applied to check for outliers and no outliers were identified. Therefore, it is appropriate to use parametric ANOVA to evaluate differences in the population means.

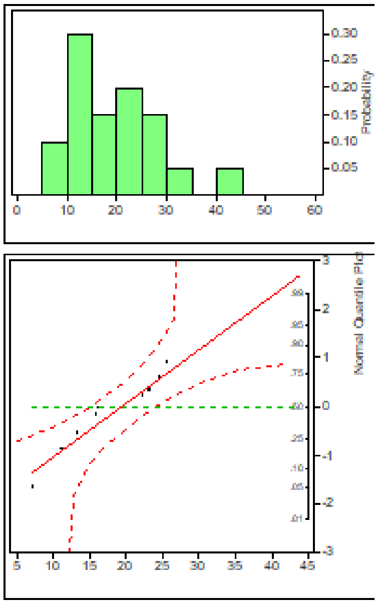
Table D 8-1 Percent Reduction in Lead Concentration Data for Example D8-1

BMP 1	BMP 2	BMP 3
7.3	9	17.1
8.9	9.5	18.1
10.2	9.7	18.3
11.2	10.5	20
12.9	11.2	25.1
13.5	17	25.8
14	17.1	27
14.9	18.5	29.4
15.8	18.6	30.2
16.3	19.5	31.6
18.2	19.8	31.8
22.4	20.6	33.5
23.3	21.2	34
23.8	21.5	37.8
24.6	22.1	38
25.4	22.3	42.8
25.7	22.3	43.9
27.9	31.7	44.7
31.8	31.9	51.5
44.1	39.8	57.7

The JMP software was used to perform the parametric ANOVA. The results are shown in Figure D 8-4. The key part of the results is the significance level, *p*, which is labeled "Prob>F" in the JMP output. For the example, *p* is <0.0001 as highlighted in Figure D 8-4. Assume that an acceptable risk of false rejection, α is set to its typical value of 0.05, and since $p < \alpha$, one will reject the null hypothesis and accept the alternative hypothesis that the mean percent reduction in lead concentration is different among the BMPs. Note for Caltrans studies, the 90 percent confidence level should be used.

Since the conclusion of ANOVA is that the three means are not equal (i.e., the three BMPs do not perform similarly), the next step in the analysis is to identify which BMPs perform differently. This step involves performing a multiple comparisons test.

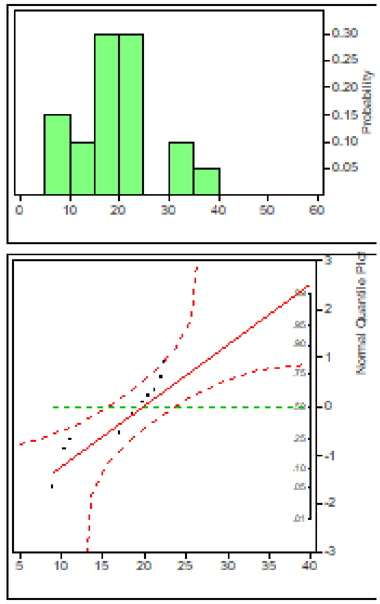
Distributions - BMP 1



Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.928649	0.1454

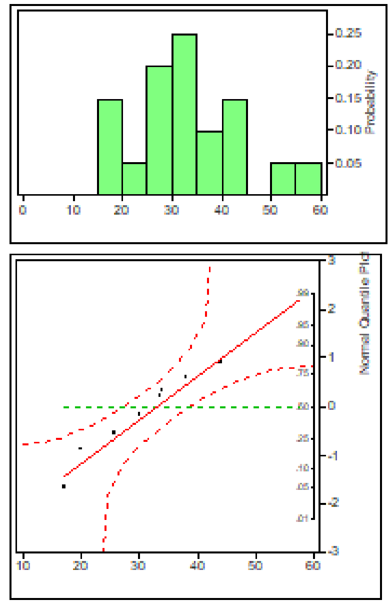
Distributions - BMP 2



Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.904780	0.0507

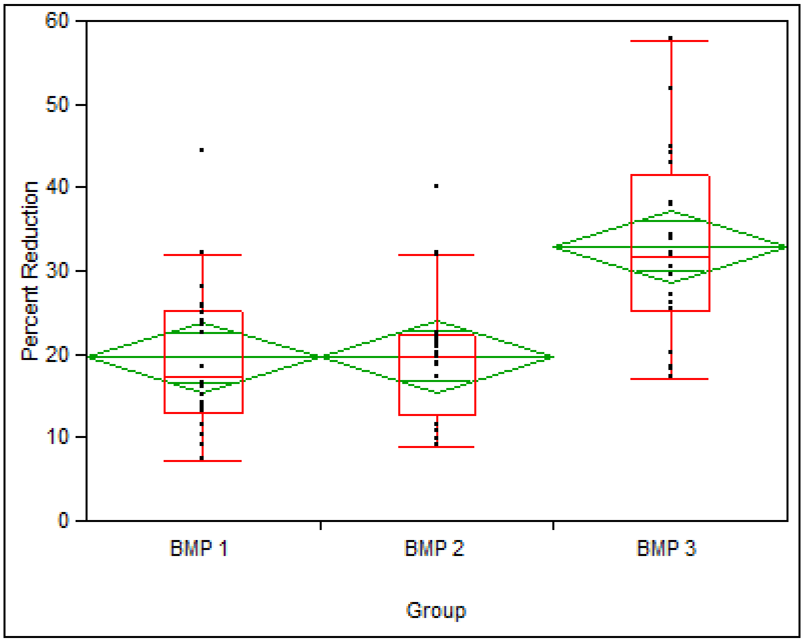
Distributions - BMP 3



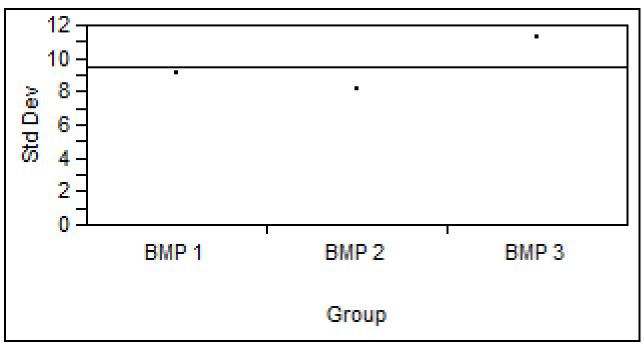
Goodness-of-Fit Test
Shapiro-Wilk W Test

W	*Prob<W*
0.959923	0.5423

Figure D 8-2 Histogram, Normal Probability Plot, and Shapiro-Wilk W Test Results for Example D8-1



Tests that the Variances are Equal



Level	*Count*	*Std Dev*	*MeanAbsDif to Mean*	*MeanAbsDif to Median*
BMP 1	20	9.00450	7.251000	7.110000
BMP 2	20	7.99539	5.630000	5.630000
BMP 3	20	11.20992	8.766500	8.655000

Test	*F Ratio*	*DFNum*	*DFDen*	*Prob > F*
O'Brien[.5]	1.0378	2	57	0.3608
Brown-Forsythe	1.2186	2	57	0.3032
Levene	1.4607	2	57	0.2406
Bartlett	1.1147	2	-	0.3280

Figure D 8-3 Box Plot and Levene's Test Results for Example D8-1

Oneway Anova
Summary of Fit

Rsquare	*0.313289*
Adj Rsquare	0.289194
Root Mean Square Error	9.498589
Mean of Response	24.07167
Observations (or Sum Wgts)	60

Analysis of Variance

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Ratio*	*Prob > F*
Group	2	2346.2003	1173.10	13.0022	<.0001
Error	57	5142.7215	90.22	-	-
C. Total	59	7488.9218	-	-	-

Means for Oneway Anova

Level	*Number*	*Mean*	*Std Error*	*Lower 95%*	*Upper 95%*
BMP 1	20	19.6100	2.1239	15.357	23.863
BMP 2	20	19.6900	2.1239	15.437	23.943
BMP 3	20	32.9150	2.1239	28.662	37.168

Figure D 8-4 Parametric ANOVA Results for Example D8-1

D8.4.2 Welch ANOVA (Unequal Variances)

If each data set (in the original or log-transformed scale) follows a normal distribution, but the variances of the data sets are found to be significantly different (using the methods described in Appendix D4), the Welch ANOVA test should be used. This test is available in statistical software packages, such as JMP. The results of this test are organized the same way as those shown in Figure D 8-4. Conclusions regarding whether the population means are the same or significantly different are drawn based on the significance level, *p*, which is labeled "Prob>F" in the JMP output. Because the results of the Welch ANOVA are displayed and interpreted the same way as those for the parametric ANOVA with equal variances, a separate example of this test is not shown.

D8.4.3 Tukey-Kramer HSD Test for Multiple Comparisons

This multiple comparison test is applied to find out which population means are different if the ANOVA or Welch ANOVA test shows that the population means are not equal. For three or more populations, there will be multiple pairs of population means and the difference in the means for each pair needs to be tested. Each test would be subject to a risk of false rejection. As the number of pairs to be compared increases, the risk of false rejection for at least one of the pairs is higher than the risk of false rejection for any one single pair. This is analogous to making multiple coin tosses and finding the probability of observing heads. The probability of a head coming up on any single toss of a fair coin is 0.5. However, if many tosses were to be made, the probability of getting a head on *at least* one of the tosses would be higher. It is desirable to make sure that the overall risk of false rejection for all pairs of means is controlled at a specified level, α (i.e., the probability (confidence) of reaching the correct conclusion of no difference among all population pairs means when, in fact, all population means are the same would be (1-α)). The Tukey-Kramer HSD test is designed to control the total risk of false rejection at the specified level, α. Equivalently, it is designed to provide the overall confidence of (1-α) in reaching the correct conclusion of no difference when all population means are the same. Its use is best illustrated with an example for the case when the parametric ANOVA is valid to use. The flowchart in Figure D 8-1 shows that the Tukey-Kramer HSD test is also used on rank-transformed data following a significant result of the KW test. The use of rank-transformed data in the Tukey-Kramer HSD test is discussed in a following section.

Example D8-2

For this example, the data from Example D8-1 involving three BMPs will be used. As discussed above, the parametric ANOVA showed that the means of the three BMPs are not equal. Now, one wants to find out which BMP(s) have different means. Figure D 8-5 shows the results of the Tukey-Kramer HSD test, which were obtained using the JMP software. The overall risk of false rejection, α, is controlled at 0.05 as shown in Figure D 8-5. Each BMP is assigned one or more letters. The BMPs that do not have a common letter have significantly different means. BMP 3 has a letter “A” and does not share the same letter with BMP 1 and BMP 2 (which both have a letter “B”), which means that BMP 3 is significantly different from BMP 1 and BMP 2, while BMP 1 and BMP 2 are similar to each other. Note for Caltrans studies, the 90 percent confidence level should be used when comparing the performance of BMPs.

Means Comparisons
Comparisons for all pairs using Tukey-Kramer HSD

*q**	*Alpha*
2.40642	0.05

Abs(Dif)-LSD	*BMP 3*	*BMP 2*	*BMP 1*
BMP 3	-7.2282	5.9968	6.0768
BMP 2	5.9968	-7.2282	-7.1482
BMP 1	6.0768	-7.1482	-7.2282

Positive values show pairs of means that are significantly different.

Level		*Mean*
BMP 3	A	32.915000
BMP 2	B	19.690000
BMP 1	B	19.610000

Levels not connected by same letter are significantly different.

Level	*- Level*	*Difference*	*Lower CL*	*Upper CL*	*Difference*
BMP 3	BMP 1	13.30500	6.07678	20.53322	
BMP 3	BMP 2	13.22500	5.99678	20.45322	
BMP 2	BMP 1	0.08000	-7.14822	7.30822	

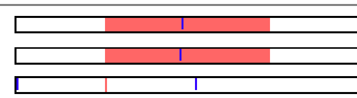


Figure D 8-5 Tukey-Kramer HSD Test Results for Example D8-2

D8.4.4 Kruskal-Wallis (KW) Test

The KW test is the nonparametric equivalent of the parametric ANOVA. It is used when at least one data set contains one or more non-detects or when at least one data set is not normally distributed. The KW test compares differences in the mean ranks of the data sets. This is equivalent to testing if the medians of the data sets are different when the shapes of the distributions are approximately similar. If the power of the Kruskal-Wallis test is low, consider using a permutation test when there are no non-detects or a Generalized Wilcoxon test if there are non-detects. Permutation tests are described in Section D6.4.4. Permutation tests are not recommended when sample sizes are less than about 20.

Example D8-3

Data from Example D8-1 will be used in this example; however, assume that some of the measurements are non-detects and the shapes of the distributions are approximately similar. Table D 8-2 shows the data that will be used for this example.

Figure D 8-6 shows the results of the JMP software for this example. The significance level, *p*, of the test is 0.0004, as highlighted in the figure. Assuming that the acceptable risk of false rejection, α, is set to 0.05, and since $p < \alpha$, one will reject the null hypothesis and accept the alternative hypothesis that the median percent reduction in lead concentration is different among the BMPs. Note for Caltrans studies, the 90 percent confidence level should be used when comparing the performance of BMPs.

Table D 8-2 Percent Reduction Lead Concentration Data for Example D8-3

BMP 1	BMP 2	BMP 3
<5	<5	<5
<5	<5	18.1
<5	<5	18.3
11.2	<5	20
12.9	<5	25.1
13.5	17	25.8
14	17.1	27
14.9	18.5	29.4
15.8	18.6	30.2
16.3	19.5	31.6
18.2	19.8	31.8
22.4	20.6	33.5
23.3	21.2	34
23.8	21.5	37.8
24.6	22.1	38
25.4	22.3	42.8
25.7	22.3	43.9
27.9	31.7	44.7
31.8	31.9	51.5
44.1	39.8	57.7

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	*Count*	*Score Sum*	*Score Mean*	*(Mean-Mean0)/Std0*
BMP 1	20	487.500	24.3750	-1.916
BMP 2	20	480.000	24.0000	-2.034
BMP 3	20	862.500	43.1250	3.958

1-way Test, ChiSquare Approximation

ChiSquare	*DF*	*Prob>ChiSq*
15.7357	2	0.0004

Figure D 8-6 Kruskal-Wallis (KW) Test Results for Example D8-3

If the KW test shows a significant result (i.e., the significance level of the KW test is less than the specified α), one should assess which population medians are different using the Tukey-Kramer HSD test on the rank-transformed data. Statistical software packages such as JMP provide a procedure to obtain the ranks of the original data. The JMP procedure calculates the average rank of all tied observations and assigns this rank to each tied observation. To use this procedure for a data set containing non-detects, all non-detects are set equal to zero. JMP will assign the average rank of all non-detects to each non-detect. Once the rank-transformed data are obtained, the Tukey-Kramer HSD test is applied the same way as described above for the parametric ANOVA. Tukey-Kramer HSD test results for Example D8-3 are shown in Figure D 8-7.

Means Comparisons
Comparisons for all pairs using Tukey-Kramer HSD

*q**	*Alpha*
2.40642	0.05

Abs(Dif)-LSD	*BMP 3*	*BMP 1*	*BMP 2*
BMP 3	-11.559	7.191	7.566
BMP 1	7.191	-11.559	-11.184
BMP 2	7.566	-11.184	-11.559

Positive values show pairs of means that are significantly different.

Level			*Mean*
BMP 3	A		43.125000
BMP 1		B	24.375000
BMP 2		B	24.000000

Levels not connected by same letter are significantly different.

Level	*- Level*	*Difference*	*Lower CL*	*Upper CL*
BMP 3	BMP 2	19.12500	7.5662	30.68378
BMP 3	BMP 1	18.75000	7.1912	30.30878
BMP 1	BMP 2	0.37500	-11.1838	11.93378

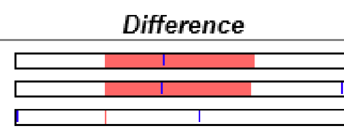


Figure D 8-7 Tukey-Kramer HSD Test Results for Example D8-3

Appendix D9 How to Develop a Linear Regression Equation

D9.1 Purpose and Organization

The purpose of this appendix is to present methods to develop a predictive relationship between a response variable y (e.g., percent removal of a pollutant concentration) and several explanatory variables x_1, x_2, …, x_m (e.g., storm characteristics such as rainfall amount, rainfall intensity, and antecedent weather conditions). When the predictive relationship is assumed to be linear in the explanatory variables, methods of multiple linear regression can be used. If only one explanatory variable is of interest and a linear relationship is appropriate, the method is called simple linear regression, which is a special case of multiple linear regression analysis.

A multiple linear regression equation between a response variable, y, and a set of explanatory variables, $x_1, x_2, \ldots, x_m$ has the following form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m + \varepsilon$$

where: β_0, β_1, …, β_m are model parameters, called regression coefficients, and ε = random error term, assumed to be normally distributed with a mean of zero and variance of σ^2.

The regression coefficients are estimated using the method of "least squares," which minimizes the sum of squared deviations between the measured y and its expected value over all measurements. The estimated regression coefficients are denoted as b_0, b_1, …, b_m.; that is, $\hat{\beta}_i = b_i$.

There are a number of assumptions that have to be met when using linear regression:

- The relationship is linear
- The errors, or residuals, of the regression are random fluctuations around the true line
- The variability in the response does not increase as the value of the explanatory variable increases. This is the assumption of equal variance.
- The observations are independent of one another (correlation between observations can be an issue with time series data.)

In JMP, linear regression can be performed in the Fit Model platform. The most useful graph for analyzing residuals is a residual by predicted plot, which plots each residual value against the corresponding predicted value. If the above assumptions are met, the residuals will be randomly scattered around the center line of zero, with no obvious pattern. If the assumptions are not met, consider transforming the data. Log transformation of the response variable is commonly used when dealing with environmental data.

Consider using quantile regression when the assumptions for linear regression are not met after transforming the data. Quantile regression is particularly useful when the response of interest is not the mean, but instead is a percentile response such as the median or 50th percentile response. Unlike ordinary least squares regression, quantile regression makes no assumptions about the distribution of the residuals nor does it assume a constant variance for the response variable. The method is also more robust to statistical outliers than standard least squares linear regression. The computation of the quantile regression coefficients is different compared with linear regression because the sum of weighted absolute residuals is minimized instead of sum of squared residuals. Quantile regression is available in the Generalized Regression option of the Fit Model platform of JMP Pro.

D9.2 Steps to Perform Multiple Linear Regression

The main steps in performing multiple linear regression are:

- Perform a screening analysis to filter potential explanatory variables.
- Run stepwise multiple regression analysis and select a model for further investigation.
- Verify key assumptions of a multiple linear regression model.
- Use the selected regression model to predict the response variable for specified values of the explanatory variables.

A common example will be used throughout this appendix so that each of the four steps can be illustrated based on the conclusions drawn from the previous step. The data for this example is presented first and then each of the steps identified above are described. The final section provides a discussion of how regression analysis could be used to compare the performance of different BMPs based on predicted effluent concentrations.

<u>Example D9-1</u>

Soil properties and site characteristics control many of the hydrologic and sediment aspects of stormwater, which in turn affect infiltration and stormwater runoff volume. The estimation of stormwater runoff volume is directly related to the runoff coefficient. The purpose of the statistical analysis described in this example is to derive a predictive relationship between the runoff coefficient and relevant soil and site characteristics. The results of the analysis could be used to identify the soil and site characteristics that have the most influence on the runoff coefficient and to develop effective design and construction practices to control runoff volume.

Data on relevant soil and site characteristics were collected or estimated at 23 sites. Using historical data on rainfall and runoff volume, the runoff coefficient was estimated at each site. Table D 9-1 lists the data on runoff coefficient and relevant soil and site characteristics collected at the 23 sites.

Table D 9-1 Runoff Coefficient and Relevant Soil and Site Characteristics Data for Example D9-1

Site & System	Avg. Runoff Coeff.	Avg Strip Width (m)	Slope (%)	Avg. Veg. Cover (%)	Est. HRT[1] (min)	Relative Compaction (%)	Dry Density (lb/ft3)	Infiltration Rate (in/hr)	Porosity (%)	Gravel (%)	Sand (%)	Silt/Clay (%)
Sacramento 2	0.31	1.1	2.2	92.5	4.58	93.5	121.6	2.96	29.2	51.8	36.9	11.3
Sacramento 3	0.32	4.6	33	83.725	4.79	81.2	105.6	2.68	38.5	31.9	36.5	31.6
Sacramento 4	0.28	6.6	33	91.5	5.95	79.7	103.6	2.35	39.6	32.5	36.5	31
Sacramento 5	0.15	8.4	33	89.775	7.67	78.4	101.9	3.14	40.6	39.2	35.8	25
Cottonwood 2	0.19	9.3	52	73.075	7.12	85.8	111.5	3.5	33.3	44	41.6	14.4
Redding 2	0.57	2.2	10	79.575	4.91	93.9	129.3	1.89	27.1	39.6	48.8	11.6
Redding 3	0.31	4.2	10	85.425	7.24	93	128.9	3.34	27.3	47.2	42.5	10.3
Redding 4	0.45	6.2	10	87.1	8.20	88.6	122.6	4.15	30.8	34.7	52.8	12.5
San Rafael 2	0.13	8.3	50	83.975	6.72	78.8	107.1	9.29	35.9	40.6	38.6	20.8
Irvine 2	0.39	3.3	11	70.15	5.46	88.4	108.7	1.54	34	24.9	59.9	15.2
Irvine 3	0.05	6	11	63.425	7.81	84.7	104.9	1.65	36.3	16.7	59.5	23.8
Irvine 4	0.16	13	11	62.225	12.43	87.6	107.8	0.92	34.6	20.1	46.5	33.4
Yorba Linda 2	0.37	1.85	14	61.05	3.59	89.2	114.7	1.26	33.4	28.1	53.4	18.5
Yorba Linda 3	0.51	4.9	14	82.375	6.44	82.5	106	0.87	38.5	25.3	53.5	21.2
Yorba Linda 4	0.58	7.6	14	74.45	8.38	87.7	112.7	1.57	34.6	17.2	60.6	22.2
Yorba Linda 5	0.17	13	14	75.575	11.56	86.8	111.6	1.81	35.2	34.2	49.6	16.2
Moreno Valley 2	0.95	2.6	13	3.05	1.13	90.7	123.4	0.72	28.9	20.3	61.5	18.2
Moreno Valley 3	0.95	4.9	13	16.3	1.65	93.3	126.6	0.57	27	29.7	53	17.3
Moreno Valley 4	0.48	8	13	21.575	2.22	92.9	125.8	0.94	27.5	16.5	59.1	24.4
Moreno Valley 5	0.51	9.9	13	18.225	2.52	93.9	127.3	1.04	26.6	13.7	70.2	16.1
San Onofre 2	0.45	1.3	8	81.2	3.43	95.9	122.7	2.25	27.4	19	63.8	17.2
San Onofre 3	0.27	5.3	10	73.8	7.46	88.5	114.7	1.25	32.2	27.1	56.8	16.1
San Onofre 4	0.07	9.9	16	69.1	9.43	85.3	108.3	0.75	36	21.7	55.7	22.6

[1]HRT = Hydraulic Residence Time

D9.2.1 Step 1. Perform a Screening Analysis to Filter Potential Explanatory Variables

When a large number of potential explanatory variables have been identified, it is desirable to reduce the number of explanatory variables. A smaller number of explanatory variables helps to increase the computational efficiency of multiple regression analysis and to obtain more stable estimates of the regression coefficients. Both graphical and numerical methods are recommended to filter the potential explanatory variables.

