# Monitoring Guidance To Support Adaptive Watershed Management



Kenneth Schiff **Emily Darin** Elizabeth Fassman-Beck Gemma Shusterman Lorenzo Flores Tony Hale Chris Beegan

# Monitoring Guidance To Support Adaptive Watershed Management

Kenneth Schiff<sup>1</sup>, Emily Darin<sup>1</sup>, Elizabeth Fassman-Beck<sup>1</sup>, Gemma Shusterman<sup>2</sup>, Lorenzo Flores<sup>2</sup>, Tony Hale<sup>2</sup>, Chris Beegan<sup>3</sup>

<sup>1</sup>Southern California Coastal Water Research Project, Costa Mesa, CA

<sup>2</sup>San Francisco Estuary Institute, Richmond, CA

<sup>3</sup>State Water Resources Control Board, Sacramento, CA

March 2022

Technical Report 1257

#### SUMMARY

Adaptive Watershed Management (AWM) initiates with a rigorous planning process to identify the most effective combination of management actions to achieve watershed goals, followed by periodic evaluations of implementation success. Where implementation achieves progress toward watershed goals, managers can stay the course until the goals are achieved. Where implementation is not achieving progress toward watershed goals, managers can make decisions about what actions to change or modify in order to make progress. In essence, managers "adapt" their implementation strategies because planning does not always predict exact outcomes.

Fundamental to AWM is monitoring, which is the primary mechanism for tracking progress toward watershed goals. However, unlike the initial AWM planning phases, there is almost no statewide guidance on AWM monitoring, nor is there guidance on how to use the monitoring information for making AWM decisions.

The goal of this document is to provide regulated and regulatory agencies across California the guidance and some hands-on tools they need to design monitoring programs to provide this critical information for AWM decision making.

In this document, we capture the guidance of an AWM Advisory Committee (Advisory Committee) to define the highest priority monitoring questions, create web applications to focus monitoring design attributes, provide data visualization tools to clarify and comprehend monitoring output, and apply case studies as examples.

Because each agency and watershed is different, this document does not try to create a one-size-fits-all monitoring plan, nor does it try to create a standardized monitoring program. Instead, flexibility and site-specific adaptations are encouraged based on the unique needs of each watershed management program. For example, each web application is built for local agencies to upload their own data for making monitoring design decisions.

## **Monitoring questions**

Critical to the success of any monitoring program are the monitoring questions: What is it that managers want to know and what AWM decisions will they make once they have an answer? The Advisory Committee spent over 150 person-hours developing and ranking monitoring questions for AWM. The Advisory Committee included a diverse array of members from regulated, regulatory, and non-governmental agencies ranging from San Francisco Bay to San Diego.

Recognizing the challenge in developing management questions that may be relevant throughout the state, the Advisory Committee ultimately prioritized three management-style questions intended to address a variety of stormwater pollutants, control measures/locations, and regulatory or program settings:

- 1. What is the temporal trend in pollutant concentrations (or loading) from MS4 outfalls?
- 2. Which site is my worst site?
- 3. What is the treatment effectiveness of my BMP?

## **Temporal Trends Question**

Fundamental to AWM decision-making is change over time. Stormwater regulators require, and regulated stormwater agency staff rely upon, watershed planning documents. Once a watershed plan begins to get implemented, improvements in water quality are expected to follow.

While many scientists like to make statistics challenging, the concepts behind trend monitoring are relatively simple. Detecting a trend is a function of the amount of change you want to observe, the variability in the data, and the amount of time to detect the change. Larger changes, less data variability, and more time make trends easier to detect and require fewer samples. Smaller changes, lots of data variability, and shorter time periods make trends harder to detect and require more samples. Thus, the level of sampling effort – how many samples per year – becomes the predominant factor for trend monitoring design.

A web application created to assist regulated and regulatory managers when designing their trends monitoring program for supporting AWM decisions (<u>Trends App</u>). The trends app can support monitoring designs for most any indicator (chemistry, bacteria, trash) and can be used for a variety of locations (receiving water, outfalls, sources), but the app is designed to be used one location at a time. The trends app can be used for dry or wet weather, but should be used separately in each of the weather conditions.

Trends can either be decreasing, increasing, or not changing. The advisory committee provided these recommendations for actions associated with these three answers:

• None if trends are decreasing.

Managers should continue what they are doing since strategies are performing as expected to reach identified goals.

• Trigger enhanced monitoring as you approach the decision-making threshold.

This can apply if trends are increasing, decreasing, or not changing. Once you are close to a threshold, additional confidence in the answer may be necessary.

• Trigger BMP effectiveness evaluations if trends are increasing or not changing.

Managers should assess if the actions to date are not performing as expected.

• Trigger source identification if trends are increasing.

Managers need to invest more effort into defining the source(s) of their problem so they can more effectively remediate it.

• Use monitoring results for improved watershed modeling/planning.