In the graphical method, a scatterplot of the response variable against each explanatory variable is prepared. This plot is used to assess whether there is a reasonable linear relationship between the two variables. The development and interpretation of scatterplots are described in Appendix D2. If a scatterplot shows mostly random noise and little structure, the explanatory variable may be excluded from further analysis. If the scatterplot shows a nonlinear relationship, an appropriate transformation (e.g., log transformation) may be used on the response variable, the explanatory variable, or both to develop a more linear relationship between the two variables. With regard to the explanatory variables, the transformation decision may be made separately for each variable. Thus, one could have a mix of raw and transformed explanatory variables in a multiple regression equation.

In the numerical method, a correlation matrix is developed that shows the simple correlation coefficient between each pair of variables. This matrix has (m +1) rows - one row for the response variable and one row for each of the m explanatory variables. Similarly, it has (m +1) columns. Thus, each cell in this matrix represents one pair of variables. In each cell, the matrix displays simple correlation coefficient between the two variables represented by that cell. This coefficient measures the strength of the linear relationship between two variables. If there is an exact linear relationship between two variables, the correlation coefficient is 1 or –1, depending on whether the variables are positively or negatively correlated. If there is no linear relationship between the two variables, the correlation coefficient tends toward zero.

The JMP software generates a matrix table to show the correlation coefficients between all pairs of variables. This table should be used to screen out explanatory variables with relatively low correlation with the response variable. Two criteria may be used to screen out explanatory variables.

The first screening criterion is based on a threshold for the correlation coefficient between the response variable and the explanatory variable under consideration. A reasonable rule of thumb is to exclude the explanatory variables with a correlation coefficient less than or equal to 0.2. The square of the correlation coefficient defines the proportion of the variability in the response variable that is explained by the explanatory variable. Thus, a correlation coefficient of less than 0.2 would mean that less than 4% of the variability in the response variable would be explained by the explanatory variable and, hence, this explanatory variable could be considered to be of little value in developing a predictive equation for the response variable.

The second screening criterion is based on the concept that if two potential explanatory variables are highly correlated, both variables in the regression equation could create numerical as well interpretation problems. A reasonable rule of thumb is to retain *only* one of the two explanatory variables that have a correlation coefficient of 0.9 or greater with each other.

The results of completing the screening analysis for Example D9-1 are discussed next.

<u>Results of Step 1 for Example D9-1</u>

Table D 9-2 shows the correlation coefficients between all pairs of original/log-transformed response variables and original explanatory variables. Table D 9-3shows similar results using log-transformed explanatory variables. The highest simple correlation is between the runoff coefficient and natural logarithm of hydraulic residence time (r = -0.79). Because the log-transformed explanatory variables generally show higher correlations with the raw runoff coefficient than raw explanatory variables, the subsequent regression analysis was performed using the raw runoff coefficient and the log-transformed explanatory variables.

Figure D 9-1 shows the scatterplot between the runoff coefficient and natural logarithm of hydraulic residence time. The scatterplot suggests a reasonably linear relationship between the two variables with a negative correlation. That is, the runoff coefficient decreases as the hydraulic residence time increases. Similar scatterplots between the runoff coefficient and other explanatory variables should be prepared to examine the nature of the relationship between those pairs of variables.

Results of Step 1 for Example D9-1 (continued)

Using the first screening criterion described above, no explanatory variables were excluded from further analysis because all correlation coefficients (in log scale of explanatory variables) with the raw runoff coefficient were higher than 0.2 (in absolute value).

The second screening criterion was applied next to the retained explanatory variables. The following pairs of explanatory variables showed a correlation coefficient of 0.9 or greater: log of dry density and log of porosity (r = -0.97), log of porosity and log of relative percent compaction (r = -0.93), and log of dry density and log of relative percent compaction (r = -0.90). Because these three variables are highly correlated with each other, only one of them should be retained for further analysis. The log of dry density was retained because, among the three variables, it showed the highest correlation with runoff coefficient.

After applying both screening criteria, the following explanatory variables were retained for further analysis:

- Ln Average Strip Width (m)
- Ln Slope Inclination (%)
- Ln Average Vegetative Cover (%)
- Ln Dry Density (lb/ft^3)
- Ln Infiltration Rate (in/hr)
- Ln Gravel (%)
- Ln Sand (%)
- Ln Silt/Clay (%)
- Ln Estimated Hydraulic Residence Time (min)

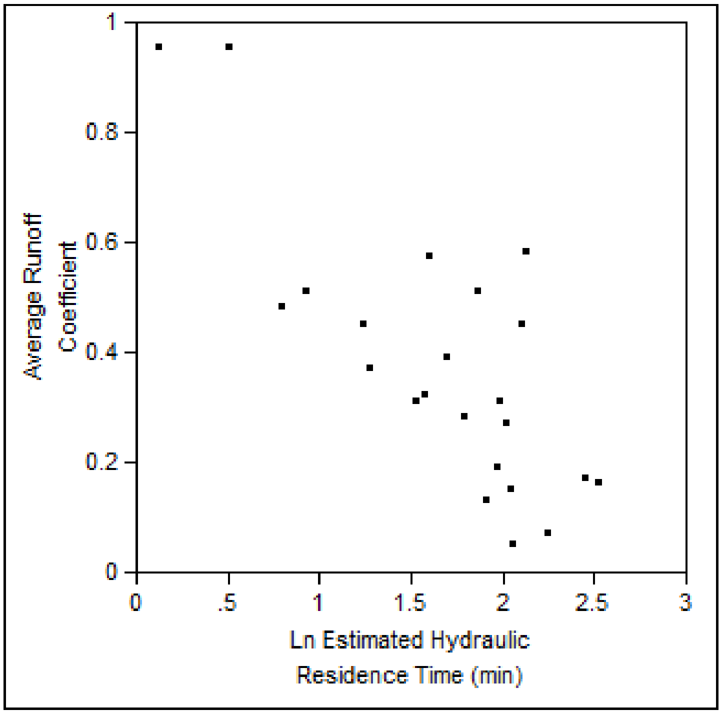


Figure D 9-1 Scatterplot Between Runoff Coefficient and Natural Logarithm of Hydraulic Residence Time in Minutes

Table D 9-2 Correlation Coefficients Between Original/Log-transformed Response Variable and Original Explanatory Variables

Variable	Avg. Runoff Coeff.[1]	Ln Avg. Runoff Coeff.[1]	Avg. Strip Width[2] (m)	Slope[2] (%)	Avg. Veg. Cover[2] (%)	Relative Compaction[2] (%)	Dry Density[2] (lb/ft3)	Infiltration Rate[2] (in/hr)	Porosity[2] (%)	Gravel[2] (%)	Sand[2] (%)	Silt/Clay[2] (%)
Ln Avg. Runoff Coeff.[1]	0.911	1	-	-	-	-	-	-	-	-	-	-
Avg. Strip Width[2] (m)	-0.445	-0.478	1	-	-	-	-	-	-	-	-	-
Slope[2] (%)	-0.340	-0.312	0.328	1	-	-	-	-	-	-	-	-
Avg. Veg. Cover[2]	-0.623	-0.410	-0.065	0.222	1	-	-	-	-	-	-	-
Relative Compaction[2] (%)	0.524	0.536	-0.371	-0.697	-0.440	1	-	-	-	-	-	-
Dry Density[2] (lb/ft3)	0.620	0.633	-0.360	-0.503	-0.454	0.8961	1	-	-	-	-	-
Infiltration Rate[2] (in/hr)	-0.367	-0.282	0.051	0.617	0.472	-0.4033	-0.169	1	-	-	-	-
Porosity[2] (%)	-0.566	-0.557	0.340	0.506	0.504	-0.9407	-0.968	0.176	1	-	-	-
Gravel[2] (%)	-0.230	-0.091	-0.183	0.337	0.561	-0.1427	0.044	0.570	0.036	1	-	-
Sand[2] (%)	0.416	0.297	-0.061	-0.557	-0.599	0.5362	0.373	-0.535	-0.451	-0.808	1	-
Silt/ Clay[2] (%)	-0.264	-0.311	0.398	0.313	-0.0002	-0.596	-0.652	-0.113	0.640	-0.403	-0.214	1
Est. Hydraulic Resid. Time[2] (min)	-0.696	-0.659	0.639	0.098	0.561	-0.4385	-0.535	0.152	0.529	0.130	-0.282	0.223

1 Potential Response Variable
2Potential Explanatory Variable

Table D 9-3 Correlation Coefficients Between Original/Log-transformed Response Variable and Log-transformed Explanatory Variables

Variable	Ln Avg. Runoff Coeff.[1]	Ln Avg. Runoff Coeff.[1]	Ln Avg. Strip Width[2] (m)	Ln Slope[2] (%)	Ln Avg. Veg. Cover[2] (%)	Ln Relative Compaction[2] (%)	Ln Dry Density[2] (lb/ft3)	Ln Infiltration Rate[2] (in/hr)	Ln Porosity[2] (%)	Ln Gravel[2] (%)	Ln Sand[2] (%)	Ln Silt/Clay[2] (%)
Ln Avg. Runoff Coeff.[1]	0.911	1	-	-	-	-	-	-	-	-	-	-
Ln Avg. Strip Width[2] (m)	-0.369	-0.426	1	-	-	-	-	-	-	-	-	-
Ln Slope[2] (%)	-0.239	-0.254	0.559	1	-	-	-	-	-	-	-	-
Ln Avg. Veg. Cover[2]	-0.701	-0.479	0.033	0.065	1	-	-	-	-	-	-	-
Ln Relative Compaction[2] (%)	0.521	0.532	-0.471	-0.722	-0.373	1	-	-	-	-	-	-
Ln Dry Density[2] (lb/ft3)	0.620	0.634	-0.397	-0.530	-0.444	0.897	1	-	-	-	-	-
Ln Infiltration Rate[2] (in/hr)	-0.468	-0.294	-0.037	0.296	0.579	-0.365	-0.179	1	-	-	-	-
Ln Porosity[2] (%)	-0.576	-0.568	0.402	0.543	0.468	-0.930	-0.974	0.208	1	-	-	-
Ln Gravel[2] (%)	-0.217	-0.082	-0.198	0.115	0.452	-0.219	-0.030	0.637	0.111	1	-	-
Ln Sand[2] (%)	0.416	0.292	-0.045	-0.357	-0.473	0.552	0.386	-0.620	-0.439	-0.802	1	-
Ln Silt/ Clay[2] (%)	-0.230	-0.306	0.442	0.481	-0.022	-0.608	-0.677	-0.221	0.639	-0.433	-0.130	1
Ln Est. HRT[2] (min)	-0.788	-0.680	0.447	0.142	0.808	-0.479	-0.572	0.398	0.595	0.267	-0.338	0.112

[1] *Potential Response Variable*
[2] *Potential Explanatory Variable*

D9.2.2 Step 2. Run Stepwise Multiple Regression Analysis and Select a Model for Further Investigation

The stepwise multiple regression analysis is used to select the most efficient set of explanatory variables among those that are retained after performing the screening analysis. In this method, the analysis is performed in a series of steps to decide which explanatory variables should be included in the regression equation. At each step, the method seeks to add a variable to the equation if a certain inclusion criterion is met. The method also seeks to exclude a variable that is already in the equation if a certain exclusion criterion is met. The inclusion criterion relates to the additional amount of variability in the response variable that would be explained by including a new explanatory variable in the equation. Similarly, the exclusion criterion relates to the loss in the amount of variability in the response variable explained if an explanatory variable is excluded. Statistical software packages such as JMP or Minitab have reasonable default values for these criteria. The stepwise regression analysis stops when no new variable could be added and none of the existing variables in the equation could be excluded. The explanatory variables in the final step and the corresponding regression equation are selected for predicting the response variable for given values of the explanatory variables.

<u>Results of Step 2 for Example D9-1</u>

Stepwise regression analysis was performed using JMP software package. The resulting regression is:

Equation D 9-1

$$\text{Runoff coefficient} = -3.06 + 0.809 \times \ln(\text{dry density}) - 0.071 \times \ln(\text{infiltration rate}) - 0.218 \times \ln(\text{hydraulic residence time})$$

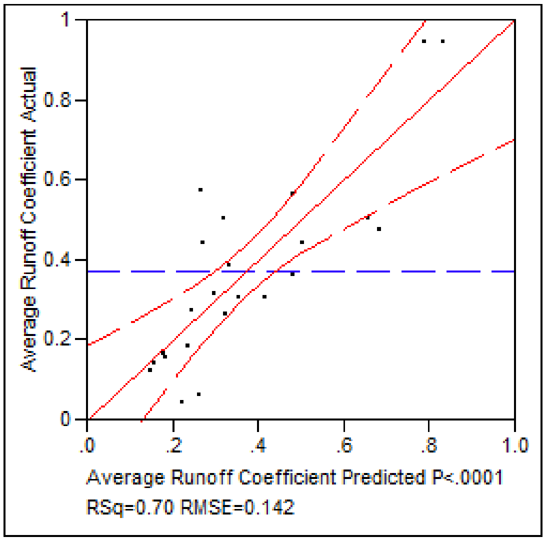
The units for the variables are: runoff coefficient (unitless), dry density (lb/ft^3), infiltration rate (inches/hour), hydraulic residence time (minutes). Figure D 9-2 shows the JMP output for the regression equation. The key outputs, highlighted in green in the figure, are described below:

- <u>Adjusted RSquare (RSquare Adj)</u> is the adjusted square of the coefficient of multiple correlation. It measures the proportion of the total variability in the response variable explained by the regression model, after adjustment for the number of explanatory variables in the model. The higher the RSquare Adj, the better is the fit of the regression model to the data. For the example, the RSquare Adj is ~ 0.65, so the regression mode explains ~ 65 percent of the site-to-site variability in the estimated runoff coefficient.
- <u>Root Mean Square Error (RMSE)</u> is the standard deviation of the random error that can be attributed to the unexplained variability in the runoff coefficients for the study sites. The smaller the RMSE, the smaller is the uncertainty in predicting a value of the response variable for a new combination of the *x* variables. For the example, the RMSE is about 0.14, indicating that the runoff coefficient estimated from the regression equation has an approximate (one standard deviation) accuracy of ±0.14. To assess the reasonableness of a regression model, one should report both RSquare Adj and RMSE. Although RSquare Adj provides an indication of how well the model fits the data, it is not sufficient to assess whether the model is reasonable for predictive purposes. Even when RSquare Adj is relatively high, the RMSE could still be large and the range of plausible predicted values based on this RMSE too wide to be of practical use.

Results of Step 2 for Example D9-1 (continued)

- Prob > F (under "Analysis of Variance") is the significance probability, *p*, for the regression model. It is the probability of getting the relationship between the response variable and the explanatory variables defined by the regression model by chance alone. Significance probabilities < 0.05 are often considered sufficient evidence that the regression model is significant. For the example, the significance probability is < 0.0001, indicating a highly significant regression model. Therefore, this model may be used to estimate the runoff coefficient at different sites as a function of the explanatory variables.
- Parameter Estimates provides information about the estimated regression coefficients. The "Estimate" column shows the means of the regression coefficients. The "Std Error" column shows each estimate's standard error. The "t Ratio" column is the ratio (Estimate/Std Error). Higher *t* ratio values indicate greater significance of the regression coefficient. The "Prob>|t|" column is the significance probability, *p*, for the estimated regression coefficient. Values of $p < 0.05$ are often considered sufficient evidence that the regression coefficient is significant (i.e., it is not zero). The *p* values for the example suggest that the log of hydraulic residence time is highly significant ($p = 0.0033$) and log of dry density ($p = 0.10$) and log of infiltration rate ($p = 0.16$) are moderately significant.

Response Average Runoff Coefficient
Whole Model
Actual by Predicted Plot



Summary of Fit

Parameter	Value
RSquare	0.696738
RSquare Adj	0.648854
Root Mean Square Error	0.142015
Mean of Response	0.374783
Observations (or Sum Wgts)	23

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	0.8803795	0.293460	14.5507
Error	19	0.3831944	0.020168	Prob > F
C. Total	22	1.2635739		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-3.060438	2.284872	-1.34	0.1962
Ln Dry Density (lb/ft3)	0.8091372	0.469107	1.72	0.1008
Ln Infiltration Rate (in/hr)	-0.07108	0.04914	-1.45	0.1643
Ln Estimated Hydraulic Residence Time (min)	-0.218341	0.06499	-3.36	0.0033

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Ln Dry Density (lb/ft3)	1	1	0.06000215	2.9751	0.1008
Ln Infiltration Rate (in/hr)	1	1	0.04219688	2.0923	0.1643
Ln Estimated Hydraulic Residence Time (min)	1	1	0.22763894	11.2871	0.0033

Figure D 9-2 Statistical Details of Regression

D9.2.3 Step 3. Verify Key Assumptions of a Multiple Linear Regression Model

Key assumptions of the multiple regression model are:

- The error terms (i.e., the difference between the actual and predicted values of the response variable) are normally distributed
- The error terms have a constant variance
- The error terms are independent (i.e., random)

These assumptions can be verified by analyzing the residuals, which are defined as the differences between the observed and predicted values of *y*. Statistics software packages such as JMP and Minitab provide the option of calculating and saving the residuals. These residuals can be analyzed using the graphical and numerical methods of data review described in Appendix D2.

Verify Error Terms Are Normally Distributed: The assumption of normal distribution should be verified with a normal probability plot as described in Appendix D2.

Verify Error Terms Have a Constant Variance: A plot of residuals versus the predicted values of the response variable should be prepared to assess whether the scatter of the residuals around the zero-residual line is relatively constant and does not show an increasing or decreasing trend.

Verify Error Terms Are Independent: If the response variable has been observed sequentially over time, one can plot residuals against time, even though time has not been explicitly incorporated as an explanatory variable in the model. If this plot shows a random pattern, the error terms can be considered to be independent. Conversely, if the time plot shows a systematic pattern (e.g., increasing residuals with time), this would suggest that the error terms are not independent over time and are, in fact, correlated.

If error terms are not normally distributed or do not have a constant variance, data transformation can often help in correcting both conditions. If the plot of residuals versus predicted values of the response variable shows that the variability of the residuals increases with increasing predicted values of the response variable, helpful transformations in this case are the logarithmic and square root transformations of the response variable. If any remedial measures are necessary, Steps 1 and 2 should be repeated after applying the remedial measures to the data.

Results of Step 3 for Example D9-1

Figure D 9-3 shows a normal probability plot of the residuals for the regression model developed in Step 2. The plot appears to be linear and data are fitted relatively well on a straight line; hence, the residuals can be assumed to be normally distributed.

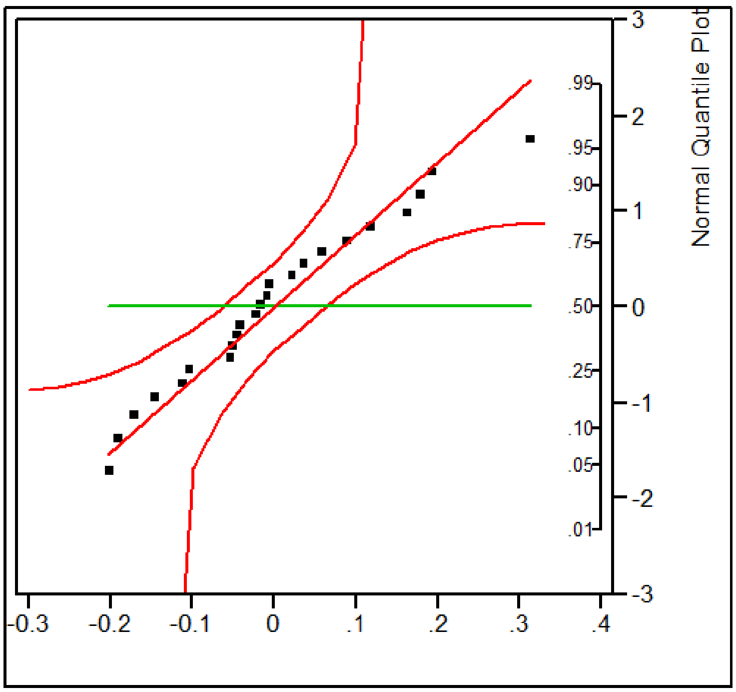


Figure D 9-3 Normal Probability Plot of Residuals for the Regression Model

Figure D 9-3 shows a plot of residuals against predicted values of the response variable. The data appear to be scattered around the zero-residual line with no specified pattern. Hence, the residuals can be assumed to have a constant variance.

Since the data for this example are not obtained in a time sequence, the check on time independence is not relevant. One can assume that the study sites for the example were selected to be sufficiently apart from each other such that they can be considered to be spatially independent. That is, if the runoff coefficient at one site is above its regression-predicted value, the runoff coefficient at any other site is not more likely to also be above its regression-predicted value.

Since the key assumptions of the regression model are verified for the example, one will assume that the regression model defined in Equation D 9-1 is valid to predict runoff coefficients at different sites, provided the data on the explanatory variables included in the regression equation are collected.

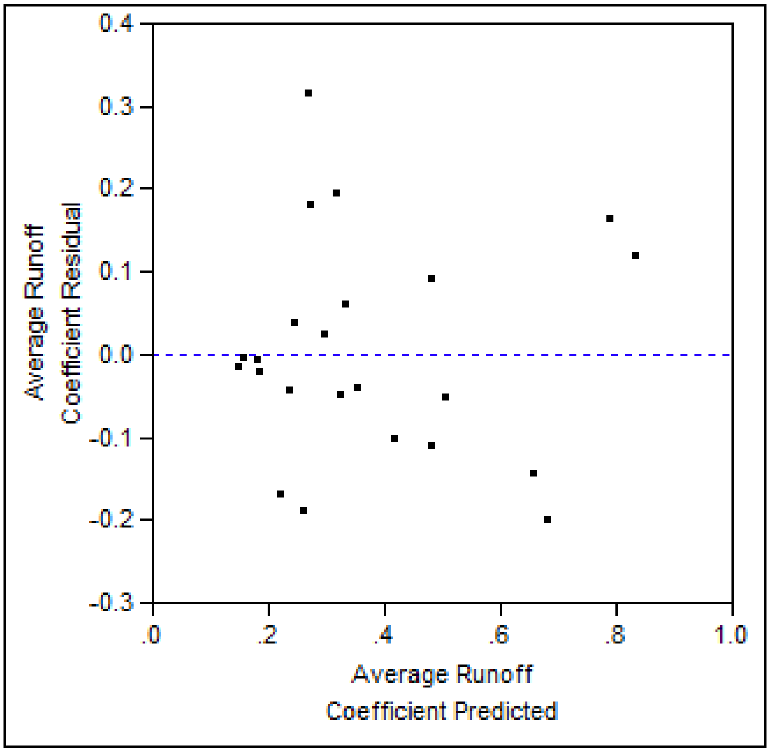


Figure D 9-4 Residuals Against Predicted Values of the Response Variable

D9.2.4 Step 4. Use the Selected Regression Model to Predict the Response Variable for Specified Values of the Explanatory Variables

Once an acceptable regression model is developed, the interest is to use the model to predict the response variable for specified values of the explanatory variables and to assess the standard deviation of the response variable at the specified explanatory variables. The predicted value and the standard deviation are then used to calculate a prediction interval on the response variable with a specified confidence. A prediction interval with a confidence of (1-α) × 100 percent confidence defines an interval that will contain the response variable for given explanatory variables (1-α) × 100 percent of the time.