Part of the problem when trends are increasing or not changing is perhaps the plan was insufficient to make the change anticipated. Managers can use the data collected to date for improving the watershed management plan or the watershed model used for creating the plan.

## **Hot Spot Question**

When approaching AWM decision-making, most managers want to know where their best and worst sites are located. The best sites are worthy of protection to ensure they don't degrade. The worst sites are prime candidates for where the management actions should be focused in the upcoming AWM cycle to improve water quality and achieve water quality goals. This section focuses on monitoring guidance for "hot spots" to define a watershed manager's worst sites, but the approach simultaneously derives "cold spots" to define the watershed manager's best sites.

A web application was created to assist regulated and regulatory managers when analyzing their hot spot monitoring program for supporting AWM decisions (<a href="https://docume.com/hot-spot-app">hot-spot-app</a>). The hot spot app can support monitoring designs for most any indicator (chemical, bacteria, trash, biology, volume). However, multiple sites need to be sampled multiple times, and a water quality threshold must exist.

Hot spots were defined as the frequency of water quality threshold exceedances, although the hot spot app also provides magnitude of exceedances. The advisory committee provided these recommendations for actions associated at sites with the greatest frequency of threshold exceedances:

- Additional follow-up sampling to confirm the timing and magnitude of water quality threshold exceedances.
- Plan and implement source tracking and identification monitoring to pinpoint remediation.
- Targeted inspection and enforcement activities.
- Targeted non-structural BMP implementation.
- Prioritized catchment for future structural BMP implementation.

## **BMP Effectiveness Question**

Structural Best Management Practices (BMPs) are one of the key tools used by managers to achieve water quality goals. These BMPs include treatment technologies such as biofilters, bioretention, infiltration dry wells, baffle boxes, and alike. These BMPs are used to reduce pollutant concentrations, reduce runoff volumes, or both. Managers expend tremendous effort to select the proper BMP for their Watershed Management Plans, select the proper design specifications for the BMP, and select the BMP location within the watershed to ensure optimal performance.

A web application was created to assist regulated and regulatory managers in assessing if their BMP is performing (BMP app). Assessing if structural BMPs are performing to expectations is a challenge because there is no universal standard for making these assessments. The BMP app incorporates both percent reduction and linear regression approaches, maximizing the advantages and minimizing the disadvantages of each approach alone. Ultimately, the effectiveness of a BMP for pollutant reduction is whether it is achieving the water quality threshold such as a standard, objective, action level, or TMDL target.

BMPs can either be performing well, performing, underperforming, or failing. The advisory committee provided these recommendations for actions associated with these three answers:

• None if the BMP is achieving the water quality threshold.

Managers should continue what they are doing since the BMP is performing as planned.

• Evaluate maintenance if the BMP is underperforming or failing.

Maintenance is critical for keeping a BMP operating at its maximum capacity for its design lifetime.

• Modify design specifications for construction or retrofit.

If a BMP is underperforming or failing, it is not operating as designed. First, managers will want to ensure the BMP was constructed to its specifications. Next, managers will want to conduct maintenance to ensure the BMP is operating as designed. If construction meets design specifications and the BMP is operating as designed, then modifying the BMP design specifications makes sense to improve performance for this (and future) BMPs. Additional AWM response options can include changing the type of BMP to be used or adding a BMP treatment train.

## **ACKNOWLEDGEMENTS**

This project would not have been possible without the Advisory Committee. A special note of gratitude goes to Committee members (from north to south): Tom Mumley, Zach Rokeach, Reid Bogert, Ian Wren, Karen Cowan, Ivar Ridgeway, Grant Sharp, Laurie Walsh, and Todd Snyder.

## TABLE OF CONTENTS

Summary	i
Monitoring questions	i
Temporal Trends Question	ii
Hot Spot Question	iii
BMP Effectiveness Question	iii
Acknowledgements	V
Table of Contents	vi
Introduction	1
Goal of this project	2
Methodological Approach	4
Create an Advisory Committee	4
Question development and prioritization	5
Monitoring design frameworks (iterative web app development)	5
Monitoring Questions	7
Temporal Trends	9
Temporal Trend Monitoring Question	9
Temporal Trend Monitoring Guidance	10
The Trends App	11
Case Study: Los Angeles County Wet Weather Receiving Water Monitoring	12
Case Study: Orange County Dry Weather Outfall Monitoring	14
BMP Effectiveness	16
BMP Effectiveness Monitoring Question	16
BMP Effectiveness Monitoring Guidance	17
The BMP Web App	19
Case Study: Biofilter	20
Spatial Hot Spot	24
Spatial Hot Spot Monitoring Question	24
Hot Spot Monitoring Guidance	25
The Hot Spot Web App	25
Case Study: Orange County Dry Weather Outfall Monitoring	26
Summary	29
References	30
Appendix A: Translating Priority Issues into Monitoring Questions	31
Appendix B: Trend Analysis Calculations	33
Power analysis	33

Optimal sample effort	34
Citations	34
Appendix C: Web App Tutorials	35

#### INTRODUCTION

Alternative Compliance Pathways (ACP) represent an increasingly popular approach to stormwater management. ACP implements stormwater best management practices (BMP) on timescales of 20 or more years to achieve compliance with watershed goals such as receiving water objectives and other NPDES permit conditions (OWP 2018). To achieve success across these time scales, ACP is based on the concept of Adaptive Watershed Management (AWM). AWM engenders a rigorous planning process to identify the most effective combination of management actions to achieve watershed goals, followed by periodic evaluations of implementation success. Where implementation achieves progress toward watershed goals, managers can stay the course until the goals are achieved. Where implementation is not achieving progress toward watershed goals, managers can make decisions about what actions to change or modify in order to make progress. In essence, managers "adapt" their implementation strategies because planning does not always predict exact outcomes.

In California, considerable guidance is available to watershed managers – both regulated and regulatory managers – for the planning phases of ACP. Sometimes called Reasonable Assurance Analysis, the planning phases typically constitute watershed modeling. These models may be relatively simple and straightforward and other times the models may be very complex and integrated (Paradigm 2017). Either modeling approach is attempting to model future water quality/water quantity outcomes based on a series of management actions.

Unlike Reasonable Assurance Analysis and model-based planning, comparatively little guidance is available to help watershed managers design or implement a routine field monitoring program that can effectively benchmark an ACP project's progress toward achieving its long-term water-quality improvement goals. There is no statewide guidance or clear-cut definitions for adaptive management decision-making criteria that can help inform managers on how and when they should make course corrections and other adjustments based on monitoring program data.

Currently, stormwater monitoring programs throughout the state vary widely in the designs, levels of effort, and insights provided by their monitoring (PG Environmental 2018). Some programs have hundreds of sites and are sampled many times during the year for a long list of contaminants, spending millions of dollars annually. Other monitoring programs are leaner, sampling a small number of sites and times during the year, spending only a fraction of the cost (Afrooz and Schiff 2018). In some cases, this discrepancy reflects the regulatory requirements for monitoring, which can be highly specific or broadly general, depending on the region of the state.

There is wide agreement that not all monitoring programs need to be or should be the same. They have different AWM goals, different landscape and climate settings, and/or different pollutants of concern. There is also wide agreement that monitoring programs should invest sufficient resources to answer their questions about what changes they should be making at periodic intervals, typically every five years or NPDES permit cycle.

## Goal of this project

The goal of this project is to provide regulated and regulatory agencies guidance to design effective and efficient monitoring programs to support AWM decision making. The project aims to provide tools for enhancing communication between regulatory and regulated parties when designing monitoring programs so that discussions can be about what decisions should be made based on monitoring results rather than arguing about what data should (or should have been) monitored.

The project provides three valuable resources to both regulated and regulatory staff:

- 1. Priority monitoring questions
- 2. Web applications for supporting technical design choices using your own data
- 3. Case studies as examples for agency staff on how to make monitoring design choices

A conceptual model of the AWM process (Figure 1) illustrates where this project supports decision making. To initiate the AWM process, water quality goals are established and documented in the Watershed Management Plan (e.g., WQIP, WMP, eWMP, etc.). The Watershed Management Plan then describes the different implementation activities to achieve these goals. A monitoring program is next designed and implemented to answer questions about progress toward those goals. The answers to those questions – either yes or no – define the adaptive components to watershed management as alternative management actions to improve progress toward the water quality goals.

This project focuses on the monitoring aspects of the adaptive watershed management process helping regulated and regulatory managers create meaningful monitoring questions, effective and efficient monitoring designs, and impactful data visualizations so that answers to the monitoring questions are clearly understood.

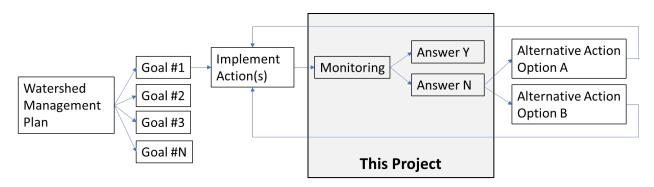


Figure 1. Conceptual model of adaptive watershed management process. Note that this project, indicated by the shaded box, focuses on the monitoring design and assessment portions of the process for assessing effectiveness of implementation actions.

While this project outlines this important part of the AWM process, it is not a cookbook. The questions defined are not the only questions watershed managers might ask and the management responses described are not the only adaptive actions that could be considered. The aim is to

provide sufficient guidance that the monitoring tools created can be used for most any indicator of concern in almost any watershed setting during wet and/or dry weather. This broad application provides productive and well-informed starting points of conversations among regulated and regulatory agencies, not a dictatorial step-by-step set of requirements.