The predicted value of the response variable, $\hat{y}_x$, for a given vector of *x* values is calculated using the regression equation in which the estimated regression coefficients are used. Thus, the following equation is generated:

Equation D 9-2

$$\hat{y}_x = b_0 + b_1 x_1 + b_2 x_2 + ... + b_n x_n$$

Because the regression equation is only an empirical relationship, the use of the equation to predict the response variable is valid only within the range of explanatory variables in the sample data. Caution should be exercised in making predictions of the response variable beyond the range of sample data (i.e., extrapolating results beyond the range of observations).

Statistical software packages provide options to calculate and save the predicted values, standard deviations, and confidence intervals of y_x for different sets of values (x_1, x_2, …, x_n).

The standard deviation, $s(y_x)$ of y_x, provided in statistical software may be considered an exact estimate because it accounts for all uncertainties. However, a reasonable approximation to the estimate of $s(y_x)$, which is relatively simple to calculate, is:

Equation D 9-3

$$s(y_x) \approx \text{RMSE} \times \sqrt{1 + \frac{1}{n}}$$

In this equation, n is the sample size and RMSE is (as previously defined) the root mean square error of the regression equation.

For a large sample size, $s(y_x)$ approaches RMSE and the probability distribution of y_x approaches the normal distribution.

A useful measure of the reliability of the regression equation to predict the response variable is the RMSE. If different regression equations are being considered for a response variable expressed on the same scale, one can compare the reliability of these equations in terms of their RMSE. An equation with lower RMSE should be preferred.

Results of Step 4 for Example D9-1

Suppose that one wants to predict the runoff coefficient at a new site that has the following values of the explanatory variables:

- Dry density = 110 lb/ft3
- Infiltration rate = 2.5 inches/hour
- Hydraulic residence time = 8 minutes

Using Equation D 9-1 , the runoff coefficient is calculated to be 0.224. The RMSE for the regression equation is 0.14. Therefore, an approximate estimate of the standard deviation of the runoff coefficient at this site is 0.14 × square root of (1 + 1/23) = 0.143, using Equation D 9-3. Assuming normal distribution, the true runoff coefficient at this site would be within the interval (0.224 ± 0.143) about 68 percent of the time.

D9.3 Use of Regression Analysis to Predict Effluent Concentration

Regression analysis may be used to compare the performance of different BMPs in a pilot study based on predicted effluent quality. The influent quality is likely to be variable at the individual BMP sites and the effluent quality for many constituents depends on the influent quality. The steps described in the previous section may be used to develop a regression equation to predict the effluent event mean concentration (EMC) as a function of the influent EMC. A reference influent EMC is defined and used to predict the effluent EMC using the regression equation at each BMP site. Procedures described in Appendix D6 can be used to test the hypothesis that the mean effluent EMC is the same at two different BMP sites.

Such an approach was used, for example, in the Caltrans BMP Retrofit Pilot Program (Caltrans 2004). Care should be taken to ensure that the key assumptions of a linear regression model, described in Step 3, are satisfied and the RMSE of the regression equation is small enough so that the predicted range of effluent concentration for a given BMP is not too wide. When the RMSE is large, important differences between BMPs may be masked by the uncertainty in the predicted effluent concentrations; that is, the ranges of predicted effluent concentrations at different BMPs may overlap.

Appendix D10 How to Evaluate Time Trends in BMP Monitoring Data

D10.1 Purpose and Organization

The purpose of this appendix is to present statistical methods to evaluate whether BMP data show significant trends over time. The methods can address a typical question of interest in BMP studies; namely, is the effectiveness of a particular BMP expressed (e.g., in terms of percent removal of a pollutant) changing over time? The focus of the methods is on evaluating the presence of a monotonic trend in a time series of sample data. For evaluating other temporal trends, such as seasonal effects, the methods of hypothesis testing described in Appendices D6 and D8 should be used.

For BMP studies, data would be collected during each storm event that meets certain threshold conditions in each of several monitoring years. Interest will be in evaluating two types of time trends:

- Intra-year trend
- Inter-year trend

For the intra-year trend, one evaluates whether or not a monitoring parameter of interest (e.g., the concentration of a certain pollutant) shows a systematic trend over different storm events in a given year. For example, one would like to assess whether the effects of the storms on pollutant concentration are different during the initial storms in a wet season, due to flushing of materials accumulated before the start of a wet season.

For the inter-year trend, one evaluates whether or not a monitoring parameter shows a systematic trend over different monitoring years? For example, one would like to assess whether BMP effectiveness is decreasing over years.

This appendix presents methods to evaluate both types of trends. A common example will be used to illustrate the application of these methods to BMP studies. The data for this example are presented first. Methods for the two types of trend analysis are described next and illustrated using the example data.

Example D10-1

Consider a BMP study that used the influent-effluent monitoring approach to collect data on TSS loads in the influent and effluent at a BMP structure over a period of four years. The influent and effluent TSS loads for each sampling event were used to calculate the percent reduction in the TSS load. Table D 10-1 shows a summary of the data on the percent reduction in TSS load for each storm event of each monitoring year.

The questions of interest for this BMP study are:

- Does the percent TSS load reduction show a systematic variation over different storm events within each monitoring year? For example, is the percent TSS load reduction different for the initial storm events of the rainy season? If the initial events consistently show a lower percent reduction in a pollutant, this might suggest the need for a supplementary BMP maintenance practice during the start of a rainy season that would increase the effectiveness of the program.
- Does the percent TSS load reduction show a trend over the monitoring years? Is the BMP effectiveness changing over time? If the analysis shows a decreasing trend in BMP effectiveness, supplementary BMP actions might be necessary.

Table D 10-1 Percent Reduction in TSS Load Data for Example D10-1

Event ID	Year 1	Year 2	Year 3	Year 4
1	27.5	20.1	9.8	13.7
2	25.4	21.4	17	8.6
3	21.2	21.7	12.7	10
4	24.7	17.8	18.8	10.2
5	28.2	17.4	14.2	8.8
6	20.9	24.9	13.8	9.3
7	29	28.9	11.7	6.4
8	22.7	19.4	14.6	9.4
9	-	25.1	12.3	10.9
10	-	22.8	-	7.9
11	-	12.4	-	6.7
12	-	-	-	10.7
Annual Percent Reduction in TSS Load	**25.5**	**20.4**	**14.1**	**8.6**

D10.2 Evaluation of Intra-Year Trends

Both a graphical method and a formal statistical test (Mann-Kendall test) are recommended for the evaluation of intra-year trends.

D10.2.1 <u>Graphical Method</u>

A time series plot is prepared that plots the monitoring parameter as a function of the time of the sampling event for each monitoring year. Such plots can be prepared using Excel graph functions. The time pattern within each monitoring year is examined visually to assess whether there are any consistent monotonic (upward or downward) trends. If the plot does not show a monotonic trend, but does show two or more distinctly different time periods (e.g., the two halves of a wet season) with regard to observed values of the monitoring parameter, the methods of hypothesis testing described in Appendix D6 or D8 should be used to evaluate whether the differences between the time periods are statistically significant.

D10.2.2 <u>Mann-Kendall Test</u>

The Mann-Kendall (M-K) test is described in detail in Gilbert (1987) and also in the USEPA Guidance Document (2006). It is used to evaluate monotonic trends. This is a nonparametric test that does not assume any particular probability distribution for the sample data. Also, non-detect values can be incorporated by assigning them a common value that is smaller than the lowest detected value. This approach is valid because the M-K test uses only the relative magnitudes of the data rather than their actual values. The recommended minimum sample size (i.e., the annual number of storm events for the present analysis) for the M-K test is eight. Statistical software packages, such as JMP, do not include the M-K test. However, the following programs/software are available to run the M-K test:

- The FORTRAN program code provided in Gilbert's book
- Visual Sample Plan (version 4) software, developed by the U.S. Department of Energy, provides a module to run the M-K test (Hassig et al. 2005)
- Kendall.exe (a DOS-executable file, USGS 2005)

Also, one can set up an Excel spreadsheet to calculate the *S* statistics (described below) and determine the M-K test result.

The M-K test follows the standard principles of hypothesis testing that were described in Appendix D6. The null hypothesis indicates there is no trend. The alternative hypothesis could be either two-tailed (i.e., there is an upward or downward trend) or one-tailed (i.e., there is a trend only in one direction, either upward or downward). The data are listed in the time order and the signs (positive or negative) of differences between a measured concentration versus all other measured concentrations before it are found. The signs are used to calculate a statistic denoted by *S*, which is the number of positive differences minus the number of negative differences. If *S* is a large positive number, measurements taken later in time tend to be larger than those taken earlier (i.e., there is an increasing trend). Conversely, if *S* is a large negative number, measurements taken later in time tend to be smaller (i.e., there is a decreasing trend).

The probability of getting a statistic greater than the absolute value of *S* by chance alone is calculated. This is the significance probability, *p*, of the test. If *p* is less than a specified threshold (typically 0.05), this is taken as evidence of a significant monotonic trend. The sign of the statistics *S* and the time series plot will inform the user whether there is an increasing or decreasing trend. If *p* is greater than, or equal to, the specified threshold, one concludes that the data do not provide sufficient evidence of a monotonic trend. The *p*-value can be determined by using Table A-12b in the USEPA's Data Quality Assessment: Statistical Methods for Practitioners (USEPA 2006a), in which the *p*-value depends on sample size, *n*, and the Mann-Kendall statistic, *S*.

Results of Intra-Year Trend Analysis for Example D10-1

Figure D 10-1 shows a time series plot of the data listed in Table D 10-1. The figure reveals no particular trend in the percent reduction in TSS loads for any of the four monitoring years. The M-K test was applied using an Excel setup. The results are shown in Table D 10-2. The significance probabilities, *p*, are 0.45, 0.44, 0.46, and 0.27 for Years 1, 2, 3, and 4, respectively, which support the conclusion of no trend drawn from the time series plot.

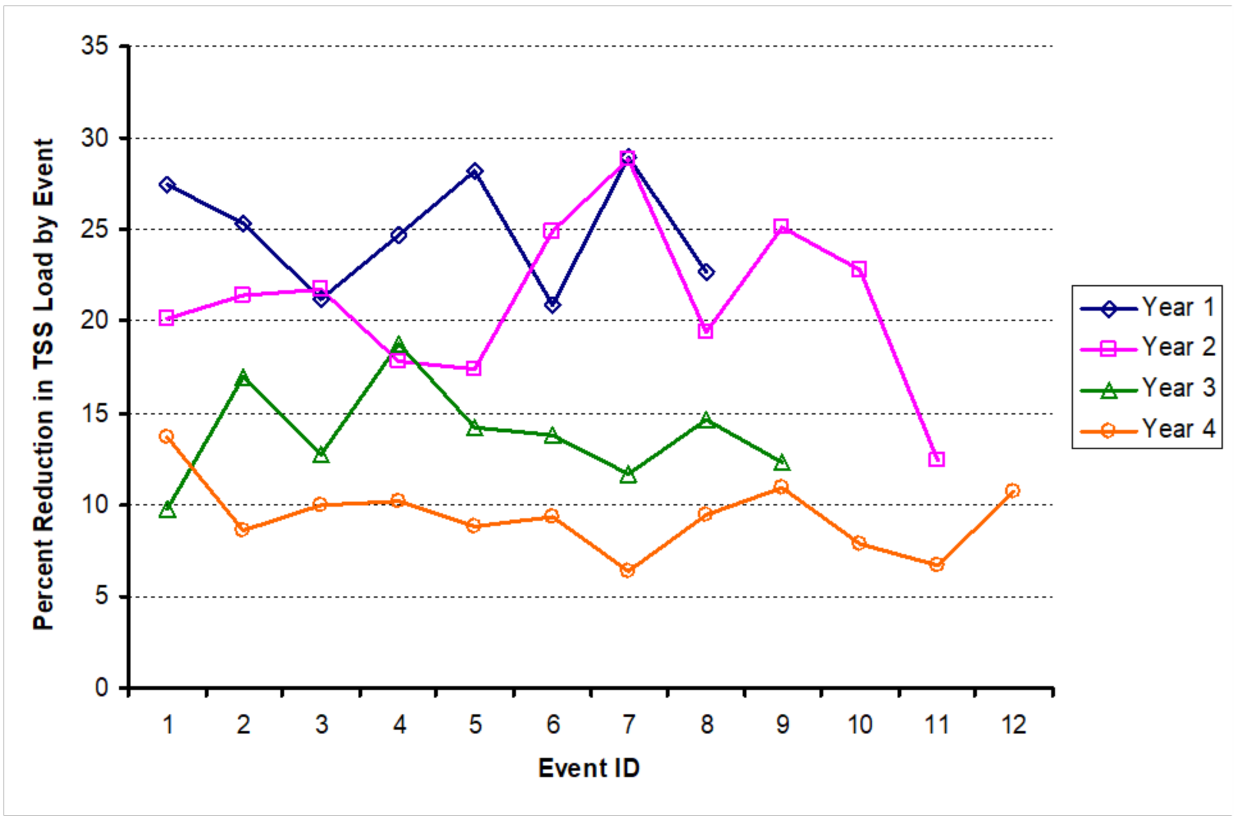


Figure D 10-1 Example D10-1 Time Series Plot of Percent Reduction in TSS Loads

Table D 10-2 Intra-Year Mann-Kendall Test Results for Example D10-1

Year	Number of Sampling Periods	Number of Positive Differences	Number of Negative Differences	S	M-K Test p-value	Trend Result
Year 1	8	13	15	-2	0.452	No Trend
Year 2	11	29	26	3	0.44	No Trend
Year 3	9	17	19	-2	0.46	No Trend
Year 4	12	28	38	-10	0.273	No Trend

D10.3 Evaluation of Inter-Year Trends

The question of interest for this analysis is whether the BMP monitoring parameters show any trends over different monitoring years. Again, both a graphical method and a formal statistical test (M-K test) are recommended for the evaluation of inter-year trends.

D10.3.1 Graphical Method

A time series plot is prepared that plots the monitoring parameter versus the monitoring year. Excel graphing functions can be used to prepare such plots. The time pattern over the monitoring years is examined visually to assess whether the plot exhibits a monotonic (increasing or decreasing) trend over the monitoring period. If there is no trend, the data points will plot randomly on either side of the overall mean value. If a trend exists, the data points will show a systematic pattern of increasing or decreasing values.

D10.3.2 Mann-Kendall Test

This test is applied as described in the previous section. As was noted previously, the recommended minimum sample size for the M-K test is eight. However, the number of monitoring years for typical BMP studies is only two to four years. If the sample size is less than four, the M-K test should not be applied. If the sample size is between four and seven, the M-K test may be performed, but its results should be used with caution and always supported by a visual inspection of the time series plots.

As noted previously, the key result of the M-K test is the significance probability, p. If p is less than a specified threshold (typically 0.05), this is taken as evidence of a significant trend. If p is greater than or equal to the threshold, one concludes that the data do not provide sufficient evidence of a trend.

<u>Results of Inter-Year Trend Analysis for Example D10-1</u>

Figure D 10-2 shows a time series plot of the annual percent reduction in TSS load. The plot suggests that there is a decreasing trend in the percent reduction in the TSS load over the four years of the BMP monitoring program. This result might be an indication of decreasing BMP performance over the four monitoring years. The M-K test was performed next. The significance probability, *p*, for the test is 0.042, suggesting the presence of a significant, decreasing trend. Because the sample size is only four, the results of the M-K test should be used with caution, primarily to reinforce the conclusion drawn from the time series plot.

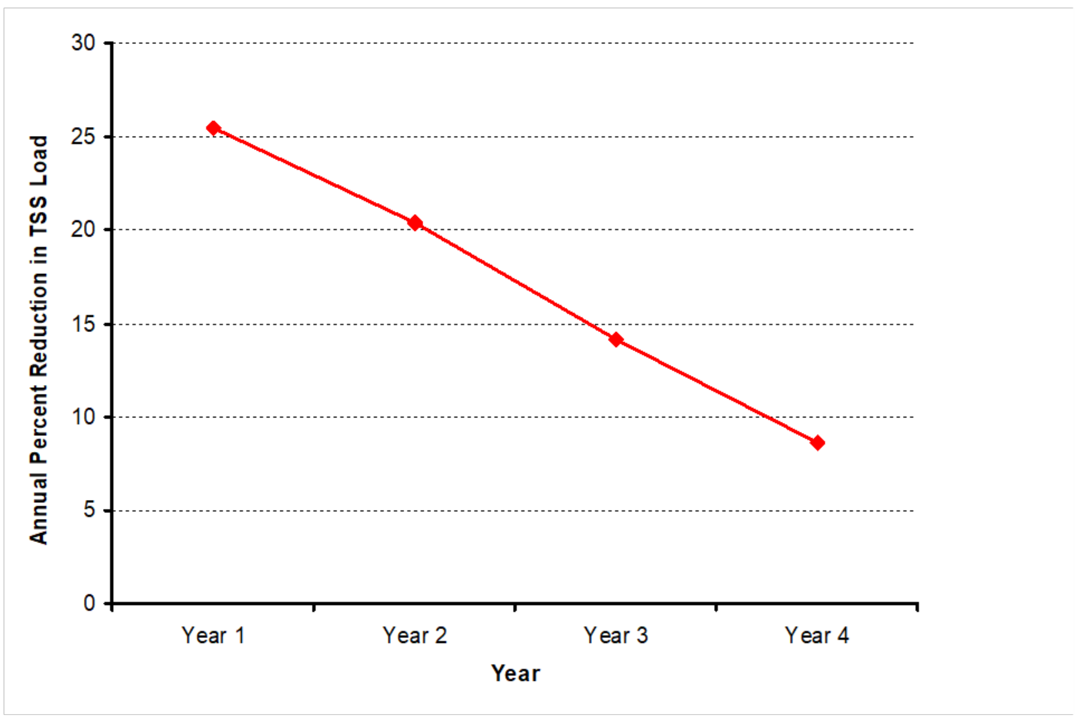


Figure D 10-2 Time Series Plot of Annual Percent Reduction in TSS Loads for Example D10-1

Table D 10-3 Inter-Year Mann-Kendall Test Results for Example D10-1

Number of Sampling Periods	Number of Positive Differences	Number of Negative Differences	S	M-K Test p-value	Trend Result
4	0	6	-6	0.042	Downward

Appendix D11 How to Compare and Rank BMPs Monitored at Multiple Sites

D11.1 Purpose and Organization

The purpose of this appendix is to present statistical methods to evaluate whether the treatment performance of BMPs is significantly different when each BMP type has been monitored at multiple locations. Comparisons between different BMPs, and even BMPs of the same type, can be difficult when influent-effluent data are in quite different (sometimes nonoverlapping) ranges due to site-to-site variability (Equation D 11-1). The performance calculated is strongly affected by the influent concentration at a particular site and can make devices evaluated at locations with low influent concentrations appear to perform less effectively in comparison to those located at sites with higher influent concentrations. One approach that can be used in this situation is regression analysis of influent and effluent concentrations to predict performance at all the devices based on a common influent concentration. This approach, described in Appendix D9.3, may be suitable if the root mean square error (RMSE) of the regression equation is small enough that the predicted range of effluent concentration for a given BMP is not too wide. When the RMSE is large, important differences between BMPs may be masked by the uncertainty in the predicted effluent concentrations; that is, the ranges of predicted effluent concentrations for different BMPs may overlap. The approach is not suitable when multiple sites have been monitored for each BMP type. Finding systematic site (or other) effects requires use of a mixed model. The mixed model methods presented here can address a typical question of interest in BMP studies; namely, is the effectiveness of a particular type of BMP (e.g., in terms of percent concentration or load reduction) different than that of another BMP that was monitored under very different conditions? The focus of the methods is on evaluating differences in treatment performance with the objective of grouping BMPs that do not show statistically significant differences and ranking these groups. These methods can be used to create the performance tiers described in Appendix L.

Mixed models add at least one random variable to an underlying model, such as a linear model (Appendix D9). The random variables of a mixed model add the assumption that observations for a given BMP type are correlated, i.e., observations from one location are correlated with observations from another location with the same BMP type. Mixed models are designed to address this correlation and do not cause a violation of the independence of observations assumption for the underlying model. There are two types of effects in a mixed model. A fixed effect is one for which one wants to draw conclusions only about the groups under consideration and would be used again if the monitoring was performed again. The type of BMP is a fixed effect. We may, say, be interested in determining the effect select BMP types have on a dependent/response variable (e.g., percent concentration reduction). Soil type is another example of a fixed effect. We may be interested in determining the effect different soils types have on, say, percent load reduction. Random effects are variables whose values are not of primary interest, but rather are thought of as a random selection from a much larger population. Site is a random effect since we have no interest in particular sites, but only how BMPs would compare within an arbitrary (i.e., hypothetical) population of sites that might have been used. A mixed model is one that contains at least one fixed effect and at least one random effect. To consider the possibility of site-to-site differences, one must use a mixed model such as a linear mixed model.

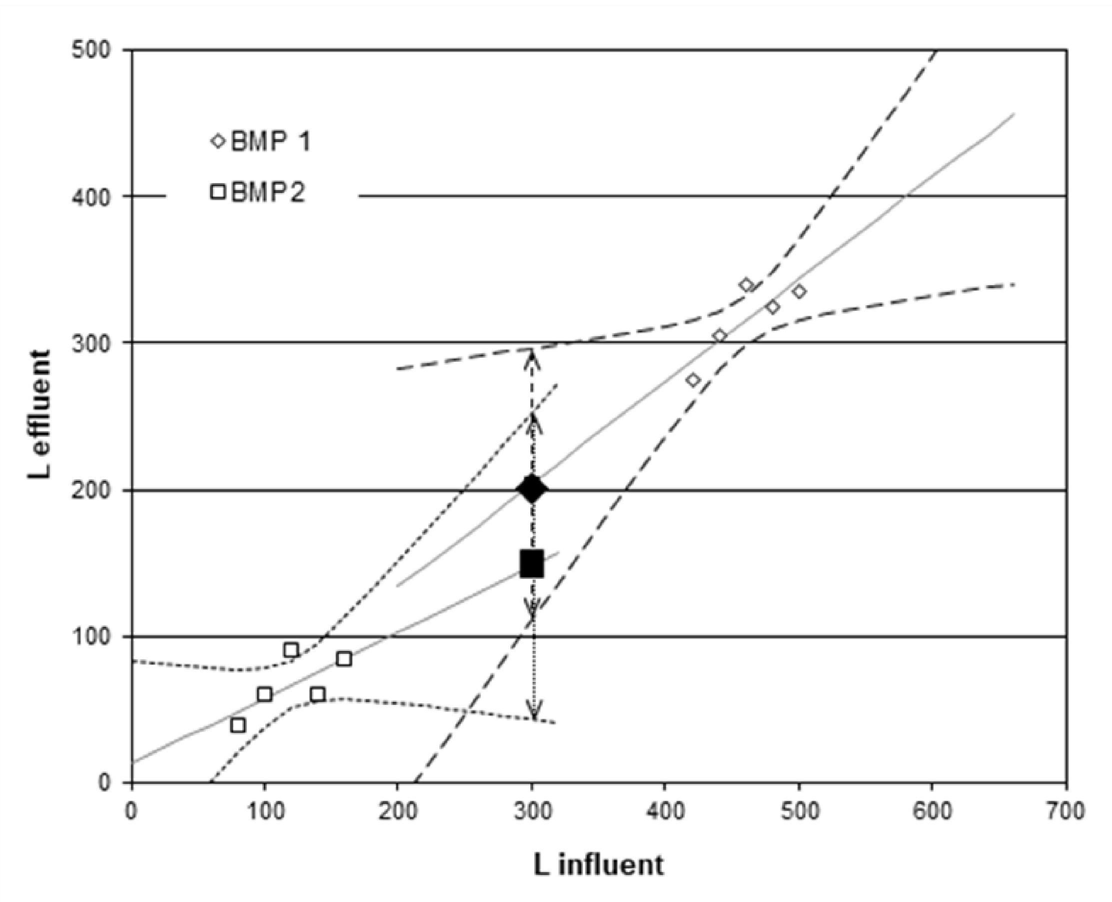


Figure D 11-1 Comparison of Two Hypothetical Nonoverlapping BMP Regressions

D11.2 Steps to Perform Linear Mixed Model Analysis

The following steps are performed to conduct this type of analysis. The subsections below describe each step.