#### METHODOLOGICAL APPROACH

This project followed a four-step approach to creating AWM monitoring design guidance:

- 1. Create an Advisory Committee
- 2. Define the list of monitoring questions
- 3. Create a web application to define monitoring guidance needs
- 4. Utilize a case study to optimize the web application

## **Create an Advisory Committee**

Fundamental to the success of this study was creating an AWM Advisory Committee (Table 1). While monitoring program study design has a technical foundation, ultimately managers use the monitoring information for management decision-making. Thus, the Advisory Committee was designed to include:

- Upper-level managers who are the individuals responsible for making the AWM decisions.
- Geographic diversity to include major municipal areas throughout California so that a full range of local issues are addressed.
- Sector diversity including Regional Water Quality Control Boards, regulated municipalities, and a non-governmental advocacy organization to ensure the monitoring information collected addresses all of the necessary perspectives for decision-making.

The charge of the Advisory Committee was to:

- Generate the primary monitoring questions for the SWRCB guidance.
  - In order to make the monitoring guidance as applicable and realistic as possible, the Advisory Committee was asked what adaptive management actions and activities might be taken once the monitoring question is answered.
- Shape the priorities for the monitoring program framework.
  - Assuming there are always more needs than resources, the Advisory Committee was asked to establish priority monitoring questions and the key pieces of AWM decision-making information.
- Build consensus on the final recommendations for the monitoring guidance.
  - The Advisory Committee was asked to reach consensus on all of the monitoring recommendations. Consensus across the diversity of sectors and geographies ensures the meaningfulness of the monitoring guidance.

Table 1. Advisory Committee membership.

Name	Organization
Tom Mumley	San Francisco Regional Water Quality Control Board
Reid Bogert	City/County Association of Governments, San Mateo County
Ian Wren	San Francisco Baykeeper
Karen Cowan	California Stormwater Quality Association
Ivar Ridgeway	Los Angeles Regional Water Quality Control Board
Grant Sharp	Orange County Public Works, Environmental Resources
Laurie Walsh	San Diego Regional Water Quality Control Board
Todd Snyder	San Diego County Watershed Protection Program

## Question development and prioritization

Monitoring questions are the most important part of designing a monitoring program. They can also be the most challenging because it is largely non-technical. Monitoring questions should be intrinsically tied to AWM concerns and specific management actions.

A well-established and impactful monitoring question is comprised of four parts, including:

- Spatial and temporal scale: the "where and when" to monitor
- Indicators: the "what" to monitor
- Benchmarks: the threshold(s) that define good *versus* bad monitoring outcomes
- AWM Action: defines what the managers, either from regulated or regulatory agencies, are going to do once they have an answer to the monitoring question. Managers for this study are defined as either Municipal NPDES Stormwater Permittee or RWQCB staff who are considering actions taken based on monitoring results.

The Advisory Committee spent three meetings – a total of nearly 15 hours – creating a list of monitoring questions, prioritizing which questions were the most important, and refining the questions to address their most pressing management needs. Then, for each of the prioritized monitoring questions, the Advisory Committee defined the four attributes above.

## Monitoring design frameworks (iterative web app development)

Web applications were developed as potential tools to help future users develop effective monitoring programs. The web apps were also an excellent tool for helping the Advisory Committee define what they thought regulatory and regulated parties would need in developing their monitoring programs. Ultimately, through the use of case studies, the web apps became the tool for the Advisory Committee to test drive decision making. This included what information they would need for making that decision, and how much confidence they would need in that information.

Web app creation followed four basic premises:

- The web apps should provide simple answers that non-technical users can understand
- The web apps should allow users to upload their own data to support developing locally-relevant monitoring programs
- The web apps should be flexible enough to allow the large range of climates, landscapes, and pollutants found in California

•	The web apps should be built with public domain software, and with sufficient documentation, so anyone can use it and modify it if they need to

## **MONITORING QUESTIONS**

The Advisory Committee created a list of their most important issues related to AWM (Table 2). The 39 issues fell into one of five categories:

- 1. Receiving water health the focus of most regulatory water quality criteria is on chemical and biological receiving water quality
- 2. Outfall monitoring links stormwater inputs from urban catchments to receiving waters including concentrations and loads
- 3. BMPs and load reductions describes maximizing effectiveness, efficiency, and longevity of management actions to mitigate stormwater concentrations and loads
- 4. Program administration wide range of activities, most of which are required in NPDES permits, to plan, invest, coordinate, and document stormwater program activities
- 5. Communication crucial for ensuring knowledge and endorsement of stormwater management activities at staff, inter-agency, intra-agency, executive management, and public levels

After creating this list of vital issues, the Advisory Committee translated them into management questions. This process can be seen in Appendix A. Issues such as "outfall monitoring for concentrations" become a management question such as "are outfall concentrations going up or down?" Clearly, this would be an important AWM focal point if one is thinking about the need for doing something different into the future. However, in order to build a monitoring program to inform a manager if outfall concentrations are changing over time, the management question needs more detail to get translated into a monitoring question. In the case of Appendix A, the monitoring question translation was "What is the trend in [pick indicators] at [pick outfall] between [pick time span]?" The increased specificity includes space and time (which outfall, which time span), an indicator (pick pollutant), and a threshold (trend).