- Step 1.. Verify Assumptions for Linear Models
- Step 2. Perform Mixed Model Analysis
- Step 3. Rank BMP Groups

D11.2.1 Step 1. Verify Assumptions for Linear Models

Linear mixed model analysis has the same assumptions as linear regression (Appendix D9.2.3), namely

- The error terms (i.e., the difference between the actual and predicted values of the response variable) are normally distributed.
- The error terms have a constant variance.
- The error terms are independent (i.e., random).

Consider the need for data transformation if scatter plots of the raw data indicate non-constant variability of the effluent measurements as a function of the influent (error terms are not normally distributed). It is, of course, important to verify the regression model's fit to the transformed data. Although the mixed model will address correlation within BMP types due to site as a random variable, other factors that may need to be considered include storm date and event size. For example, correlation when data for multiple BMPs are collected during the same storms (the influent and effluent values in the same storm are correlated). It is important to realize that in the absence of replication (i.e. only one site is monitored) a mixed model will assume that there is no site-to-site variability within a given BMP, which may not be a reasonable assumption.

D11.2.2 Step 2. Perform Mixed Model Analysis

Perform a mixed-model statistical analysis of BMP types to identify BMP groups where the likelihood of similar performance is greater than 0.1 (a confidence of 90 percent is used as a test of difference so that BMPs tend to stay in larger, overlapping groups when data are scarce or highly variable).

The relationships between influent and effluent concentrations or between influent and effluent mass loadings are modeled using mixed model analysis of covariance methods. Unlike the linear regression approaches outlined in Appendix D9.3, this approach treats the collections of BMPs of a given type as a random sample from a population of BMPs of that type, and its conclusions are applicable to that population, rather than merely to the particular BMPs for which data were collected. This is a more relevant question, since one wishes to learn not only how existing BMPs are performing, but also about the possible consequences of choosing a particular BMP type where no BMP has been installed. This is precisely what the mixed model analysis does, by treating the individual BMP (site) as a random effect rather than a fixed effect.

For a BMP of a type t, the model fit to the data is:

Equation D 11-1

$$Y_{tij} = [b_{0t} + (b_{0ti} – b_{0t})] + [b_{1t} + (b_{1ti} – b_{1t})] X_{tij},$$

In Equation D 11-1, Y_{tij} and X_{tij} represent the paired effluent and influent values for a BMP of type t, i indicates which individual BMP (site) the observation came from, and the data are the j-th values from that individual BMP.

This model assumes that the relationship between the influent and effluent values is linear, but that each individual BMP can have a somewhat different linear relationship. The significance of the main effect of BMP type, or of differences in slopes between BMP types, is conducted using the standard mixed model method.

Differences among BMP types are evaluated using least squares means. This method compares types at a common influent level, so that in comparing types t and s, the hypothesis being tested is:

Equation D 11-2

$$H_0: b_{0t} + b_{1t}\, \mu_X = b_{0s} + b_{1s}\, \mu_X,$$

In Equation D 11-2, μ_X is the arithmetic mean of the influent values.

The influent and effluent values are either analyzed as their raw (untransformed) values or following log transformations, and those values can be either concentrations or mass loadings, calculated as the product of the measured concentration and the measured flow (influent or effluent). The use of log-transformed values has two consequences. First, whenever the flow or concentration is zero, no log transformation can be calculated, and so data adjustments are necessary as described below. Second, when the analysis is based on log-transformed values, then the least squares means are in fact comparisons conducted at the geometric mean of the influent values. The log-transformed analysis is preferred for two reasons. First, log-transformed data come closer to satisfying the model assumptions of normality and constant residual variance. Second, whereas an arithmetic mean influent value can be unreasonably impacted by a few outliers (extremely high values) among the influent values, the impact of outliers will be lessened when the values are log transformed.

To avoid log-transformation of zero values for flow, a numerical value can be inserted in the analysis of mass loadings, defined as F_{min} and C_{med} (and then log transformed), where F_{min} is the minimum detectable non-zero flow, and C_{med} is the median concentration for the constituent. This is done to avoid combining a negligible flow with a negligible concentration. When a concentration falls below the detection limit, the analysis (log or raw) can be based on inserting the detection threshold itself for that value, and the corresponding mass load estimate is the product of the observed flow and the detection threshold.

In JMP, a mixed model analysis can be conducted in the Fit Model platform of JMP Pro. The Fit Model platform's Mixed Model option fits a wide variety of linear models for continuous responses with complex covariance structures. These models include random coefficients, repeated measures (replication), spatial data (sites), and data with multiple correlated responses.

D11.2.3 Step 3. Rank BMP groups

A BMP may be statistically comparable to other BMPs. Groups of statistically equivalent BMPs can be constructed by considering each BMP individually and identifying all the other BMPs that are equivalent. These groups can then be collapsed into groups where all BMPs in the group are equivalent to all other BMPs in the group. Methods to rank BMPs are provided in Appendix L.

Appendix E Contents

1.0 Introduction

- State that the document serves as a study plan for evaluating a BMP
- State which type of BMP will be studied and describe what is already known about the BMP
- State the particular features that will be assessed
- State whether the BMP is being studied for technical feasibility, performance, operation and maintenance, and/or life cycle costs
- State the years the BMP will be studied and the locations

2.0 Problem Statement and Study Goals

- Write the problem statement developed in Step 1
- State the study goal(s) developed in Step 1

3.0 Study Question and Objectives

- Describe the BMP as developed in Step 2
- List the relevant evaluation criteria selected in Step 2
- Identify the relevant study variables from Step 2
- List the Caltrans standards applicable to the criteria, as determined in Step 3
- List the study questions developed in Step 3
- List the study objectives finalized from Step 6
- Discuss key assumptions that have the potential to influence the study outcome

4.0 Study Methodology and Analytical Approach

- Specify the study type identified in Step 4
- Describe how select variables will be controlled
- Describe the variables to be monitored
- Describe the data collection methods and analytical procedures
- Identify the number of sites to be studied and the number of samples required per site
- Specify the statistical analysis techniques to be used to analyze the data

5.0 Schedule and Cost

- Provide a general timeline for the pilot study
- Provide a cost estimate.

6.0 Constraints and Optimization

- State project constraints
- List trade-offs due to optimization

7.0 Reporting Requirements

- List reports that will be required for the pilot study. Examples include:
    - Site Selection Technical Memorandum
    - Design report
    - Design lessons learned
    - Construction report
    - Construction lessons learned
    - As-Built plans
    - OM&M plan
    - Vector monitoring and abatement
    - Cost summary and analysis
    - Data evaluation
    - Post-Storm Technical Memorandum
    - Operational monitoring reports and inspection forms
    - Monitoring data
    - Data validation results
    - O&M lessons learned
    - Electronic data submittal
    - Caltrans SWIS
    - ASCE International Stormwater BMP Database
    - Other

8.0 References

- List references as appropriate.

Open Graded – Rubberized Asphalt Paving
SITE EVALUATION FORM
OVERLAY SITE

Date 5/1/06 Reviewer Mark Van Dam
District 10 County San Joaquin
City ______ Route/Post Mile: 5 North Bound / PM 4.5
Site number (per map): ______
Address ______
Resident Engineer: ______ Telephone ______
Contract #: 100M2704
Photograph#'s: ______

DESCRIBE SITE

Describe: Frontage road paralleling freeway provides access to highway shoulder. Will require curbing to gather water.

REPRESENTATIVENESS

Site Surroundings and land uses: ☐ Urban ☒ Rural
Describe: farm land

Approximate Tributary Area (acres): 0.25 acres / depends on curb added

Type of Overlay (Open, Rubber-Open, Rubber-Gap): Rubber - Open

Width of shoulder (w/o overlay) 12 ft asphalt / 40 ft dirt

Date overlay completed: fall 2005

Upstream Site Topography (flat, steep, etc.): flat

Downstream Conveyance or Receiving Water Type: (e.g., storm drain, open channel, street, pond): drain inlet

Is the drainage area 100 percent representative of site type? ☒ yes ☐ no
Describe: ______

Annual ADT ______________ ☒ NA

Illegal dumping	☐ yes	☒ no
Illicit connections	☐ yes	☒ no
High groundwater table	☐ yes	☒ no
Erosion	☐ yes	☒ no

Describe: ______________________________

Runoff from landscaped areas	☐ yes	☒ no
Is site representative of Caltrans Highways and/or Facilities?	☒ yes	☐ no

Explain off roadway

Are there exposed materials that warrant additional analyses beyond the standard suite defined in the Caltrans Guidance Manual: ☐ yes ☒ no

If so, then list:

Other observations:
will require curbing to gather / concentrate overland flow from highway.

Overall Representativeness Rating (0 to 10): 9

PERSONNEL SAFETY

Personnel safety issues? ☐ yes ☒ no
(e.g., Proximity to traffic lanes, steep embankments traffic, toxic gases, explosives, slips/trips/falls, fast-moving water, confined spaces, etc.)

*Describe:*______________________________

Overall Personnel Safety Rating (0 to 10): 10

SITE ACCESS

Vehicular site access? ☒ yes ☐ no
Describe: from frontage road, or shoulder

Are there maintenance pull outs or frontage/access roads? ☒ yes ☐ no
Describe: frontage Rd. Bevis St.

Continuous access during storms? ☒ yes ☐ no
*Describe:*______________________________

Site remote and/or travel to site subject to traffic delays? ☒ yes ☐ no
Describe: approx. 90 min from SAC

Site subject to flooding? ☐ yes ☒ no
*Describe:*______________________________

Will monitoring activities interfere with Caltrans/motorists ☐ yes ☒ no
*Describe:*______________________________

Does private property have to be crossed? ☐ yes ☒ no
*Describe:*______________________________

Overall Site Access Rating (0 to 10): 9

EQUIPMENT SECURITY

Is the site subject to vandalism? ☒ yes ☐ no
Describe: site is remote

Is the site subject to other damage (e.g., errant drivers)? ☒ yes ☐ no
Describe: depending on location

Overall Equipment Security Rating (0 to 10): 8

FLOW MEASUREMENT CAPABILITY

Type of runoff from site: ☒ Curb and gutter ☒ Overland flow ☐ Other

Describe: need curb/gutter to concentrate overland flow

Room for Primary Measurement (flume): yes

Available area for flow measurement: 40' wide dirt shoulder

Potential sampling location (with access to flow):

☐ Storm drain inlet ☐ Ditch, swale ☐ Culvert ☐ Pipe

☒ Other (describe) flume / pipe downstream of flume

Comments ____________________

Will the sampling location be unaffected by site activities? ☒ yes ☐ no

Describe: ____________________

Will the sampling point be unaffected by upstream BMPs? ☒ yes ☐ no

Describe: ____________________

Is the sampling location subject to backwater conditions? ☐ yes ☒ no

Describe: ____________________

Will uniform flow be established? ☒ yes ☐ no

Describe: if enough curb installed

Will the depth of flow be at least ¾ inches? ☒ yes ☐ no

Describe: if enough curb installed

Is flow measurement feasible? ☒ yes ☐ no

Describe: ____________________

Will flow measurement be accurate? ☒ yes ☐ no

Describe: ____________________

Can the sampler be placed within 99 feet from the sampling point? ☒ yes ☐ no

Describe: ____________________

Will there be less than 26 feet of vertical lift? ☒ yes ☐ no

Describe: ____________________

Is there room for a 6' x 6' concrete pad? ☒ yes ☐ no

Describe: ____________________

Can the site be easily instrumented? ☒ yes ☐ no

*Describe:*________________________________

Can a primary flow measurement device be easily retrofitted? ☒ yes ☐ no

*Describe:*________________________________

Is grab sampling possible? ☒ yes ☐ no

*Describe:*________________________________

Is automated sampling possible? ☒ yes ☐ no

*Describe:*________________________________

What type of sampler is most suitable? ☒ portable ☐ refrigerated

What type of flow meter is most suitable?

☒ Bubbler ☐ Area/Velocity

☐ Ultrasonic ☒ Combination/Other primary - flume

Overall Flow Measurement Capability Rating (0 to 10): 9

ELECTRICAL POWER AND TELEPHONE ACCESS

Electrical Power Available? ☐ yes ☒ no

Telephone lines available? ☐ yes ☒ no

Clear cellular phone reception at site? ☒ yes ☐ no

Overall Electrical Power and Telephone Access Rating (0 to 10): 8

CALTRANS SOURCES

Contributing off-site runoff ☐ yes ☒ no

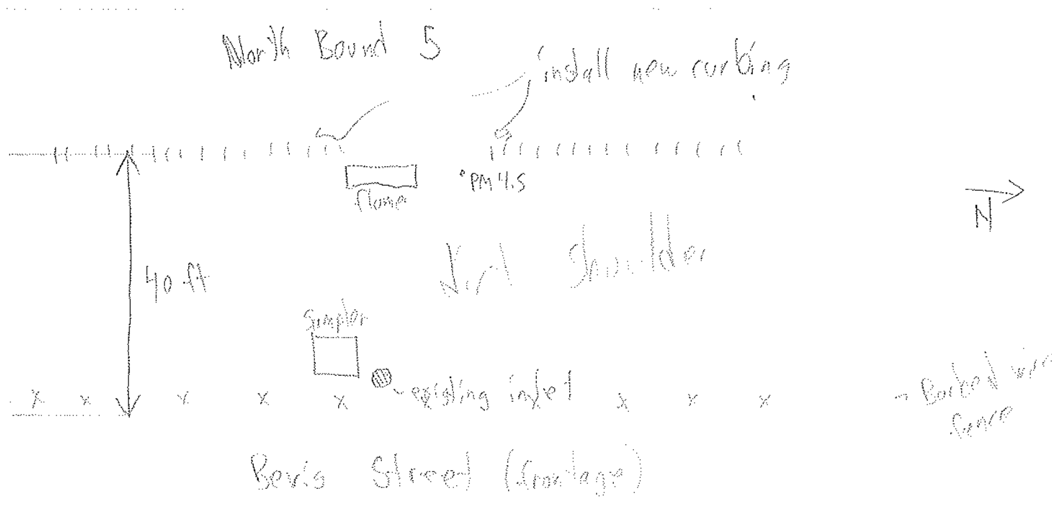
Adjacent industrial sites ☐ yes ☒ no

Adjacent commercial farming ☒ yes ☐ no

Metal guard rails ☐ yes ☒ no

Overall Caltrans Sources Rating (0 to 10): 9

Note: Obtain site plans or draw a sketch and attach copies of photos.



Open Graded – Rubberized Asphalt Paving
SITE EVALUATION FORM
PAIRED SITE

Route/Post Mile: ______________________________

Photograph #'s:____

1.0 Executive Summary

- Project Description
- Candidate Sites
- Screening Criteria Matrix
- Siting Criteria Scoring Matrix
- Results

2.0 Introduction

- Project Description
- Project Background
- Candidate Sites

3.0 Siting Criteria Development

- Siting Criteria Selection
    - Pilot Design Requirements/BMP Design Requirements
    - Monitoring Requirements
    - Safety Requirements
    - Implementation Issues
- Screening Criteria
- Evaluation Criteria
- Weighting Factors

4.0 Data Collection and Analysis

- Data Collection, Analysis, and Review of Existing Data
    - Caltrans Documents Reviewed
    - Other Documents Reviewed
- Preliminary Sizing Approach

5.0 Siting Deviations that Impact Study Plan

- Study Plan Deviations
- Impacts to Pilot Study Objectives

6.0 Site Evaluations and Analysis

- Assumptions
- Screening Criteria Matrix
- Siting Criteria Valuation Matrix
- Siting Criteria Normalization Matrix
- Siting Criteria Scoring Matrix

7.0 Summary of Results and Recommendations

Provide a summary of the activities, findings, and recommendations.

8.0 References

List references as appropriate.

9.0 Appendices

- Field Investigation Forms (including photos)

1.0 Introduction

- Project Description
- Project Background
- Project Location
- Units
- Report Organization

2.0 Existing Site Conditions

- Data Sources
- Topography
- Geotechnical Investigations
- Climate and Hydrology
- Soils
- Groundwater
- Stormwater Conveyance and Treatment
- Utilities
- Hazardous Wastes
- Environmental Setting
- Other Hazards

3.0 Design Criteria

- BMP Design and Monitoring Criteria
    - Water Quality Volume/Flow
    - Monitoring and Maintenance
    - Safety
    - Flood Control
    - Structural Design
    - Electrical/Power
    - Vegetation
- General Design Criteria
    - Existing Facilities & Utilities
    - Hazardous Wastes
    - Landscaping and Irrigation
    - Environmental

4.0 Hydrologic and Hydraulic Analysis

- Methodology
  - Design Criteria
  - Model Description and Inputs
- Watershed Delineation Rainfall Analysis
  - IDF Parameters
  - Rainfall Depths
  - Design Storms (85th percentile, 24-hour storm event)
  - WQV/WQF
- Existing Condition Modeling Results
  - Inflow
  - Outflow
- Project Condition Modeling Results
  - Outflow Hydrographs
  - Outflow Hydraulics

5.0 Design Deviations that Impact Study Plan

Describe deviations from the Study Plan, why they were necessary, and how they might impact the study objectives.

6.0 Design Summary

- BMP Design
  - Inlet Structures
  - Outlet Structures
  - Treatment System (including vegetation that provides treatment)
  - Monitoring
  - Maintenance
  - Safety
  - Flood Control
  - Structural Design
  - Electrical / Power
  - Existing Facilities
  - Utilities
  - Landscaping and Irrigation
  - Hazardous Wastes
  - Environmental

- Non-Standard Special Provisions
- Design Costs
- Construction Cost Estimate
- Lessons Learned

7.0 Construction Support

- Shop Drawings Requiring Consultant Review
- Items of Work Requiring Special Attention
- Key Measurements / Observations to be Made

8.0 OM&M Support

- Monitoring Approach and Features
- Operations & Maintenance Approach and Features

9.0 References

List references as appropriate.

10.0 Appendices

- Existing Plans
- Watershed Delineation Maps
- Intensity-Duration-Frequency Curves
- Hydrology & Hydraulic Calculations
- Design Calculations
- CD Containing Final PS&E

1.0 Introduction

- Project Description
- Project Location
- Report Organization

2.0 Bid Phase Summary

- Bid Addenda
- Bid Summary

3.0 Construction Phase Summary

- Shop Drawing Review
- Requests for Information / Clarification
- Contract Change Orders
    - Plans
    - Special Provisions
    - Schedule
- Acceptance Testing
- Final Construction Cost
- Lessons Learned

4.0 Construction Deviations that Impact Study Plan

Described deviations from the study plan, why they were necessary, and how they may affect the study objectives.

5.0 OM&M Support

- Changes that impact monitoring
- Changes that impact operations & maintenance

6.0 References

List reference as appropriate.

7.0 Appendices

- RFI / RFC Log
- CCO Log

- Daily Construction Reports
- Weekly Construction Site Visit Reports
- Final Cost Breakdown
- CD Containing As-Built Plans
- Example Environmental Commitment Record
- Certificate of Environmental Compliance

1.0 Introduction

- Study Description
- Study Objectives
- BMP Description
- General Scope of Activities (Maintenance and Monitoring)
- Project Organization and Responsibilities

2.0 Description of Site

- Site Access
- Health and Safety Plan (reference in appendix)
- BMP Features

3.0 Operation and Maintenance

- Routine Inspection and Maintenance Requirements
- Equipment and Tools Needed
- Checklist for Inspectors and Maintenance Personnel

4.0 Operation and Maintenance Costs

- Inspection and Maintenance Cost Tracking Database

5.0 Monitoring

- Data Quality Objectives
- Analytical Constituents
- Monitoring Equipment
- Monitoring Preparation and Logistics
    - Weather Tracking
    - Storm Selection Criteria
    - Storm Action Levels
    - Communication/Notification Procedures
    - Monitoring and Equipment Preparation
    - Sample Collection, Preservation, and delivery
- Quality Assurance/Quality Control
- Laboratory Sample Preparation and Analytical Methods
- Data Management and Reporting Procedures

6.0 References

- List of all cited references

7.0 Appendices

- Site Maps and Plans
- Equipment and tools needed
- Checklists for Inspectors and maintenance personnel
- Equipment Data Sheets
- Field Forms
- Vector Control and Management Procedures
- Health and Safety Plan

Appendix K Foreword and Contents

As part of its stormwater Best Management Practice (BMP) evaluation process, Caltrans staff and consultants will summarize the results of a BMP pilot study into a package to be reviewed by the Caltrans various Stormwater Advisory Teams (SWATs) for potential approval. Caltrans BMP evaluation process is summarized in Section 1 of the Pilot Study Guidance Manual (PSGM). The requirements of such a SWAT package are presented in Section 9 of the PSGM.

The SWAT package contents presented herein serve merely as an example of what a SWAT package should be comprised of a how it should be assembled. The specific bioretention details presented are not reflective of actual recommendations presented to Caltrans for including bioretention as an approved BMP. This example package includes:

1.0 Problem Statement

Caltrans needs to add BMPs that can reduce the volume of runoff discharged from its roadways and facilities. This would likely be accomplished through BMPs that infiltrate. Bioretention is a promising BMP that has been studied and documented as having effective water quality capture treatment and infiltration potential (BMP Database 2016 Summary Results). Caltrans therefore conducted the State Route 73 bioretention pilot study on to evaluate the treatment potential of BMPs following BMP evaluation criteria listed in Caltrans Stormwater Management Plan (Caltrans 2016) and the processes outlined in the PSGM.

2.0 BMP Pilot Study Summary

2.1. Background

2.1.1. Location

<Describe BMP location including highway, post-mile, latitude, longitude, and map.P

2.1.2. Schematic

Figure K- 1 shows a schematic of the pilot study BMP. Note that the configuration of this BMP differs from other bioretention facilities studied and implemented by others, particularly in its use of a liner and 30 inches of planting soil. Configurations used by others are often unlined and use only 12-24 inches of planting soil Alternative designs are described further below.