After translating all of the critical AWM issues into management questions and then monitoring questions, the Advisory Committee prioritized which questions to focus on for monitoring guidance and web app development. These follow in the next three sections.

The issues prioritized by the Advisory Committee included:

- Trends in receiving water or outfalls
- Structural BMP effectiveness
- Spatial explicitness to detect hot spot outfalls or receiving waters

It is critical to acknowledge that the non-prioritized issues and questions by the Advisory Committee may still be important for local AWM programs and monitoring guidance could still be developed to address these management needs. However, not all of the monitoring would be traditional chemistry, toxicity or biology. For example, monitoring programs to assess communication may require social science approaches to quantify the information needed by managers for AWM. Also, the Advisory Committee felt that most regulated agency managers are already collecting and interpreting program administrative information needs. These include inspections, housekeeping, asset management, and alike. While some additional investment and documentation may be in order for these program administration information needs, general

statewide guidance is not necessary because remediation actions are very localized and sitespecific.

## Table 2. List of Critical Issues for the Advisory Committee.

#### Receiving Water Health

- Healthy stream ecosystems
- Public health protected
- Trash minimized
   Are receiving waters healthy/supporting BU
- 5. Improvements in water quality

#### **Outfall Monitoring**

- 6. Spatial explicitness loads, concentrations
- 7. Outfall concentrations compared to numeric thresholds
- Effluent relative to receiving water concentrations
- 9. Effluent numerics
- 10. Load reduction
- 11. Outfall impacts on receiving waters

#### BMPs and Load Reduction

- 12. BMP effectiveness
- 13. Multi-benefit improvement projects (i.e., WQ and flow modification)
- 14. Project implementation with multiple goals and partnership with multiple agencies
- 15. Source control
- 16. Post construction O&M effectiveness
- 17. Volume capture
- 18. Pounds of trash removal increasing over

#### **Program Administration**

- 19. Maximum return on investment (activities for achieving receiving water objectives)
- 20. Pollution prevention IC/ID, housekeeping

- 21. Inspections response
- 22. Action implementation completion and response - trust
  - a. Monitoring
  - b. **Enforcement actions**
  - c. Reporting
- 23. Leadership position of authority/chain of command/experience & knowledge
- 24. Account for activities
- 25. How effective at integrating program pieces, synergy
- 26. Accomplishments relative to planned activities
- 27. Management Reactions to persistent problems
- 28. Relative effectiveness of different approaches/strategies
- 29. Training and qualifications of staff
- 30. Project implementation with multiple goals and partnership with multiple agencies
- 31. Watershed area treated asset management
- 32. Implementation of monitoring activities
- 33. Amount of green infrastructure implemented or planned
- 34. Cost effectiveness of monitoring relative to information
- 35. Cost of implementation activities relative to effectiveness

#### Communication

- 36. Communication transparency, story or messaging to meet objective
- 37. Communication amongst parties
- 38. Community outreach success
- 39. Communicate progress

## **TEMPORAL TRENDS**

Fundamental to AWM decision-making is change over time. Stormwater regulators require, and regulated stormwater agency staff rely upon, watershed planning documents. Once a watershed plan begins to get implemented, improvements in water quality are expected to follow.

What is not always clear, and can be dramatically different from watershed to watershed, is how monitoring data is collected to assess if water quality improvements are occurring. This section describes how to refine the monitoring question, and what is most important in designing the monitoring program to answer a temporal trends question.

## **Temporal Trend Monitoring Question**

The four attributes of a temporal trend question are shown in Table 4. This design guidance can apply to outfall or receiving water sites, and it can apply to virtually any pollutant including chemistry, toxicity, biology, or even volume.

Trends can either be decreasing, increasing, or not changing. The Advisory Committee provided these AWM recommendations for actions associated with these five answers:

• None if trends are decreasing.

Managers should continue what they are doing since strategies are performing as expected to reach identified goals.

• Trigger enhanced monitoring as you approach the decision-making threshold.

This can apply if trends are increasing, decreasing or not changing. Once you are close to a threshold, additional confidence in the answer is necessary.

• Trigger BMP effectiveness evaluations if trends are increasing or not changing.

Managers should assess if the actions to date are not performing as expected.

• Trigger source identification if trends are increasing.

Managers need to invest more effort into defining the source(s) of their problem so they can more effectively remediate it.

• Use for improved watershed modeling/planning.

Part of the problem when trends are increasing or not changing is perhaps the plan was insufficient to make the change anticipated. Managers can use the data collected to date for improving the watershed management plan or the watershed model used for creating the plan.