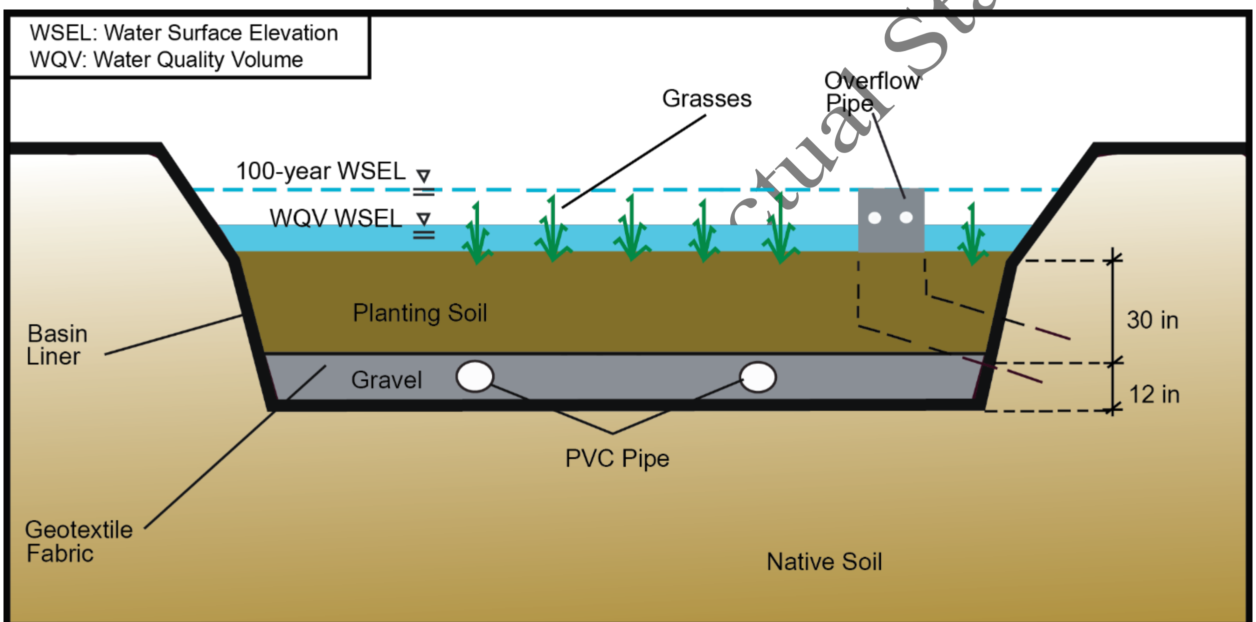


Figure K- 1. Study Bioretention BMP Schematic

2.1.3. Project Plan Sheets and Redlines

Appendix X provides the project plan sheets and redlines for the BMP.

<Include BMP project plan sheets and redlines>

Appendix Y provides the design calculations for NPDES permit compliance.

<Include design calculations for NPDES permit compliance, including:

- *WQV from the 85th percentile 24-hour event for the site location*
- *Media filter rates*
- *Drainage design hydrology and hydraulic calculations per the Highway Design Manual (Caltrans 2018e) for the locations>*

2.2. Monitoring Activities

- About 22 events monitored
- Wet seasons monitored: 2006-07, 2007-08, 2014-1,5 and 2015-16
- *<Include other details as appropriate.>*

2.3. Results

Study results are presented in the Performance Summary section of this SWAT Package.

2.4. Alternatives

The bioretention BMP used in this study differs from the typical bioretention BMPs that have been studied and implemented more recently. Figure K- 2 illustrates the difference in layers and lining with the SR 73 design on the left and more typical recent design on the right.

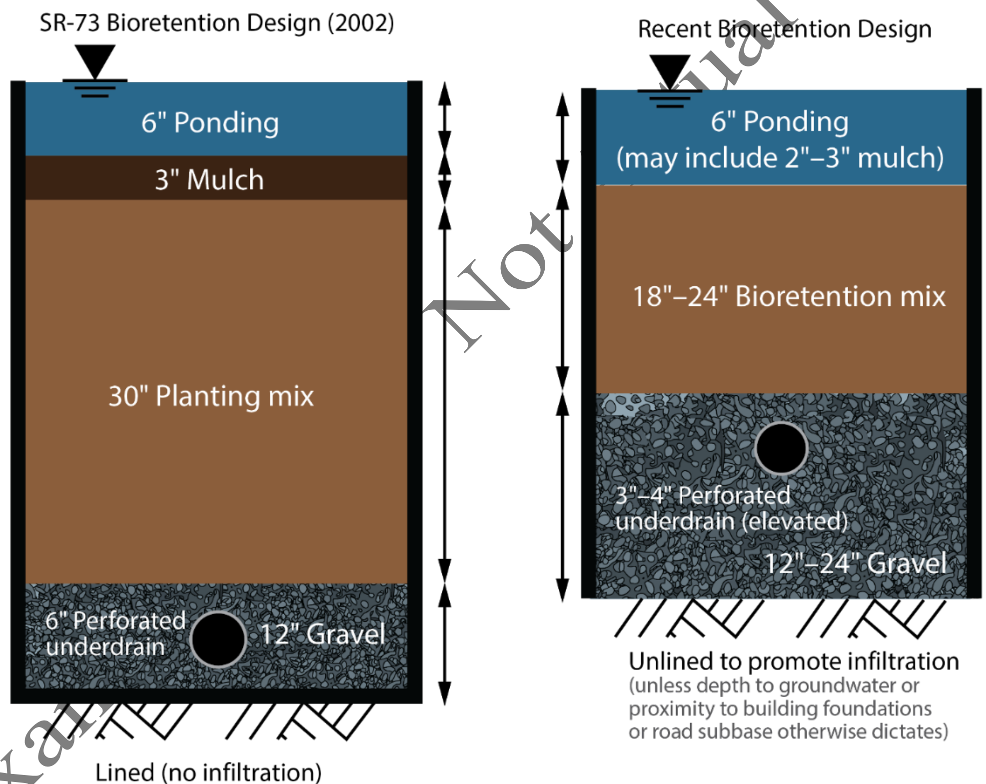


Figure K- 2. SR-73 and Recent Bioretention Design Comparison Schematic

3.0 BMP Fact Sheet

3.1. General Description

- Bioretention BMPs are depressed landscapes into which runoff is directed and allowed to collect, filter, and sometimes infiltrate. These BMPs come in a variety of configurations. A schematic is shown in Figure K- 3.
- All include a few inches of ponding depth (often 4 to 6 inches).
- A riser (often a drop inlet that is raised above the planting media) allows a means of bypass in case of overflows.
- *<Include other details as appropriate.>*

3.2. General Description

Table K- 1 presents the relevant treatment mechanisms for the BMP.

Table K- 1. Treatment Mechanisms

Mechanisms	Applicable (Yes) or Not Applicable (No)
Infiltration	Yes (for unlined systems)
Evapotranspiration	Yes
Filtration and/or Adsorption	Yes
Biochemical Transformation	Yes
Rainfall and Runoff Harvest	No
Sedimentation	Yes
Floatation	No
Plant Uptake	Yes
Trash Capture	Yes

**Does not include incidental and minor losses from treatment mechanisms that are not deliberately allowed or enhanced by the BMP design or operation*

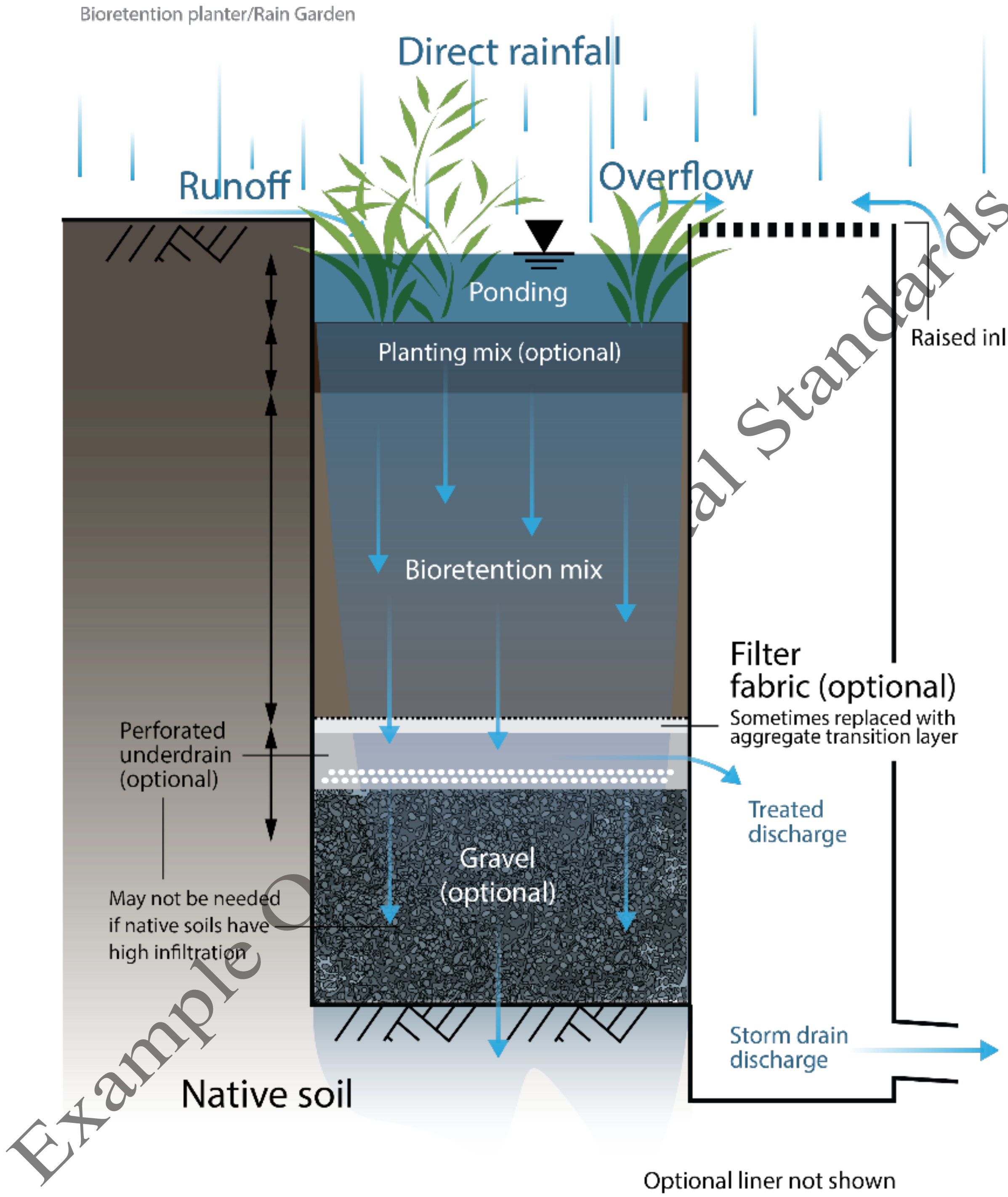


Figure K- 3. Bioretention BMP Schematic (OWP 2018)

3.3. Variations and Alternative Names

- Bioretention basins
- Rain gardens
- Infiltrating stormwater planters
- *<List others as appropriate>*

3.4. Advantages & Limitations

3.4.1. <u>Advantages</u>

- Bioretention BMPs can be both inexpensive and add aesthetic appeal
- Can be more cost-effective because it can have less concrete than other BMPs
- *<List others as appropriate>*

3.4.2. <u>Limitations</u>

- Requires terracing for steeper slopes
- Limited to a small contributing drainage area
- *<List others as appropriate>*

3.5. Siting Considerations

- The site should be relatively flat
- Use irrigation during plant establishment and during dry season in some climates.
- *<List others as appropriate>*

3.6. Design Considerations

- Contributing drainage area
- Drawdown time
- *<List others as appropriate>*

3.7. Construction Considerations

- Use appropriate construction BMPs to prevent sediment loading and erosion
- Use light construction equipment to avoid subgrade compaction
- *<List others as appropriate>*

3.8. Maintenance Considerations

- Identify and promote desired plant species; increased density decreases weeds
- Remove unwanted plant species and litter
- *<List others as appropriate>*

3.9. References

<List all relevant references>

4.0 Performance Summary

4.1. Results of Caltrans Pilot Studies

This section provides the monitored treatment performance at the SR 73 Basin 1149L pilot site. This basin was a lined bioretention BMP.

Table K- 2, Table K- 3, Table K- 4, and Table K- 5 present the observed percent reduction in concentrations of monitored constituents based on 22 events monitored between 2006 and 2016. The monitoring results indicate:

- Very effective treatment of TSS, total and dissolved metals, TKN, and PAHs (PCBs were not sampled)
- Less effective treatment of total phosphorus, orthophosphate, and nitrate. However, the design tested did not include a raised underdrain, which is known to improve nitrate treatment, and literature studies suggest that orthophosphate export only occurs during the first few years of media use.
- Variable treatment of bacteria. The mean bacteria treatment data are affected by one or two high export events that skew calculations of mean removals. The median treatment data, which show some treatment, are considered more reliable.

4.2. Result Limitations

The results represent performance from a lined bioretention planter. Water quality from an unlined planter of similar configuration and well-draining underlying soils (hydrologic group A and B) are expected to perform comparably, with increased load reductions due to likely reduction in discharge volumes that result from infiltration. Concentration and load reductions from unlined planters overlying C or D type soils are expected to perform comparably to the tested planter, with slightly better load reductions.

Table K- 2 Concentration Percent Removals (2006-08, 2013-16)

Storm Event	TSS	Total Cadmiu	Dissolved Cadmium	Total Chromiu	Dissolved Chromium	Total Copper	Dissolved Copper
10/13/2006	78%	75%	52%	40%	-6%	59%	33%
12/9/2006	-27%	29%	NC	4%	20%	65%	67%
12/16/2006	93%	74%	NC	71%	21%	79%	50%
12/27/2006	90%	51%	NC	65%	-38%	70%	56%
1/30/2007	PO/SD	78%	64%	73%	3%	78%	76%
2/11/2007	82%	35%	26%	26%	18%	47%	48%
2/22/2007	91%	76%	17%	66%	12%	75%	51%
4/20/2007	83%	78%	59%	49%	17%	66%	53%
12/7/2007	PO/SD	NC	NC	-83%	-36%	31%	45%
12/18/2007	PO/SD	70%	NC	52%	5%	66%	39%
1/4/2008	81%	91%	77%	73%	NC	83%	34%
1/23/2008	76%	83%	NC	52%	29%	81%	38%
2/3/2008	64%	NC	NC	55%	23%	75%	46%
2/21/2008	43%	NC	NC	46%	NC	63%	27%
2/28/2014	88%	75%	CN	65%	CN	72%	CN
4/2/2014	60%	44%	CN	47%	CN	84%	CN
11/1/2014	93%	59%	CN	61%	CN	69%	CN
12/13/2014	98%	82%	CN	50%	CN	78%	CN
12/17/2014	97%	56%	CN	19%	CN	81%	CN
10/5/2015	95%	CN	17%	76%	CN	85%	82%
1/5/2016	95%	81%	14%	68%	49%	87%	59%
3/8/2016	93%	CN	49%	61%	CN	87%	69%
Mean	78%	67%	42%	47%	9%	72%	51%
Median	88%	75%	49%	54%	17%	75%	50%

Abbreviations:

CN = Constituent not analyzed/sample not collected
PO/SD = Possible outlier/suspect data point
NC = Not calculated because influent and effluent at or below RL

Notes:

1 The TSS Overall Mean calculation does not include possible outlier/suspect data
2 The 2015-16 data is still under review and subject to change
3 Total nitrogen calculated as sum of total TKN and total nitrate, with assumption that nitrite is negligible. Inclusion of total nitrogen values calculated as sum of total TKN and total nitrate results in no net removal of total nitrogen.
4 Microbiological data are from grab samples
5 Insufficient data available for total coliform, total orthophosphate analysis

Table K- 3. Concentration Percent Removals (2006-08, 2013-16)

Storm Event	Total Lead	Dissolved Lead	Total Nickel	Dissolved Nickel	Total Selenium	Total Zinc	Dissolved Zinc
10/13/2006	63%	PO/SD	63%	57%	CN	87%	78%
12/9/2006	13%	NC	56%	56%	CN	67%	83%
12/16/2006	89%	NC	64%	32%	CN	89%	73%
12/27/2006	79%	NC	58%	51%	CN	84%	73%
1/30/2007	80%	NC	73%	74%	CN	84%	85%
2/11/2007	17%	NC	51%	59%	CN	64%	62%
2/22/2007	85%	NC	68%	57%	CN	83%	72%
4/20/2007	71%	NC	67%	63%	CN	76%	74%
12/7/2007	-62%	NC	-4%	18%	CN	-8%	33%
12/18/2007	82%	NC	56%	40%	CN	77%	66%
1/4/2008	87%	NC	74%	NC	CN	87%	72%
1/23/2008	76%	NC	63%	NC	CN	85%	71%
2/3/2008	79%	NC	61%	NC	CN	77%	57%
2/21/2008	67%	NC	54%	NC	CN	69%	53%
2/28/2014	89%	CN	CN	CN	-100%	91%	CN
4/2/2014	88%	CN	CN	CN	9%	92%	CN
11/1/2014	90%	CN	CN	CN	-123%	95%	CN
12/13/2014	92%	CN	CN	CN	-144%	94%	CN
12/17/2014	96%	CN	CN	CN	-250%	93%	CN
10/5/2015	85%	62%	CN	CN	CN	97%	98%
1/5/2016	93%	21%	CN	CN	CN	96%	86%
3/8/2016	94%	52%	CN	CN	CN	95%	89%
Mean	71%	45%	57%	51%	-122%	81%	72%
Median	83%	52%	62%	56%	-123%	86%	73%

Abbreviations:

CN = Constituent not analyzed/sample not collected
PO/SD = Possible outlier/suspect data point
NC = Not calculated because influent and effluent at or below RL

Notes:

1 The TSS Overall Mean calculation does not include possible outlier/suspect data
2 The 2015-16 data is still under review and subject to change
3 Total nitrogen calculated as sum of total TKN and total nitrate, with assumption that nitrite is negligible. Inclusion of total nitrogen values calculated as sum of total TKN and total nitrate results in no net removal of total nitrogen.
4 Microbiological data are from grab samples
5 Insufficient data available for total coliform, total orthophosphate analysis

Table K- 4. Concentration Percent Removals (2006-08, 2013-16)

Storm Event	Total Phosphorous	Dissolved Orthophosphate	Total Nitrogen	Total Nitrate	Dissolved Nitrate	Total TKN
10/13/2006	78%	75%	52%	40%	-6%	59%
12/9/2006	-27%	29%	NC	4%	20%	65%
12/16/2006	93%	74%	NC	71%	21%	79%
12/27/2006	90%	51%	NC	65%	-38%	70%
1/30/2007	PO/SD	78%	64%	73%	3%	78%
2/11/2007	82%	35%	26%	26%	18%	47%
2/22/2007	91%	76%	17%	66%	12%	75%
4/20/2007	83%	78%	59%	49%	17%	66%
12/7/2007	PO/SD	NC	NC	-83%	-36%	31%
12/18/2007	PO/SD	70%	NC	52%	5%	66%
1/4/2008	81%	91%	77%	73%	NC	83%
1/23/2008	76%	83%	NC	52%	29%	81%
2/3/2008	64%	NC	NC	55%	23%	75%
2/21/2008	43%	NC	NC	46%	NC	63%
2/28/2014	88%	75%	CN	65%	CN	72%
4/2/2014	60%	44%	CN	47%	CN	84%
11/1/2014	93%	59%	CN	61%	CN	69%
12/13/2014	98%	82%	CN	50%	CN	78%
12/17/2014	97%	56%	CN	19%	CN	81%
10/5/2015	95%	CN	17%	76%	CN	85%
1/5/2016	95%	81%	14%	68%	49%	87%
3/8/2016	93%	CN	49%	61%	CN	87%
Mean	78%	67%	42%	47%	9%	72%
Median	88%	75%	49%	54%	17%	75%

Abbreviations:

CN = Constituent not analyzed/sample not collected
PO/SD = Possible outlier/suspect data point
NC = Not calculated because influent and effluent at or below RL

Notes:

1 The TSS Overall Mean calculation does not include possible outlier/suspect data
2 The 2015-16 data is still under review and subject to change
3 Total nitrogen calculated as sum of total TKN and total nitrate, with assumption that nitrite is negligible. Inclusion of total nitrogen values calculated as sum of total TKN and total nitrate results in no net removal of total nitrogen.
4 Microbiological data are from grab samples
5 Insufficient data available for total coliform, total orthophosphate analysis

Table K- 5. Concentration Percent Removals (2006-08, 2013-16)

Storm Event	Total PAHs	Total PCBs	Pesticides	Fecal Coliform	Enterococcus
10/13/2006	33%	63%	PO/SD	63%	57%
12/9/2006	67%	13%	NC	56%	56%
12/16/2006	50%	89%	NC	64%	32%
12/27/2006	56%	79%	NC	58%	51%
1/30/2007	76%	80%	NC	73%	74%
2/11/2007	48%	17%	NC	51%	59%
2/22/2007	51%	85%	NC	68%	57%
4/20/2007	53%	71%	NC	67%	63%
12/7/2007	45%	-62%	NC	-4%	18%
12/18/2007	39%	82%	NC	56%	40%
1/4/2008	34%	87%	NC	74%	NC
1/23/2008	38%	76%	NC	63%	NC
2/3/2008	46%	79%	NC	61%	NC
2/21/2008	27%	67%	NC	54%	NC
2/28/2014	CN	89%	CN	CN	CN
4/2/2014	CN	88%	CN	CN	CN
11/1/2014	CN	90%	CN	CN	CN
12/13/2014	CN	92%	CN	CN	CN
12/17/2014	CN	96%	CN	CN	CN
10/5/2015	82%	85%	62%	CN	CN
1/5/2016	59%	93%	21%	CN	CN
3/8/2016	69%	94%	52%	CN	CN
Mean	51%	71%	45%	57%	51%
Median	50%	83%	52%	62%	56%

Abbreviations:

CN = Constituent not analyzed/sample not collected
PO/SD = Possible outlier/suspect data point
NC = Not calculated because influent and effluent at or below RL

Notes:

1 The TSS Overall Mean calculation does not include possible outlier/suspect data
2 The 2015-16 data is still under review and subject to change
3 Total nitrogen calculated as sum of total TKN and total nitrate, with assumption that nitrite is negligible. Inclusion of total nitrogen values calculated as sum of total TKN and total nitrate results in no net removal of total nitrogen.
4 Microbiological data are from grab samples
5 Insufficient data available for total coliform, total orthophosphate analysis

4.3. Supporting Information: Summaries from the International BMP Database

Table K- 6 provides performance summaries of bioretention BMPs reported in the International BMP Database 2014 Pollutant Category Statistical Summary Report: Solids, Bacteria, Nutrients, and Metals. They give a perspective of the performance of various BMPs reported by others, for comparison to Caltrans bioretention BMP pilot study results.