Table 4. Defining the four attributes of the trends monitoring question.

Management Question: What is the temporal trend in pollutant concentrations (or loading) from MS4 outfalls?		
Space and time scale:	Space: Outfalls selected based on local input	
	Time: Number of samples per year, minimum one permit cycle	
Indicator(s) to be measured:	Various based on local regulatory issues:	
, ,	Bacteria, Metals, Nutrients, Sediment, Pesticides, PCBs, trash	
Benchmark for comparison:	Based on local input:	
·	TMDL Target(s), Water Quality Objective, Action Level	
Action to be taken once you have an answer:	Nothing if trends decreasing	
•	Triggers enhanced monitoring as you approach the threshold	
	Triggers source identification	
	Use for improving watershed modeling/planning	

## **Temporal Trend Monitoring Guidance**

While many scientists like to make statistics challenging, the concepts behind trend monitoring are relatively simple. Detecting a trend is a function of amount of change you want to observe, the variability in the data, and the amount of time to detect the change. Larger changes, less data variability, and more time make trends easier to detect and require fewer samples. Smaller changes, lots of data variability, and shorter time periods make trends harder to detect and require more samples. Thus, sampling effort – how many samples per year – becomes the predominant factor for trend monitoring design.

Defining the optimal sampling effort for detecting trends is described by power analysis. Power analysis incorporates amount of change and sampling effort into one chart (Figure 2). When few samples are collected only large changes can be detected. Small changes require the most samples. The "sweet spot" is at the inflection point of the curve, where fewer samples dramatically reduces a manager's ability to detect a trend and more samples does not substantially increase a manager's ability to detect the same trend. For the recommended guidance in this document, we will use power curves of detectable change and sampling effort, a set time interval of 10 years, and then let the users own monitoring results describe variability in the data.

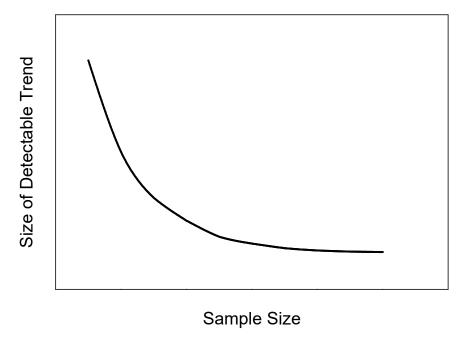


Figure 2. Idealized power curve. As sample size increases, smaller and smaller trends can be detected. After the inflection point, more samples do not dramatically improve trend detection. If managers cannot a priori define how much change they want to see, the inflection point defines the "optimal" sample size. In some cases, if the desired size of trend is small, a sufficient number of samples may never be collected.

## The Trends App

A web application was created to assist regulated and regulatory managers when designing their trends monitoring program for supporting AWM decisions (<u>Trends App</u>). The trends app can support monitoring designs for any indicator (chemical, bacteria, trash) as long as it is continuous data. The trends app can be used for a variety of locations (receiving water, outfalls, sources), but is designed to be used one location at a time. Finally, the trends app can be used for dry or wet weather, but should probably be used separately in each of the weather conditions.

The trends app has three elements: data upload and inventory, visualization of existing monitoring data, and power analysis for selecting optimal frequency. For a complete step-by-step tutorial, visit the trends app online documentation or use the tutorial in Appendix B of this report.

The data upload and inventory allows users to utilize their own data in the app. This is a key feature to ensure the sampling frequency estimates are specific to the user's location. User data can either be downloaded from CEDEN, or users can fill in the attached data template which mimics CEDEN formats. Either way, the data formats must match the template exactly or erroneous results may occur.

Once data is uploaded, an inventory of the data file is automatically created, helping ensure the app reads the users input files correctly. The inventory includes watersheds, sampling locations, parameters measured, and other data types.

Visualization of the existing data spotlights a temporal plot of concentrations at the user's site over time. In this way, users can see if trends are already occurring at their site. Options for adding a trend line and downloading regression statistics can be created with the touch of a radio button.

Power analysis consists of two parts. First, the web page produces a power analysis using the current sampling frequency. Users can quickly see what amount of change they can currently detect, and if they are sampling efficiently. The second part allows users to optimize their sampling effort by inputting a different number of samples per year and/or a different number of years. In some cases, optimal frequency may require less effort than currently used. In other cases, more effort may be required than currently used to see optimal trends. Finally, in some cases, no matter how much effort is used, significant trends may not be detected for decades. Regardless, the trends app will help users decide the value of their effort.

The trends app makes assumptions monitoring program designers should be aware of. The largest is that the data variability in the future will mimic the variation observed in the past (and used to drive the power curves). While not an unfair assumption, changes in variability could occur from changes in sources, source strength, storm size, antecedent dry period, measurement method, and other factors.