Table K- 6. Pilot Study Results for TSS (mg/L)

BMP Type	# of Studies	# of Influent, Effluent Data Points	Q1 Influent, Effluent (mg/L)1	Q2 Influent [95% Conf. Interval] (mg/L)1,2,3	Q2 Effluent [95% Conf. Interval] (mg/L)1,2,3,4	Q3 Influent, Effluent (mg/L)1
Biostrip	19	361, 282	20.0, 10.0	44.1 [39, 48]	19 [15.9, 21]	90.0, 35.0
Bioswale	23	399, 346	9.0, 10.0	27.7 [21, 31.6]	21.6 [17.8. 24]	67.0, 43.0
Bioretention	22	461, 393	18.0, 4.9	38.1 [31, 42]	9.9 [7, 10]	86.0, 20.0
Composite	10	202, 174	42.4, 8.0	87.6 [75.1, 101.5]	18.4 [14, 19.3]	178.8, 36.5
Detention Basin	22	321, 336	21.0, 10.0	68.2 [52.3, 77.3]	23.3 [19.5, 26]	128.0, 47.0
Media Filter	23	381, 358	21.1, 3.0	450.9 [42.8, 58]	8.4 [6.3, 9.8]	110.5, 19.9
Porous Pavement	8	356, 220	35.0, 14.0	90.3 [69, 115]	24.9 [21.5, 27]	230.0, 44.4
Retention Pond	56	923, 933	15.0, 4.3	47.7 [40, 54]	11.5 [10, 12.3]	139.8, 28.0
Wetland Basin	19	395, 385	11.0, 3.5	24.5 [19.1, 28.9]	9.4 [7.4, 11]	63.3, 20.6
Wetland Basin/ Retention Pond	75	1318, 1318	13.3, 4.0	37.9 [34, 41.6]	10.9 [9.6, 11.7]	110.0, 25.4
Wetland Channel	8	171, 151	12.0, 8.0	18.9 [16, 21]	14.4 [10, 16]	47.5, 27.0

[1]*Q1 = 25^{Th} percentile; Q2 – median (50^{Th} percen12.0tial); Q3 = 75^{th} percentile*
[2]*Computed using the BCa bootstrap method described by Efron and Tibishirani (1993)*
[3]*Values in brackets are the lower and upper confidence limits for the 95% confidence level*
[4]*Hypohtesis testing showed statistically significant decreases for this BMP category.*

<Tabulate other constituents as appropriate>

4.4. Results of Performance-Based Ranking Analysis

The results from the performance-based ranking analysis for each of Caltrans target design constituents for the SR73 pilot study are summarized below. The analysis followed that outlined in Appendix L of the PSGM.

4.4.1. Constituents

The following four constituents were evaluated for the BMP, as directed in the PSGM:

- TSS
- Total Phosphorus
- Total Nitrogen
- Total Copper

4.4.2. Quality Assurance and Outlier Analysis

A review of the data for the four selected constituents indicated that certain storm events contained outlier results per procedures in the PSGM, Appendix K. The outlier storm events and constituents for the BMP were:

- 10/13/2006, Total N
- 1/30/2007, Total N
- 10/13/2006, Total P
- 1/21/2006, Total Cu

These data were flagged for further scrutiny to determine if any censoring was justified. As a result, it was noted that the 10/13/2006 storm event was the first monitored storm event after construction of the bioretention BMP was completed. Initial nutrient removal performance of new bioretention media may not represent long-term behavior. Therefore, nutrient data from this storm event were excluded from all analysis. All other data were used.

An assessment of quality assurance/quality control (QA/QC) results did not identify additional data that should be excluded.

4.4.3. Performance Efficiency Calculations

Section 5.3 and the T-1 Checklist (Appendix E) of the PPDG (Caltrans 2019) presents a process for selecting from approved BMPs. The process was developed to address Total Maximum Daily Load (TMDL) and 303(d) scenarios. Since TMDLs are typically load-based, a load-based approach was used for the latest PPDG rankings. Also, using a load reduction approach makes sense for infiltration BMPs, since both volume and pollutants are retained. Bioretention BMPs, which can allow infiltration, were therefore assessed using the load-based methodology from the PSGM Appendix L.

Although bioretention BMPs can be built to allow infiltration, the SR73 bioretention BMP tested by Caltrans was constructed with an impermeable liner; therefore, the measured water quality data represented performance for the < 20 percent infiltration range (actual infiltration was 0 percent). For this scenario, the Sum of Loads Method described in Appendix L of the PSGM was used to calculate a load reduction efficiency:

$$\eta_c = \frac{\sum C_i - \sum C_e}{\sum C_i} * 100\%$$
Eq. (a)

$$\eta_L = \eta_C + \eta_v - \eta_C \cdot \eta_v$$
Eq. (b)

Where:

C_i = influent concentration from each storm event (i)
C_e = effluent concentration from each storm event (i)
η_c = concentration reduction efficiency calculated from Eq. (b)
η_v = assumed volume reduction (infiltration) efficiency = 0%
η_L = load reduction efficiency for <20% volume reduction efficiency

The impermeable liner is not a design requirement. This BMP can be constructed without an impermeable liner, so infiltration can occur. The load reduction efficiencies for the remaining two infiltration ranges (20-50% and >50%) were calculated based on an assumed infiltration percentage (20 and 50%, respectively) and observed concentrations reductions using Eq. b where n_v is set to 20% and 50%, respectively.

4.4.4. <u>Performance Tier Recommendation</u>

The calculated load reduction efficiencies for the bioretention BMP are listed in Table K-7. For each efficiency-infiltration range combination for each constituent, a performance tier was assigned following the thresholds listed in Appendix M of the PSGM and Caltrans 2018:

- Tier 0: Superior BMPs (100% load reduction of design storm)
- Tier 1: More effective BMPs (>60% load reduction efficiency)*
- Tier 2: Less effectives BMPs (20-60% load reduction efficiency)*
- Tier 3: Not effective BMPs (<20% load reduction efficiency)

  **55% threshold is used for total phosphorous*

<Insert examples of calculations for one constituent.>

The resulting performance tiers are listed in Table K- 7. Reclassification of the tiers was not required. Note that the results for this bioretention BMP are based on design guidelines from the early 2000s. This design included a 30-inch media layer. Current bioretention designs established by others specify an 18- to 24-inch media layer. Treatment performance is not expected to change by decreasing the media layer.

Table K- 7. Bioretention Basin 1149L Load Reduction Efficiencies and Performance Tiers

Constituent	Infiltration Range*	Load Reduction Efficiency	Performance Tier
TSS	< 20%	78%	Tier 1
TSS	20%–50%	82%	Tier 1
TSS	> 50%	89%	Tier 1
Total P	< 20%	28%	Tier 2
Total P	20%–50%	43%	Tier 2
Total P	> 50%	64%	Tier 1
Total N	< 20%	57%	Tier 2
Total N	20%–50%	66%	Tier 1
Total N	> 50%	78%	Tier 1
Total Cu	< 20%	74%	Tier 1
Total Cu	20%–50%	79%	Tier 1
Total Cu	> 50%	87%	Tier 1

**Load reduction efficiencies for the < 20% infiltration range were calculated using Eq. (a). The load reduction efficiencies for the other two infiltration ranges were calculated using Eqs. (b&c).*

5.0 Siting Recommendations

5.1. Intended Applications

Results of the SR 73 pilot study and that reported by others were used to develop the recommendations presented herein and are intended for statewide use. Recommendations for modifications appropriate to site-specific conditions are also provided.

5.2. Siting Constraints

Table K- 8 presents recommended siting constraints.

Table K- 8. Recommended Siting Constraints

Site Feature	Constraint
Drainage area	• No constraints
Slope/ elevation change/ hydraulic head	• Flat to shallow slopes (<10%); not effective on steep slopes • No head requirements
<List others as appropriate.>	• *<List others as appropriate.>*

6.0 Design Recommendations

Table K- 9 presents recommended design criteria. These are based on a combination of findings from the configuration tested by Caltrans in the SR 73 pilot study and those implemented and tested by others. Note that some design recommendations differ from the design features studied by Caltrans. Such recommendations were selected to improve function and feasibility, without impacting the treatment performance (based on best professional judgement). Table K- 10 lists the modifications and reasoning for making them.

Table K- 9. Design Recommendations

Design Component	Constraint
Drainage Area	• Best for sites with drainage area of less than 1 acre.
Ponding zone	• Maximum of 6”
Mulch	• Consult with District Landscape Architect as to whether mulch is needed for vegetation health and what type
<List others as appropriate.>	• <List others as appropriate.>

Table K- 10. Modified Design Components and Reasoning*

Design Component	Modification and Reasoning
Lining	• BMP studies by Caltrans was not lined. Design recommendations allow BMP to not be lined to enhance infiltration and volume reduction, as long as depth to groundwater, groundwater contamination, or other impacts from infiltration are not an issue.
<List others as appropriate.>	• <List others as appropriate.>

**Modifications were made to improve function and feasibility without impacting treatment, based on best professional judgement.*

7.0 Recommendations for Specifications and Special Provisions

7.1. Plastic Pipe

Plastic pipe shall conform to the provisions in Section 64, "Plastic Pipe," of the Standard Specifications and these special provisions.

7.2. Plastic Underdrain Pipe

Plastic underdrain pipe shall conform to the provisions in Section 68, "Subsurface Drains," of the Standard Specifications and these special provisions.

7.3. Bioretention Planter

Materials for construction of bioretention areas shall conform to the provisions in Section 20-2, “Materials” of the Standard Specifications and these special provisions.

7.3.1. Seeds

Seeds shall conform to *<continue information as appropriate>.*

7.3.2. Bioretention Soil Mix

Bioretention Soil Mix (BSM) shall *<insert information as appropriate>*

7.3.3. Gravel Backfill

Gravel backfill shall conform to the *<list information as appropriate>.*

7.3.4. Quality Assurance

Quality assurance conformance testing shall *<list information as appropriate>*

7.4. Geomembrane

Geomembrane shall function as *<insert text as appropriate>.*

7.4.1. Materials

Geomembrane shall be *<insert text as appropriate>.*

7.4.2. <List others as appropriate.>

7.5. <List others as appropriate.>

8.0 Maintenance Recommendations

8.1. Description

The recommendations below are intended for modifying an existing section of Caltrans Maintenance Staff Guide or developing a new section to incorporated maintenance needs for bioretention BMPs. In developing these recommendations, the most recent version of the Maintenance Staff Guide (CTSW-RT-18-314.20.1) was reviewed.

8.2. Implementation

Table K- 11 lists recommended maintenance indicators and how frequently inspections should be done to identify the indicated conditions. Maintenance activities to be conducted in response to the indictors are also provided. Table K- 12 shows relevant cut sheets from Caltrans Maintenance Staff Guide that provide specific details for the maintenance activities. Chemical vegetative treatment control measures will not be used on the vegetation in the bioretention BMPs.

Table K- 11. Maintenance Frequency, Activities, and Indicators

Frequency	Description	Maintenance Indicator	Responsive Action
Annually	Inspect BMP for dead or diseased vegetation.	Dead or diseased vegetation	Remove and replace with drought tolerant and flood tolerant grasses[1,2]
Annually	Flush underdrains and piping with high water pressure using clean out access.[3]	Not applicable	Not applicable
Annually, 3 days after a 0.75-inch storm event[4]	Inspect BMP for standing water	Drain time exceeds 96 hours after end of rain event	Remove sediment, trash, and debris. Check and flush underdrain[3] Notify the District Maintenance Stormwater Coordinator/Vector Control District
Semi-annually	Inspect BMP for accumulated trash/debris	Trash/debris present	Remove and dispose of trash and debris.
<List others as appropriate>	*<List others as appropriate>*	*<List others as appropriate>*	*<List others as appropriate>*

[1] DO NOT USE synthetic fertilizers or pesticides; doing so introduces pollutants into the storm drain system and surrounding environment.
[2] Other types of vegetation may be acceptable, but only outside the clear recovery zone.
[3] The slurry generated by the cleaning should be captured and routed back to the bioretention BMP. Apply wet slurry over soil. Avoid concentrating the slurry discharge in one spot and ensure sediment depth below marker on staff gage.
[4] If no such event occurs before April 1, conduct wet season inspections in April.

Field measurements of maintenance indicators are made by visual observation. Frequencies provided are for the minimum required level of service. Greater maintenance frequencies may be required depending on the particular site and level of traffic. Design elements include an impermeable liner, filter fabric, a top soil layer, and an underdrain collection system. Not all elements may be present. Therefore, not all measures described below will apply.

Table K- 12. Relevant Maintenance Staff Guide Cut Sheets

Maintenance Staff Guide Section	Applicable Cut Sheets
E Family - Landscaping	• B-39 Manual Veg Control, • B-40 Landscaped Mechanical Vegetation Control/Mowing • B-41 Landscaped Tree and Shrub Pruning, Brush Chipping, Tree & Shrub Removal
<List others as appropriate>	*<List others as appropriate>*

9.0 Life-Cycle Costs Comparison

Life-cycle costs are developed following the approach used in the Life Cycle Costs Based on 2007 BMP Operation and Maintenance Cost Analysis Technical Memorandum (CTSW-TM-08- 176-11.1). *<Briefly describe process.>*

Bioretention construction costs are based on the BEES estimate for site State Route 73 basin 1149L. No actual construction costs are available. Bioretention maintenance costs were assumed to be similar to maintenance costs for infiltration basins.

Table K- 13 compares estimated 2016 life-cycle costs for the Bioretention (1149L) with approved BMPs. Although bioretention costs are expected to be higher than those for earthen Austin Sand Filters because of a larger footprint, the difference shown in Table K- 13 is substantially higher than expected. The differences can be explained by:

- Uncertainty in the construction cost for basin 1149L
- Unrealistically low-cost estimates for the two Austin Sand Filters (earthen). Both the Retrofit Study Report (Caltrans 2003) and the Life Cycle Costs Based on 2007 BMP Operation and Maintenance Cost Analysis Technical Memorandum (Caltrans 2008) acknowledge the lower than expected costs.

Table K- 13. 2016 Life Cycle Costs Based on 2007 BMP Operation and Maintenance Cost Analysis Technical Memorandum (CTSW-TM-08-176-11.1)

BMP	Present Worth Adjusted Capital Cost /cf	Present Worth Maintenance Cost /cf	Present Worth Total Cost /cf
Wet Basin (1)	$70.50	$18.42	$88.93
Multi Chamber Treatment Trains (2)	$76.37	$6.44	$82.81
Delaware Sand Filter (1)	$77.88	$2.17	$80.05
Bioretention (1149L)	$65.50	$3.24	$68.73
Traction Sand Trap	$61.10	$1.30	$62.40
Austin Sand Filter – Concrete (5)	$58.94	$2.22	$61.16
Biofiltration Swale (6)	$30.62	$3.58	$34.21
Infiltration Trench (2)	$29.86	$1.72	$31.58
Biofiltration Strip (3)	$30.46	$0.58	$31.05
Gross Solids Removal Device (15)	$24.97	$4.32	$29.28
Detention Device (4)	$24.04	$4.34	$28.37
Austin Sand Filter – Earthen (2)	$17.83	$2.71	$20.54
Infiltration Basin (2)	$15.03	$3.24	$18.27

Note: All costs normalized to WQV

10.0 References

Caltrans 2009. Discharge Characterization Report Study Update. CTSW-RT-09-182.44.2. January.

Caltrans 2010. Performance-Based Ranking of Treatment BMPs. CTSW-RT-10-239.01. February.

Caltrans 2017. Stormwater Quality Handbooks: Project Planning and Design Guide. CTSW-RT-17-314.24.1. July.

Caltrans 2018. Performance-Based Ranking of Treatment BMPs, Addendum #1. CTSW-RT-18-350.04.01. June.

Koerner, R.M., Koerner, G.R., 2015. Lessons learned from geotextile filter failures under challenging field conditions. Geotextile. Geomembranes 43, 272-281.

OWP at Sacramento State (OWP). 2018. Stormwater Capture Guidance for Schools (California Practices to Used Runoff Effectively). December 2018.

1.0 Introduction

This document presents methods for comparing the treatment performance of a candidate best management practice (BMP) to that of approved BMPs for the purpose of approving the candidate BMP for future Caltrans use. The results of the methods are documented in a “SWAT package” that is submitted to Caltrans Stormwater Advisory Teams (SWATs) for consideration. Chapter 9 of the Caltrans Pilot Study Guidance Manual (PSGM, Caltrans 2021) describes the contents to be included in the SWAT package, including a comparison of performance between Caltrans approved BMPs and candidate BMPs. This document describes how Caltrans makes such comparisons.

Several analytical steps are required before comparing the performance, including analysis to show that the candidate BMP provides statistically significant treatment. These analyses are described in Appendix D of the PSGM.

2.0 Background

In 2010, Caltrans compared the performance of several treatment BMPs and documented results in The Performance-Based Ranking of Treatment BMPs (Caltrans 2010a). The BMPs evaluated were those for which Caltrans had collected pilot study data over the last decade. The BMP performance was compared based on load and concentration reductions for each Target Design Constituent (TDC) listed in the Caltrans 2007 Project Planning and Design Guide (PPDG, Caltrans 2007a). Based on the performance findings, Caltrans developed new BMP design and selection methodologies, and updated the PPDG to reflect them (Caltrans 2010b). In 2012 and 2017, Caltrans updated the PPDG again, but these updates did not include any new treatment BMPs or alter the BMP selection process.

In 2018, Caltrans evaluated additional BMPs for which pilot study data had become available, documenting the findings in a report referred to as Addendum #1 (Caltrans 2018f). The 2018 evaluation did not assess all the TDCs defined in the 2017 PPDG (which were the same as in the 2010 PPDG), but instead evaluated a subset of TDCs that were considered to be representative of various constituent types. Performance evaluations were conducted for percent load reductions; concentration reductions were not evaluated under the assumption that most Total Maximum Daily Loads (TMDLs, the primary compliance criteria driving Caltrans BMP implementation policies) are load-based. Using the findings from Addendum #1 (Caltrans 2018f), Caltrans developed methods for designing and selecting among the previously approved and newly approved BMPs, and updated the PPDG in 2019 (Caltrans 2019e).

The 2010 Performance-Based Ranking (Caltrans 2010a) and 2018 Addendum #1 (Caltrans 2018f) serve as the basis for instructions contained herein to compare the performance of new BMPs and assign a classification for selection. The instructions are organized as follows:

- Constituents of Concern
- Framework for Load-Based Comparisons and Classification
- Steps for Load-Based Comparisons and Classification
- Steps for Concentration-Based Comparisons and Classification

3.0 Constituents of Concern

Data for the following constituents should be used in comparing BMPs, consistent with the constituents used in Addendum #1 of Caltrans Performance-Based Ranking of Treatment BMPs (Caltrans 2018f):

- Total suspended solids (TSS)
- Total phosphorus
- Total nitrogen
- Total copper

In the past, total zinc has been considered as a constituent for ranking, but Caltrans has not formally included it to date. Total zinc may be included in the future if deemed appropriate.

These four constituents comprise a subset of Caltrans TDCs. Caltrans 2019 PPDG defines a TDC as "…a pollutant that has been identified during Departmental runoff characterization studies to be discharging with a load or concentration that commonly exceeds allowable standards and which is considered treatable by currently available Department-approved Treatment BMPs." Caltrans full set of TDCs were selected based on the results of past discharge characterization studies conducted by Caltrans, pollutant categories listed in the Caltrans 2013 National Pollutant Discharge Elimination System (NPDES) permit (State Water Board 2012), and TMDL and 303-d listings.

Caltrans 2013 NPDES permit (State Water Board 2012) categorized the TMDLs to which Caltrans is subject into the following eight TMDL pollutant categories:

1. Sediment/nutrients/mercury/siltation/turbidity TMDLs
2. Metals/toxics/pesticides TMDLs
3. Trash TMDLs
4. Bacteria TMDLs
5. Diazinon TMDLs
6. Selenium TMDLs
7. Temperature TMDLs
8. Chloride TMDLs

This categorization assumes (as stated in Caltrans 2013 permit) that pollutant types within the same category can be addressed by similar control measures, monitoring, and adaptive management (State Water Board 2012). Attachment IV of the 2013 permit lists the specific implementation requirements for each pollutant category.

Notably, not all permit-listed pollutant categories are represented in the constituents of concern evaluated for BMP performance comparisons presented in this document. The four constituents that are evaluated fall into two of these permit categories; TSS, total phosphorus, and total nitrogen represent category #1 (sediment/ nutrients/ mercury/ siltation/ turbidity), while total copper represents category #2 (metals/ toxics/ pesticides TMDLs). The other permit-listed pollutant categories have not been included in the BMP treatment comparisons, primarily because constituents would be adequately controlled by institutional BMPs such as street sweeping, illicit discharge prevention, public education, shade, and prohibition/discontinuation of product use. These measures, as stated in the permit, are assumed to adequately control pollutant categories 3 (trash), 5 (diazinon), 6 (selenium), 7 (temperature), and 8 (chloride). The permit also states that such measures, combined with infiltration, can adequately address category 4 (bacteria). Finally, the permit states that pollutant category 1 (sediment/ nutrients/ mercury/ siltation/ turbidity) should be controlled by preventing and minimizing erosion and sediment discharge (i.e., source control and treatment BMPs) and pollutant category 2 (metals/ toxics /pesticides) should be controlled likewise, plus through infiltration to address dissolved constituents.

Based on these requirements, Caltrans relies on treatment and institutional BMPs to address all permit-listed pollutant categories completely or partially, using infiltration to supplement control of categories 2 and 4. In fact, Caltrans prioritizes fully infiltrating BMPs in its BMP selection process.

In the future, Caltrans may modify the constituents evaluated for BMP treatment comparisons based on changes in control strategies, improved monitoring strategies, or if regulations specify appropriate surrogates (e.g., assume sediment removal is representative of nutrient removal).

4.0 Framework for Load-Based Comparisons and Classification

The most recent Caltrans BMP performance ranking (Addendum #1, Caltrans 2018) placed each approved BMP into one of four tiers for each constituent based on the observed load reductions:

- Tier 0: Superior BMPs (100% load reduction of design storm)*
- Tier 1: More effective BMPs (>60% load reduction observed in pilot study)**
- Tier 2: Less effectives BMPs (20-60% load reduction observed in pilot study)**
- Tier 3: Not effective BMPs (<20% load reduction observed in pilot study)***

**The 2019 PPDG directs designers to first consider BMPs that fully infiltrate, but does not use the term "Tier 0"*

***55% threshold is used for total phosphorous (Caltrans 2010a)*

****Tier 3 does not appear in the 2017 PPDG (because they are ineffective).*

Note that the observed load reductions for each BMP are dependent on the site-specific volume reductions resulting from infiltration. Therefore, for any particular constituent, a BMP may appear in multiple tiers, depending on the BMP's observed volume reduction. For example, in a pilot study that monitored unlined and lined versions of a BMP, the unlined BMP may fall into Tier 1, while the lined version of that BMP may fall into Tier 2. Lining the BMP reduces the volume reduction contribution, resulting in a drop in the overall load reduction from Tier 1 to Tier 2. Caltrans therefore categorizes BMPs not only by observed load reductions, but also by the following volume reduction ranges:

- <20% volume reduction (infiltration)
- 20-50% volume reduction (infiltration)
- >50% volume reduction (infiltration)

Table L 1 shows an example of the approved BMPs categorized by BMP performance tiers for TSS load reductions, as documented in Addendum #1 of The Performance-Based Ranking for Treatment BMPs (Caltrans 2018f). The red, bold font indicates the candidate BMPs evaluated in 2018 (the other BMPs in Table L 1 were based on the 2010 ranking). Note that although no Tier 0 BMPs are shown in the table, Caltrans current practice, as documented in the 2019 PPDG, directs designers to first consider BMPs that can fully infiltrate the design storm (i.e., Tier 0 BMPs). Also, note that the 2019 PPDG does not include BMPs that are not effective, so no Tier 3 BMPs appear in the PPDG.

Table L 1 Example of Load-Based BMP Performance Categorization for TSS (Caltrans 2018)

Tier	Infiltration < 20%	Infiltration 20%–50%	Infiltration > 50%
1	**AMF: All** **Bioretention** Delaware Filter Detention Basin: Unlined MCTT **OGFC** Strip: As/Ad > 0.2 Strip: As/Ad < 0.1 Wet Basin	**AMF: All** **Bioretention** Delaware Filter Detention Basin: Unlined MCTT **OGFC** Strip: All Swales Wet basin	**AMF: All** **Bioretention** Delaware Filter Detention Basin: Unlined MCTT **OGFC** Strip: All Swales Wet basin
2	- - -	- - -	- - -
3	Detention Basin: Lined	Detention Basin: Lined	Detention Basin: Lined

Notes:

1. This table is adapted from Table 4.2, 2010 PBR TBMP (Caltrans 2010a).

[1]*AMF = Austin media filter; AA = activated alumina; MCTT = multi-chambered treatment train; OGFC = open-graded friction course; As = area of the strip; Ad = drainage area*
[2]*Red text indicates candidate BMPs that were evaluated in 2018.*

5.0 Steps for Load-Based Comparisons and Classification

This section presents steps to compare and classify BMPs based on load-reduction performance. This assumes that one BMP type (e.g. BMP A) was studied at multiple locations (BMP A1, BMP A2, BMP A3). Figure L 1 provides a flow chart.

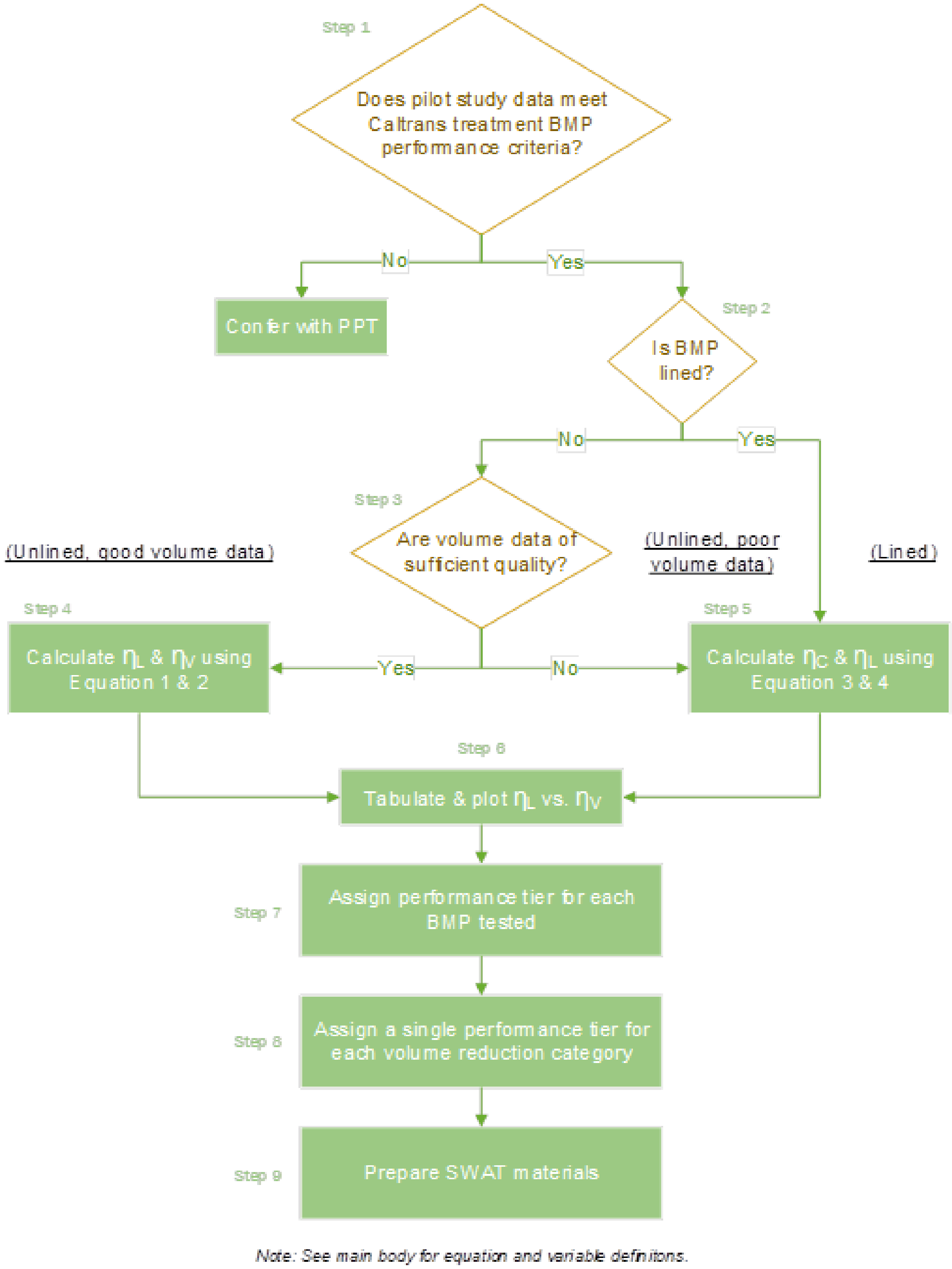


Figure L 1 Steps for Load-Based Ranking of Candidate BMPs

Step 1. Confirm pilot study data meet the Caltrans treatment BMP performance criteria. Section 4.2.1.3 of Caltrans Stormwater Management Plan (SWMP, Caltrans 2016) states that the criteria are listed in Appendix A of Caltrans Treatment Technology Report (Caltrans 2018b). Table L 2 lists the specific criteria from the Treatment Technology

Report and the table notes provide relevant statistical sections in the PSGM (Caltrans 2021).

If all performance criteria are met, go to Step 2. If statistically significant treatment or other criteria were not met, confer with the Pilot Project Team (PPT, which is described in Section 1 of the PSGM, Caltrans 2021) to reconsider recommendations for approval.

Table L 2 Treatment Performance Criteria

Treatment Performance Criteria[1,2]

• Study provides results from full-scale field testing of a stabilized (erosion-free) post-construction transportation-related impervious drainage area.
• Sampling and analysis was conducted in accordance to the Caltrans Stormwater Monitoring Guidance Manual, or other recognized protocol, such as the Urban Stormwater Runoff Monitoring Guidance (USEPA et al. 2009).
• Testing was conducted at flow rates and volumes typical of Caltrans drainage areas (areas vary, but usually are between 0.1 and 15 acres. Flow and volumes can be found by using Caltrans Basin Sizer).
• Mean (influent) concentrations from study were below the 90th percentile of statewide characterization data.
• Data was collected from at least eight storm events over a minimum period of two years and demonstrate a statistically significant removal ($p \leq 0.1$), which may require monitoring additional storm events.[3]
• Particle size distribution (PSD) during study was similar to the proposed field conditions (e.g., state whether traction sand was applied).
• The study's mean removal estimate corroborates the performance claim.
• The rainfall record for the study area or its vicinity during the evaluation period is documented and reported with the study results.
• Not all of the above criteria are met, but at least one of the following apply:
  o Statistically significant (p-value ≤ 0.1) constituent removal was established from independent stormwater field monitoring for at least one year[4]
  o Removal efficiency based on best professional evaluation of unit operations and processes that are well established for treatment of other waters
  o Load reduction of nutrients or BOD due to partial infiltration
  o Statistically significant (p-value ≤ 0.1) constituent removal was established from independent laboratory testing that follows the Technology Assessment Protocol – Ecology (TAPE) from Washington State (ECY 2008), and testing used a volume of water equivalent to one year of runoff for a typical installation. Alternatively, a laboratory loading using actual stormwater could be used as with the Tahoe Small Scale Research Facility

[1]Performance criterion applies to statewide approvals. Influent data for regional BMP approval may be compared to regional datasets (Caltrans 2009).
[2]The 2016 SWMP states "Appendix A of the Treatment Technology Report Provides specific performance criteria that are used for BMP evaluation." Criterion taken from Caltrans 2018b.
[3]Refer to PSGM Appendix D7, How to Compare Two Paired Data Sets (Caltrans 2021)
[4]Refer to PSGM Appendix D6, How to Compare Two Independent Data Sets (Caltrans 2021)

<u>Step 2.</u> If the BMP is lined, skip to Step 5. If the BMP is unlined, go to Step 3.

<u>Step 3.</u> Unlined BMP:

Check the quality of the pilot study's volume data (the quality of the concentration data were checked in Step 1). Recall that the volume data are calculated from the measured flow data, whose quality is evaluated using Caltrans Hydrologic Analysis Tool. Volume and flow aspects that are checked for quality include:

- Total event influent volume > total event effluent volume
- Any falsely elevated flow measurements due to bubbler clogs or flume/conveyance clogging are negligible

If the volume data are of sufficient quality, proceed to Step 4. If the volume data are not of sufficient quality, skip to Step 5.

<u>Step 4</u>. Unlined BMP with good quality concentration and volume data:

Calculate the load reduction efficiencies (η_L) for each candidate BMP location using Equation 1:

Equation 5-1

$$\eta_L = \frac{\sum C_i V_i - \sum C_e V_e}{\sum C_i V_i} * 100\%$$

Where:

C_i, V_i = influent concentration and volume, respectively, for each storm event (i)
C_e, V_e = effluent concentration and volume, respectively, for each storm event (i)
η_L = load reduction efficiency

Then, calculate the volume reduction efficiencies (η_v) for each candidate BMP location using Equation 2. If the calculated volume reduction efficiencies cover all three infiltration categories (less than 20%, 20 to 50%, and greater than 50%), then go to Step 6. If volume reduction efficiencies are missing for one or more of these categories, proceed to Step 5 to estimate what the load reduction would have been if volume reduction had occurred in this category.

Equation 5-2

$$\eta_v = \frac{\sum V_i - \sum V_e}{\sum V_i} * 100\%$$

Where:

V_i = influent volume for each storm event (i)
V_e = effluent volume for each storm event (i)
η_v = volume reduction (infiltration) efficiency

Step 5. Lined BMPs, unlined BMPs with poor volume data, or unlined BMPs that are not represented in all three infiltration categories:

Calculate the concentration reduction efficiencies (η_c) for each unlined BMP tested where the volume data are of poor quality or non-existent, or each lined BMP tested. Equation 3 shows the calculation.

Equation 5-3

$$\eta_c = \frac{\sum C_i - \sum C_e}{\sum C_i} * 100\%$$

Where:

C_i = influent concentration from each storm event (i)
C_e = effluent concentration from each storm event (i)
η_c = concentration reduction efficiency

Then, calculate the load reduction efficiencies (η_L) using Equation 4.

For lined BMPs that will always be lined, η_v in Equation 4 is assumed to be 0% (ignoring evapotranspiration losses) making η_L equivalent to η_c. In this situation, concentration reductions will represent load reductions for the less than 20% volume reduction category and the 20-50% and greater than 50% volume reduction categories are not applicable.

For lined BMPs that could be constructed without a liner, calculate η_L for each volume reduction category (i.e., less than 20%, 20 to 50%, and greater than 50%) using the lowest volume reduction efficiency (η_v) for each category (i.e. use values of 0%, 20%, and 50%).

For unlined BMPs that have poor flow data, calculate η_L using the lowest values (η_v) from the volume reduction categories for which the data are poor (0%, 20%, and/or 50%).

For unlined BMPs with good flow data where volume reduction efficiencies are missing for a particular volume reduction category, calculate η_L using the lowest volume reduction value for the relevant category (again, 0%, 20%, 50%).

Equation 5-4

$$\eta_L = \eta_C + \eta_v - \eta_C \cdot \eta_v$$

Where:

η_c = concentration reduction efficiency calculated from Eq. (3)

η_v = 0.0, 0.20, or 0.50, corresponding to assumed volume reduction (infiltration) efficiency [0%, 20% or 50%]*

η_L = load reduction efficiency

Source: Currier and Bonham, 2019

**Note: If BMP is always lined and no infiltration is allowed, only calculate η_L for n_v = 0.0; there will be no volume reduction. Per Eq. (4), for n_v = 0.0, n_L will have the same value as n_c.*

Go to Step 6.

Step 6. Tabulate and plot the load reduction efficiency versus the volume reduction (infiltration) efficiency for each candidate BMP location.

Note that if you used Equation 4 to calculate the load reduction efficiency for an unlined BMP with poor quality volume data or a BMP was lined but it could be constructed without a liner, you will plot those efficiencies against the volume reduction efficiencies you assumed for that calculation (i.e., 0%, 20%, and 50%). If you calculated the load reduction efficiency using Equation 4 for a BMP that is always lined, you will only have data for n_v = 0.0. Using values for n_v = 0.2 and 0.5 is only appropriate to model the performance of a BMP that can be constructed in an unlined condition.

Go to Step 7

Step 7: Using the table and plot, determine and tabulate the applicable T-1 Checklist performance tier for each BMP tested:

- Tier 0: Superior BMPs (100% load reduction of design storm)
- Tier 1: More effective BMPs (>60% load reduction efficiency)*
- Tier 2: Less effectives BMPs (20-60% load reduction efficiency)*
- Tier 3: Not effective BMPs (<20% load reduction efficiency)

**55% threshold is used for total phosphorous (Caltrans 2010a)*

Note that the next step will assign a single performance tier to the candidate BMP type for each volume reduction category. Due to variability of the data, the BMPs tested can fall into different tiers, even for the same volume reduction (infiltration) category (e.g., BMP A1 falls into Tier 1 and BMP A2 falls into Tier 2 for the 20-50% volume reduction category). The next step will address such conflicting results.

Go to Step 8.

Step 8: For each volume reduction category, check if the tested BMPs fall into more than one performance tier (e.g., BMP A1 falls into Tier 1 and BMP A2 falls into Tier 2 for the 20-50% volume reduction category). If there is only one BMP in a volume reduction category, assign the BMP that performance tier for that volume reduction category.

If there is more than one performance tier for a volume reduction category, examine the influent data for the lower tier classification and decide whether it should be dropped from consideration based on atypical conditions (e.g., influent data below Caltrans statewide discharge data). If so, for that volume reduction category classify the candidate BMP in the lowest tier after dropping the BMPs with atypical issues. Classify the candidate BMP tier as the lowest remaining tier of all the BMPs tested for that volume reduction category. For example, if for the 20-50% volume reduction category, candidate BMP A1 falls into Tier 1 and BMP A2 falls into Tier 2, but the influent concentrations for BMP A2 were lower than typical Caltrans influent concentrations, BMP type A would be classified as Tier 1. If the influent data were typical of Caltrans historic data, and no other reasons for reclassification are justified, BMP type A would be classified as Tier 2.

Go to Step 9.

Step 9: Present load vs. volume reduction graphs and draft tiers in the SWAT package. Show any performance reclassifications in the graphs and explain reclassifications in footnotes to the graphs.

5.1. Load-based Classification Example

The example herein shows how to assign performance tiers assuming multiple types of candidate BMPs were evaluated in pilot studies. Assume that two types of unlined BMPs were studied (BMP A and BMP B). Three type A BMPs and three type B BMPs were monitored (for a total of six BMPs), each at a different location (i.e., locations A1, A2, A3, B1, B2, and B3). Consequently, the individual BMP notations are A1, A2, A3, B1, B2, and B3. Assume that the storm-specific study data (not shown) were used in Equations 1, 2, and 3 above to calculate the volume, load, and concentration reduction efficiencies for each BMP studied as shown in Table L 3.

Note that Type B BMPs only fall into the 20-50% and >50% volume reduction categories; no Type B BMPs fall into the 0-20% volume reduction category. Therefore, to estimate load reductions for this poorly infiltrating category, Equation 4 was used to calculate load reduction efficiencies using an assumed 0% volume reduction efficiency for all three concentration reduction efficiencies given in Table L-3. The resulting load reduction efficiencies for the 0-20% volume reduction category are presented in Table L 4.These load reduction efficiencies represent the expected performance of the Type B BMPs for the hypothetical scenario of lining the installations at locations B1, B2, and B3. These load reduction efficiencies also represent the expected performance of unlined Type B BMPs that are constructed in conditions that are expected to achieve less than 20% volume reduction.

Table L 3 Example of Volume, Load, and Concentration Reduction Efficiencies for Total Copper for Two Types of Candidate BMPs

Candidate BMP Type	Candidate BMP Location	Volume Reduction (Infiltration) Efficiency (η_v)	Load Reduction Efficiency (η_L)	Concentration Reduction Efficiency[a] (η_c)
A	A1	63%	92%	79%
A	A2	49%	69%	39%
A	A3	15%	50%	41%
B	B1	75%	74%	-4.0%
B	B2	44%	72%	50%
B	B3	27%	56%	40%

[a] Concentration reduction efficiencies for BMP types A and B are shown but were not used in this ranking example because of the availability of good quality, storm-specific volume and concentration data. Concentration reduction efficiencies are shown to demonstrate that load reduction can occur with low or negative concentration reductions.

Table L 4 Example of Volume, Concentration, and Load Reduction Efficiencies for Total Copper for Lined Type B BMPs assuming no Volume Reduction

Candidate BMP Type	Theoretical Lined Candidate BMP Location	Theoretical Volume Reduction (Infiltration) Efficiency (η_v)	Concentration Reduction Efficiency (η_c)	Load Reduction Efficiency (η_L)
B	B1	0%	-4%	-4%
B	B2	0%	50%	50%
B	B3	0%	40%	40%

Figure L 2 shows the results of the load-based comparison for total copper for the Type A, Type B, and (theoretical) Type B-lined BMPs. Table L 5 shows the resulting classification. Candidate A BMPs with >50% volume reduction (location A1) and 20-50% volume reduction (location A2) were both classified as Tier 1 because the BMP tested for each volume reduction category fell into those tiers. The Candidate A BMP for the 0-20% volume reduction category (location A3) fell into Tier 2 and was therefore assigned that tier. The one Candidate B BMP that had >50% volume reduction (location B1) had a load reduction greater than 60%, so it fell into Tier 1 and was assigned accordingly. Two Candidate B BMPs had 20 to 50% volume reduction (locations B2 and B3). One of these had a load reduction greater than 60% (location B2), resulting in a Tier 1 category, while the other had a load reduction less than 60% (location B3), resulting in a Tier 2 category (blue triangles in Figure L 2). Candidate B was therefore assigned Tier 2 for the 20 to 50% volume reduction category in Table L 5 because that was the lowest tier of the locations tested for that BMP type (data review did not identify a reason for neglecting the data). Finally, although no type B BMPs tested had less than 20% volume reduction, the theoretical lined BMPs (which assumed a 0% volume reduction) were used (i.e., the green triangles in Figure L 2; theoretical locations B1, B2, & B3). The lowest concentration reduction was -4% (theoretical location B1), so the load reduction can be assumed to be this same amount, resulting in a Tier 3 category (i.e. not effective). The <20% volume reduction range for Candidate B was therefore assigned Tier 3. Ranking based on theoretical treatment should proceed with caution. Where a theoretical BMP performance at a given location is substantially different from that at other locations, as is the case with theoretical B1 in Figure L 2, the data for the "outlier" BMP location should be reviewed for representativeness. For example, the poor concentration and theoretical load reduction could of have resulted from unrepresentative monitoring of newly installed filter media that is capable of leaching or atypically low influent concentrations. This type of issue should have been caught during or before Step 1, when concentration data were reviewed for representativeness and accordance with Caltrans treatment performance criteria.

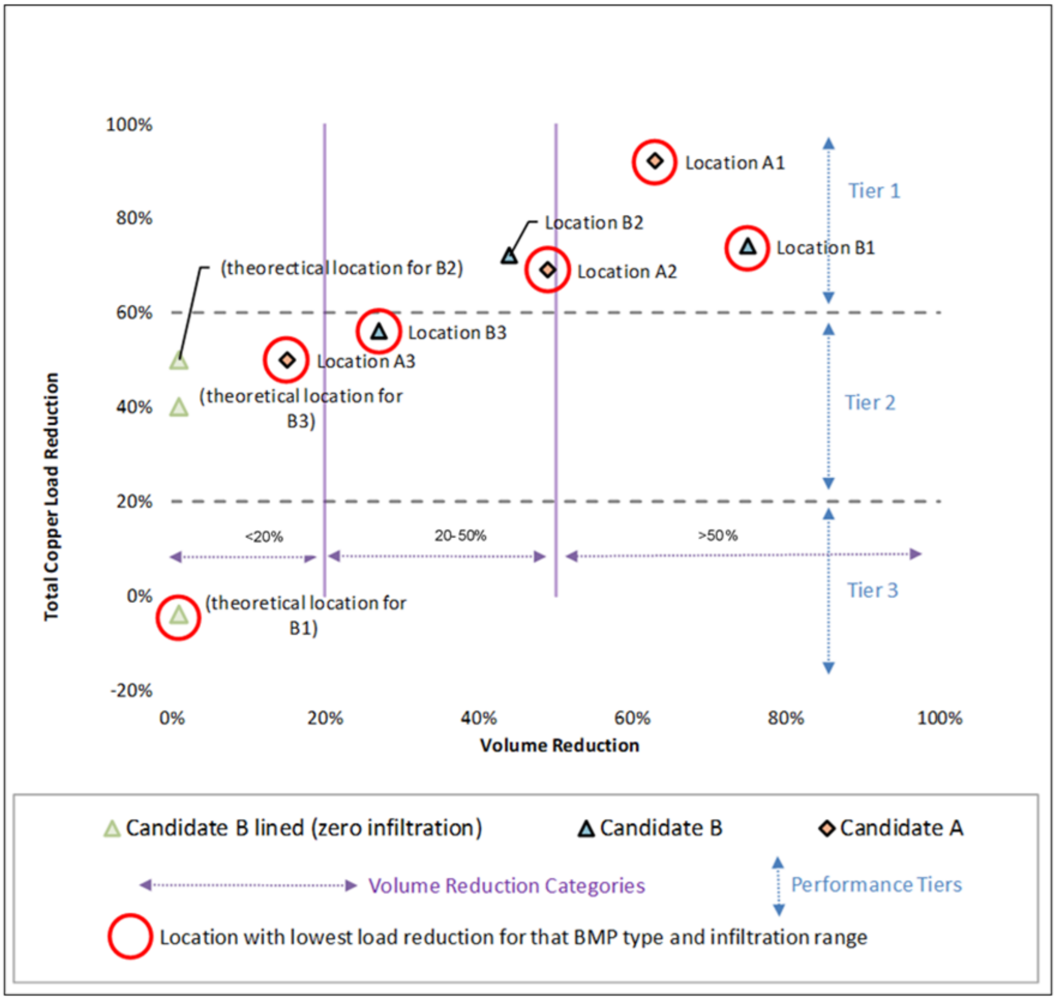


Figure L 2 Example of Total Copper Comparisons

Table L 5 Example of Recommended Performance Tiers for Total Copper

Volume Reduction Range	Candidate BMP A	Candidate BMP B
< 20%	Tier 2	Tier 3*
20%–50%	Tier 1	Tier 2
> 50%	Tier 1	Tier 1

**Classified as Tier 3, not effective, based on poor concentration reduction estimates. If this classification is adopted, Candidate B would not appear in the PPDG T-1 checklist performance tiers for that volume reduction category because the PPDG does not consider Tier 3. Also, note that candidate B had one location (B1) that passed the initial evaluation step of statistically significant load removal even though concentration reduction at that location was negative.*

6.0 Steps for Concentration-Based Comparisons and Classification

This section presents steps to compare and classify BMPs based on concentration-reduction performance. The primary purpose of concentration-based comparisons is to confirm the load-based ranking developed in the previous section for the <20% volume reduction category.