The calculations used for power analysis utilized the underlying variance from the uploaded and selected data, was de-trended prior to analysis, and assumes  $\alpha = 0.05$  and  $\beta = 0.8$ . For a detailed description of the data analysis used for power analysis, consult Appendix B.

In order to best describe the trends app, and best illustrate how the app can be used, two case studies were used. The first is for wet weather and the second is for dry weather. They are described separately in the next sub-sections.

## Case Study: Los Angeles County Wet Weather Receiving Water Monitoring

The first case study examined four years of wet weather monitoring data from the Los Angeles County MS4 receiving water monitoring program. One aim of this monitoring program was to track trends in the Los Cerritos Channel. In its current configuration, the monitoring program samples up to five storms per year and measures a large variety of pollutants. For this case study example, we utilized total zinc at site 000NONPJ.

Based on the web app's data upload and inventory tab, the monitoring program had sampled 15 events at site 000NONPJ, and total zinc concentrations ranged from 43 to 407  $\mu$ g/L.

Based on the web app's data visualization tab, monitoring indicated that there had been a declining trend in concentrations between the 2016-17 and 2020-21 wet seasons (Figure 3). While there is variability within years, this declining trend in concentration was statistically significant. Zinc concentrations dropped over  $100 \, \mu g/L$  over the four wet seasons or approximately 17% per year on average.

We did not follow up to ask dischargers why this reduction occurred, but from an Advisory Committee point of view, the action associated with this monitoring question is to continue existing implementation activities; these management actions are creating the desired effect. However, further reductions are still necessary because the water quality goals have not been

consistently reached (approximately 95.6  $\mu$ g/L, after adjustment for median hardness and dissolved to total Zn concentrations). Therefore, the action item is continued monitoring to continue documenting declines.

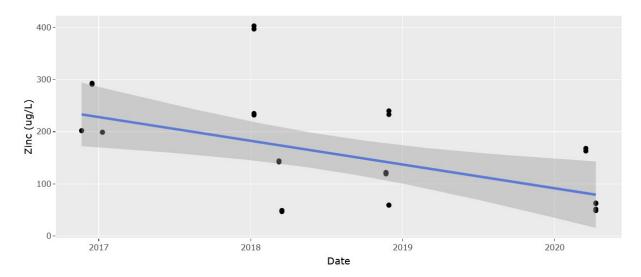


Figure 3. Total zinc concentrations sampled during wet weather at station 000NONPJ in the Los Cerritos Channel. There is a declining trend in concentrations at this receiving water station over the time period monitored.

Based on the web app's power curve tab, the sampling frequency utilized by the current monitoring program (15 samples over 4 years) is effective at detecting trends, but not very efficient (Figure 4, left). The existing effort would have been sufficient to detect average concentration changes of 5% per year (20% over four years). However, the actual reduction was closer to 50% per year (200% over four years) and could have been detected by sampling one-half the current effort. Sampling at 50% of current effort (2 to 3 samples per year for four years) would have been able to detect a >10% change per year (roughly >40% after 4 years).

Of course, assessing the necessary effort to detect actual change after-the-fact is a bit deceiving since managers do not know if/what the trend will be before data is collected. Moreover, it is rare that managers can succinctly agree *a priori* to how much change they would like to detect prior to sampling. Therefore, most monitoring programs try to find the optimal sampling effort to detect trends. This "sweet spot" can be found in the web app (Figure 4, right). By selecting different sample sizes, managers can quickly target optimal sampling frequency by trying different numbers of samples per year over different numbers of years. In some cases, such as this case study, less effort than currently utilized is necessary for optimal effort. In some cases, however, more effort than originally thought may be necessary. In in other cases, the amount of effort to detect optimal changes may be unrealistic or untenable, in which case managers need to re-align their expectations about responses to monitoring questions or identify alternate monitoring questions for AWM.

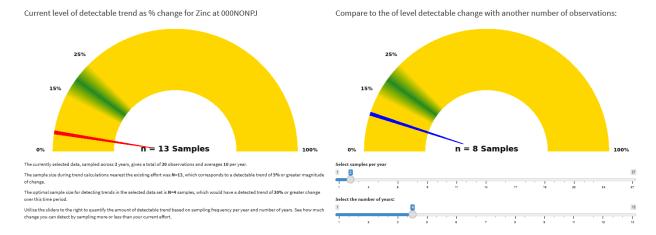


Figure 4. Power analysis gauges defining amount of detectable change per year over four years based on different levels of sampling effort. A priori optimal sampling effort occurs at the inflection point of the power curve, approximated by the green band on the dial. Left) Current effort from uploaded and selected data (4 storms per year), which can detect about a 5% change per year; Right) approximately half the effort (2 storms per year), which can detect 10% change per year. Both sampling frequencies would detect the actual decrease in total zinc concentration.