While all candidate BMPs are compared based on load-reduction performance, not all of them need to be compared based on concentration-reduction performance. *Concentration-based* analysis is most important in situations where volume reduction is minor and the majority of treatment is provided by concentration reduction. Volume reduction is considered minor for the <20% volume reduction category in the previous load-based section. The ranking approach used in the previous section (i.e., the load-based ranking) assumed that the candidate BMPs experienced typical Caltrans influent quality and that any differences in influent quality did not have a major effect on relative treatment performance. This assumption may not be valid for the <20% volume reduction category because *load-based* analysis depends on both volume reduction and concentration reduction, and when volume reduction is low, concentration differences will have a greater impact.

Concentration-based comparisons should be considered for candidate BMPs in the <20% volume reduction category that have load reduction efficiencies that are dependent on influent quality. BMPs that have load reduction efficiencies dependent on influent quality are those that show minor slopes in linear regression plots of influent concentration vs. effluent concentration.

Candidate BMPs are expected to experience differences in influent quality because stormwater runoff can vary one location to another due to dissimilar drainage area characteristics, changes in site conditions over time such as ADT, material accumulation, etc., and other factors. The effect of these differences in influent quality on the treatment performance (and ranking) of a candidate BMP largely depends on the constituent of concern and the treatment processes the BMP uses. For example, Caltrans monitoring has shown that BMPs that utilize filter media (i.e., rely on filtration as the treatment process) provide more consistent effluent quality than those that rely on sedimentation. Although a BMP based on filtration is expected to provide more consistent effluent quality, its treatment performance in terms of load reduction efficiencies is likely to be more sensitive to influent quality. This is because calculation of load reduction efficiency is based on the difference between influent and effluent quality, and as influent concentration increases the difference will increase if effluent quality is consistent.

Figure L-3 shows a stepwise approach for performing concentration-based analysis.

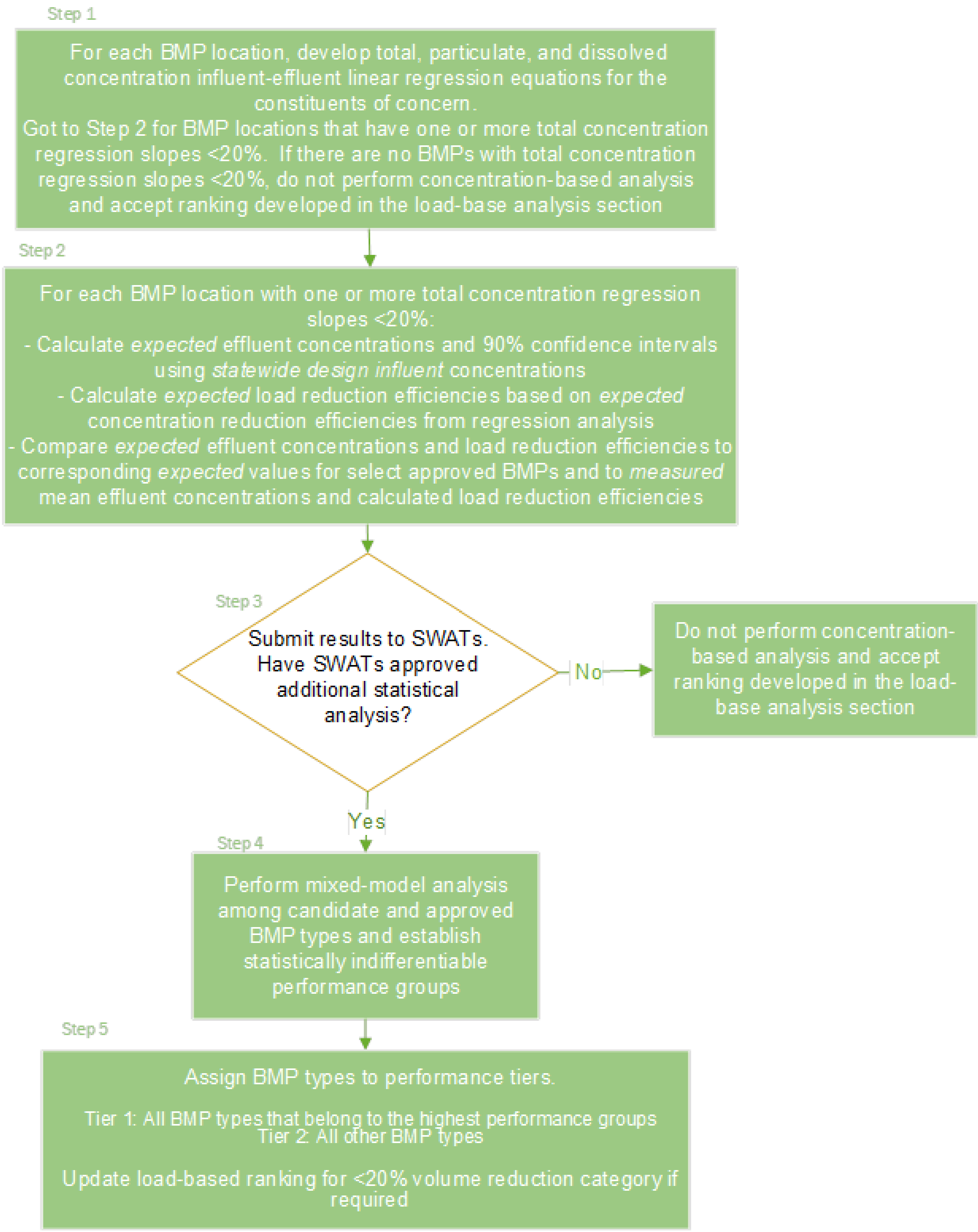


Figure L 3 Steps for Concentration-Based Ranking of Candidate BMPs

Step 1. For each candidate BMP location in the <20% volume reduction category from the previous section (i.e., the load-based ranking), develop regression equations of influent concentration versus effluent concentration for the constituents of concern based on event data collected during the pilot study. Follow the methodology presented in Appendix D of the PSGM (Caltrans 2021).

The regressions should be developed for total, particulate, and dissolved concentration data. Particulate and dissolved concentrations have very different degrees of reduction depending on the removal mechanisms within a BMP, so evaluating them separately will aid determination of the significance of different treatment processes.

Identify BMPs with one or more total concentration regression slopes that are minor, i.e., BMPs that have total concentration regression slopes that are not substantially different from zero. Assume regression slopes of <20% are minor.

Go to Step 2 if one or more candidate BMP locations have minor total concentration regression slopes (<20%). If no candidate BMPs have minor total concentration regression slopes, do not perform concentration-based analysis and accept the ranking developed in the previous section (i.e., the load-based ranking) for the <20% volume reduction category.

Step 2. For each candidate BMP location and constituent of concern, use the regression equations for that BMP to calculate *expected* effluent concentrations and 90% confidence intervals at the appropriate *statewide design influent concentrations* shown in Table L 6. Then, use the *expected* effluent concentrations and the *statewide design influent concentrations* to calculate *expected* concentration reduction efficiencies.

Next, using Equation 4 from the previous section calculate the *expected* load reduction efficiencies using the *expected* concentration reduction efficiencies from the regression analysis.

Compare the *expected* effluent concentrations and 90% confidence intervals for the candidate BMPs to the corresponding *measured* mean effluent concentrations and 90% confidence intervals for the load-based analysis in the previous section and to the corresponding *expected* effluent concentrations and 90% confidence intervals for approved sand filters and swales (Table L-6). [Measured mean effluent concentrations may need to be calculated if not available.] Also compare the *expected* load reduction efficiencies to the corresponding *measured* load reduction efficiencies from the previous section.

Go to Step 3.

Step 3. Submit the results from Steps 1 and 2 to the SWATs. The regressions and comparisons of *measured* and *expected* effluent concentrations and *measured* and *expected* load reduction efficiencies are informational so that SWATs may decide if further detailed statistical analysis is needed (Step 4). The SWATs may request other comparisons, such as grouping effluent quality from all BMPs of the same type (as opposed to the performance of individual BMPs). *Expected* effluent concentrations and associated regression equations for other approved BMPs and constituents are available in the BMP Retrofit Pilot Program Final Report (Caltrans 2004).

Proceed to Step 4 if SWATs request detailed statistical analysis. If the SWATs do not request additional statistical analysis, do not perform concentration-based analysis and accept the ranking developed in the previous section for the <20% volume reduction category.

Table L 6 Design Influent Concentrations and Expected Effluent Concentrations for Comparing Candidate BMPs to Select Approved BMPs

Constituent (units)[1]	Statewide Design Influent Concentrations[1]	Austin Sand Filter Expected Effluent Concentration (+/- 90% C.I.)[2]	Swale Expected Effluent Concentration (+/- 90% C.I.)[2]
TSS (mg/L)	49	7.8 (6.6, 9.0)	32 (22, 42)
Nitrate as Nitrogen (mg/L)	0.19	0.55 (0.41, 0.68)	0.22 (0.03, 0.40)
TKN (mg/L)	0.90	0.43 (0.27, 0.59)	1.1 (0.81, 1.4)
Part. P	0.26[3]	0.07 (0.05, 0.09)	0.22 (0.11, 0.33)
Ortho-phosphate	0.071	0.04 (0.09, 0.40)	0.40 (0.28, 0.52)
Part. Cu (μg/L)	7.5[4]	2.0 (1.4, 2.6)	3.7 (2.3, 4.9)
Diss. Cu (μg/L)	7.4	7.3 (6.5, 8.1)	7.4 (5.7, 9.1)
Part. Zn (μg/L)	37[4]	11 (7.9, 14)	18 (11, 25)
Diss. Zn ((μg/L)	35	19 (12, 26)	22 (9.5, 34)

[1] Use these values in Step 3 to calculate expected effluent concentrations from the influent-effluent regression equations developed for candidate BMPs from Step 2. These values are geometric means that were calculated as the transform of mean of ln(concentrations) for locations at elevations lower than 4,000 ft. (Caltrans 2009, Appendix C).

[2] Expected effluent concentrations and confidence intervals were recalculated from those presented in the Caltrans BMP Retrofit Pilot Program Final Report (Caltrans 2004) by using the design influent concentrations in this table. Regression and uncertainty equations for additional BMP types and constituents are also available in Caltrans BMP Retrofit Pilot Program Final Report (Caltrans 2004).

[3] The statewide design influent for particulate phosphorus is taken from the Retrofit Pilot study (2004) because particulate was not analyzed in the statewide Discharge Characterization Report Update (2009).

[4] Influent concentrations for particulate fractions were calculated by subtracting the antilog-transform of the average log-transformed dissolved value from the antilog-transform of the average of all the log-transformed total values, and thus are only an approximation of the geometric mean.

<u>Step 4</u>. The statistical analysis in this step should be conducted at the request of the SWATs when there are substantial differences between *expected* and *measured* load reduction efficiencies and *expected* and *measured* effluent concentrations for a candidate BMP. This statistical analysis requires use of mixed-models.

A mixed-model statistical analysis is used when data have more than one major source of variability. In BMP performance ranking, variability in influent concentration as well as BMP type might substantially impact treatment results. The mixed-model analysis controls for the effect of variation in BMP type and influent concentration on BMP effluent quality. It analyzes BMP effluent quality about a single influent concentration. Post hoc analysis is conducted after the mixed-model analysis to group BMP types having statistically indifferentiable (referred to as equivalent herein) effluent quality. The post hoc analysis compares the effluent quality of each BMP type to all other BMP types. Common post hoc analysis methods worth consideration include Fisher's least-squares difference (LSD), Benferroni, and Tukey's Honestly Significant Difference (HSD). For comparing among more than a couple (i.e., 3 or more) BMP types, use methods that calculate an appropriate pairwise error rate rather than using the family error rate (USGS 2020). HSD is preferred, because BMP studies will likely require comparison among many BMP types with unequal group sample sizes.

Appendix D of the PSGM (Caltrans 2021) contains guidance on using regressions (Appendix D9) and developing a mixed model (Appendix D11) to account for differences in performance due to differences in influent quality. Many statistical packages can integrate post hoc analysis into the model runs. Statistical textbooks and software tutorials are recommended for further details.

In the post hoc analysis, groups of statistically equivalent BMPs are constructed by considering each BMP type individually and identifying all the other BMPs types that are statistically equivalent. However, not all BMPs that are equivalent to the BMP being compared are equivalent to each other. Statistical software uses all pairwise comparisons to construct groups in which each BMP in that group is equivalent to every other BMP in that group. Each pairwise comparison is not typically displayed because it is a rather large matrix. The chosen statistical software typically collapses this larger matrix into groups where each BMP in the group is equivalent to all other BMPs in the group.

Table L 7 shows an example of the results of internal pairwise comparisons. The example draws from a mixed model analysis of TSS using ANCOVA with influent concentration as the covariate (Caltrans 2010a). ANCOVA assumes equal sample size. See Appendix D of the PSGM (Caltrans 2021) for further guidance on methods. All BMP comparisons are made within a single run of the statistical software.

An explanation of the pairwise comparisons made within statistical tools is aided by examining the results of the grouping in order of BMP performance. In the first column of Table L 7, candidate BMPs are numbered from best performing (1) to worst (6) (lowest effluent quality to highest). These are relative performance ranks based on estimates of effluent quality at a select influent concentration and they do not consider statistical differentiation. The BMPs are presented in order of point-estimate performance to aid in the evaluation of pairwise comparisons. BMPs that are equivalently performing (statistically indifferentiable effluent quality) to a candidate BMP are marked in each row of Table L-7.

Table L 7 Example of Categorizing Candidate BMPs into Concentration-Based Groups of Equivalent Performance[1, 2]

BMP Number by Performance (Effluent Quality) Rank	Candidate BMP	Wet Basin (BMP 1)	MCCT (BMP 2)	Delaware Filter (BMP 3)	Austin Filer (BMP4)	Strip HRT>5 (BMP 5)	Strip – HRT<5 (BMP 6)	EDB (BMP 7)	Swale (BMP 8)	Groups of Equivalent Performance by BMP Number	Performance Group
1	**Wet Basin**	X	✓	✓	✓	✓				12345	A
2	**MCTT**	✓	X	✓	✓	✓	✓			23456	B
3	**Delaware Filter**	✓	✓	X	✓	✓	✓	✓		3567	C
4	**Austin filter**	✓	✓	✓	X	✓	✓			23456	B
5	**Strip – HRT>5**	✓	✓	✓	✓	X	✓	✓	✓	5678	D
6	**Strip – HRT<5**		✓	✓	✓	✓	X	✓	✓	5678	D
7	**EDB**			✓		✓	✓	X	✓	5678	D
8	**Swale**					✓	✓	✓	X	5678	D

[1] Reproduced from Caltrans (2010a)

[2] For each candidate BMP identified in the 2nd column, the equivalently performing BMPs (BMPs that have statistically indifferentiable effluent quality) are indicated in columns 3 through 10 with a check mark. However, the group of equivalently performing BMPs (all BMPs equivalent to each other and not just the candidate BMP) must be confirmed by observation of the pairwise comparisons in the other rows for all the BMPs shown to be equivalent to the candidate BMP. For example, in the pairwise comparisons to the MCTT (BMP Number 2), it is observed that the MCTT is equivalent to wet basins (BMP Number 1) and Strip – HRT < 5 (short strips, BMP Number 6), and all BMP numbers in between. However, in the pairwise comparisons for wet basins (BMP Number 1), Strip – HRT < 5 (short strips, BMP Number 6) was determined not to be equivalent (the same result can also be observed in the pairwise comparisons to short strips). Consequently, wet basins are excluded from the performance group that it is identified in the row of pairwise comparisons for MCTT. This group is different from the group identified in the pairwise comparisons to wet basin in the previous row (Performance Group A), so this performance group is identified in the final column with the next consecutive letter (Performance Group B). Additional explanation is provided in Steps 4c to 4h.

The step-by-step process is explained below.

<u>Step 4a.</u> Starting with the wet basin (BMP Number 1), which was the best performing BMP in terms of effluent quality, there are no statistically significant differences in performance between the wet basin and the multi-chamber treatment train (MCTT), the Delaware filter, the Austin filter and longer strips (Strip – HRT>5, BMP Number 5). Because all five of these BMPs are statistically indifferentiable from each other, the BMPs

are identified as members of the same performing group, shown in Table L-7 as BMP Numbers 1, 2, 3, 4 and 5. This group is identified in Table L-7 as Performance Group A.

Step 4b. The next highest performing BMP in terms of effluent quality is the MCTT (BMP Number 2). As shown in the row corresponding to MCTT in Table L-7, MCTTs are found to be statistically indifferentiable from BMP Numbers 1,3, 4, 5 and 6. However, BMP Number 6 (short strip, Strip-HRT<5) is differentiable from wet basins (the wet basin box is unchecked in the row corresponding to short strips in Table L-7). Therefore, a new performance group is created that excludes wet basins. As a result, the new performance group comprises BMP Numbers 2, 3, 4, 5 and 6. This group is identified in Table L-7 as Performance Group B.

Step 4c. Next, from Table L-7, all BMPs that are statistically indifferentiable in terms of effluent quality from the third highest performing BMP, Delaware filters (BMP Number 3), are identified. As observed in the row corresponding to Delaware Filters in Table L 7, the indifferentiable BMPs are BMP Numbers 1, 2, 4, 5, 6, and 7. However, BMP Number 7 (EDB) is differentiable from wet basins, MCTT, and Austin filters (unchecked boxes in the row corresponding to EDB in Table L-7). Therefore, a third performance group is created that excludes these BMPs (BMPs 1,2,4). The resulting group comprises BMP Numbers 3, 5, 6, and 7. This group is shown in Table L-7 as Performance Group C.

Step 4d. Next, from Table L-7, all BMPs that are statistically indifferentiable in terms of effluent quality from the fourth highest performing BMP, Austin filters (BMP Number 4), are identified. As observed in the row corresponding to Austin filters in Table L 7, the indifferentiable BMPs are BMP Numbers 1, 2, 3, 5, and 6. However, BMP Number 1 (wet basin) is differentiable from BMP Number 6 (unchecked box in the row corresponding to short strips [Strip – HRT<5] in Table L-7). So BMP Number 1 must be excluded from the performance group. The resulting performance group is identical to that developed in the pairwise analysis for MCTTs, so no new performance groups are created. This group comprises BMP Numbers 2, 3, 4, 5, and 6. The group is shown in Table L-7 as Performance Group B.

Step 4e. Next, from Table L-7, all BMPs that are statistically indifferentiable in terms of effluent quality from the fifth highest performing BMP, long strips (Strip-HRT>5, BMP Number 5), are identified. As observed in the row corresponding to long strips in Table L 7, the indifferentiable BMPs are BMP Numbers 1, 2, 3, 4, 6, 7, and 8. Here, BMPs Numbers 1, 2, and 4 are differentiable from BMP Number 7 (row corresponding to EDB). Similarly, from Table L-7, BMPs Numbers 1, 2, 3 and 4 are differentiable from BMP Number 8 (row corresponding to Swale). Therefore, BMP Numbers 1, 2, 3, and 4 are excluded and the resulting group of equivalent performance comprises BMP Numbers 5, 6, 7, and 8. This group is shown in Table L-7 as Performance Group D.

Step 4f. Next, from Table L-7, all BMPs that are statistically indifferentiable in terms of effluent quality from the sixth highest performing BMP, short strips (BMP Number 6), are identified. As observed in the row corresponding to short strips (Strip – HRT<5) in Table

L 7, the indifferentiable BMPs are BMP Numbers 2, 3, 4, 5, 7, and 8. In this case, BMP Numbers 1, 2, 3, and 4 are differentiable from BMP Numbers 7 and 8 (rows corresponding to EDB and Swale). Therefore, BMP Numbers 1, 2, 3, and 4 are excluded. The resulting group of equivalent performance are BMP Numbers 5, 6, 7, and 8, which is the same as Performance Group D. Therefore, no new groups are needed as a result of these pairwise comparisons to short strips.

Step 4g. Next, from Table L-7, all BMPs that are statistically indifferentiable in terms of effluent quality from the seventh highest performing BMP, EDBs (BMP Number 7), are identified. As observed in the row corresponding to EDB in Table L 7, the indifferentiable BMPs are BMP Numbers 3, 5, 6, and 8. However, BMP Number 3 is differentiable from BMP Number 8 (row corresponding to Swale). Therefore, BMP Numbers 8 is excluded and the resulting group of equivalent performance comprises BMP Numbers 5, 6, 7, and 8, which is the same as Performance Group D. No new groups are needed as a result of these pairwise comparisons to EDBs.

Step 4h. Finally, in the pairwise comparison of swales (BMP Number 8), BMP Numbers 5, 6, and 7 are considered indifferentiable from swales in terms of effluent quality. This is the final pairwise comparison, so consideration of other pairwise comparisons is not needed. The resulting group of equivalent performance comprises BMP Numbers 5, 6, 7, and 8, which is the same as Performance Group D. Again, no new group is needed as a result of these pairwise comparisons to swales.

Table L 8shows the performance grouping that results from interpretation of the step-by-step pairwise comparisons. Each group in Table L 8 contains BMP types that are not differentiable at a 90% confidence level.

Table L 8 Categorizing Candidate BMPs into Concentration-Based Performance Tiers

BMP Number by Performance (Effluent Quality) Rank	Candidate BMP	Performance Group A	Performance Group B	Performance Group C	Performance Group D	Resulting Concentration Reduction Performance Tier
1	Wet Basin	X				1
2	MCTT	X	X			1
3	Delaware Filter	X	X	X		1
4	Austin Filter	X	X			1
5	Strip – HRT>5	X	X	X	X	1
6	Strip – HRT<5		X	X	X	2
7	EDB			X	X	2
8	Swale				X	2

Note that membership into performance groups does not always follow the hierarchy of performance in terms of effluent quality. For example, even though the point estimate for Delaware filter effluent is lower (i.e., better) than Austin filters, a higher uncertainty in performance by the Delaware filter results in classification within Performance Group C as well as A and B, while Austin filters are members of Performance Group A and B only. This could be a result of limited sample size, since only one Delaware filter was tested compared to five Austin filters. Within the post hoc analysis procedure illustrated above in Step 4a through 4h, all BMP types are compared to each other and BMPs with equivalent performance are collapsed into groups of performance. This exercise is performed within the statistical software. In Step 5, the resulting performance groups in Table L-8 are used to place BMPs into the performance tiers that were established by Caltrans in the PPDG to support BMP selection (Appendix E of Caltrans 2019e).

Step 5. Following the Caltrans practice to select BMPs according to performance tiers in the PPDG, classify the BMPs into one of two concentration reduction performance tiers. Traditionally, Tier 1 includes all of those BMPs that are statistically similar to the BMP with the highest performance, and Tier 2 are all other BMPs. Therefore, BMPs that fall into Performance Group A are assigned to Performance Tier 1, and all other BMPs are assigned to Performance Tier 2. Table L 8 reflects the application of this procedure.

The placement of the BMPs into concentration-based tiers is reported in the BMP study's SWAT package (Appendix K), and eventually incorporated into future versions of the

PPDG in accordance with Caltrans approvals and protocols. In the SWAT package, compare the concentration-based performance tiers with those developed for the <20% volume reduction category in the previous load-based analysis section. The SWATs will determine whether to keep the load-based reduction ranking or go with the concentration-based reduction ranking, if different.

A limitation of the mixed model approach is that rankings are based on a single point of comparison of influent quality that may not represent project-specific conditions (Caltrans 2010a). In the SWAT package, provide mean influent concentrations and 90% confidence intervals for the candidate BMPs to facilitate application of best professional judgment by the SWAT in determining BMP selection policy.