## Case Study: Orange County Dry Weather Outfall Monitoring

This case study examined seven years of dry weather monitoring data from the Orange County MS4 outfall monitoring program. One aim of this monitoring program was to track trends in discharges within the Santa Ana River watershed, which has water body contact recreation at its terminus. In its current configuration, the monitoring program samples up to five times per year and measures a large variety of pollutants. For this case study example, we utilized Enterococcus concentrations at site ANAB0151@WES.

Based on the web app's data upload and inventory tab, the monitoring program had sampled a total of 16 events over four years at site ANAB0151@WES, ranging in concentration from 390 to 56,000 cfu/100 mL.

Based on the web app's data visualization tab, monitoring indicated there had been a decreasing but non-significant trend in Enterococcus concentrations over the four dry seasons (Figure 5). Enterococcus concentrations decreased from a geomean of 17,000 cfu/100 mL in 2016 to 13,720 cfu/100 mL in  $2019 \text{ (r}^2 = 0.007, p=0.77)$ , or approximately 5% per year over the four dry seasons on average. Detecting the change in Enterococcus concentrations was hampered by large sample-to-sample variability.

We did not follow up to ask dischargers why a statistically significant reduction did not occur over this time period, but from an Advisory Committee point of view, the action associated with this monitoring question is to conduct additional source tracking, assess inspection and response activities, and evaluate BMP effectiveness to determine if the BMPs are functioning.

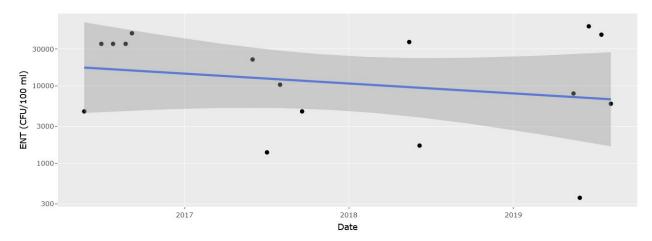


Figure 5. Enterococcus concentrations sampled during dry weather at station ANAB0151@WES in the Santa Ana River Watershed. There is no statistically significant change in concentrations at this outfall discharge station over the time period monitored.

Based on the web app's power curve tab, the sampling frequency utilized by the current monitoring program could only detect trends of 15% change per year over the four-year sampling period (Figure 6, left). Even doubling the effort would not have detected the 5% change per year that occurred in the monitoring data (Figure 6, right); 32 samples in four years would have detected a 10% change per year. This should help set expectations for both regulated and regulatory agencies about how much change they can expect to detect when monitoring for trends, impacting not only what adaptive actions they should be considering, but when they should be considering these decisions.

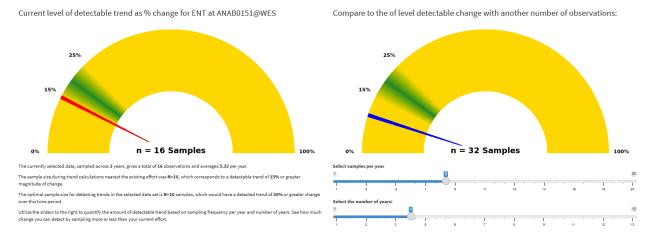


Figure 6. Power analysis gauges defining amount of detectable change based on different levels of sampling effort over the four-year monitoring period. A priori optimal sampling effort occurs at the inflection point of the power curve, approximated by the green band on the dial. Left) Current effort from uploaded and selected data (N=16) which can detect about 15% change per year; Right) Doubling the sample effort over the four-year sampling period (N=32) can detect about 10% change per year.

## **BMP EFFECTIVENESS**

Structural Best Management Practices (BMPs) are one of the key tools used by managers to achieve water quality goals. These BMPs include treatment technologies such as biofilters, bioretention, infiltration dry wells, baffle boxes, and alike. These BMPs are used to reduce pollutant concentrations, reduce runoff volumes, or both. Watershed Management Plans expend tremendous effort to select the proper BMP, select the proper design specifications for the BMP, and select the BMP location within the watershed to ensure optimal performance.

If water quality goals are not being achieved, such as a lack of trend in the previous web app, managers need to start assessing why there has been no change over time before they can select effective options for AWM implementation. There could be a multitude of reasons why water quality goals aren't being achieved, but in this section we develop tools to let managers know if their existing structural BMPs are performing to planning expectations.

## **BMP Effectiveness Monitoring Question**

The four attributes of a BMP effectiveness monitoring question are shown in Table 5. This monitoring design guidance can apply to virtually any BMP with both an inlet and an outlet, and it can apply to virtually any pollutant including chemistry, toxicity, microbiology, or volume. It can apply to concentrations or loads, but we focus on concentrations for this guidance. Benchmarks consisting of water quality thresholds are an important parameter for judging BMP effectiveness.

BMPs can either be performing well, performing, underperforming, or failing (Figure 6). The advisory committee provided these recommendations for actions associated with these three answers:

- Nothing if the BMP is achieving the water quality benchmark.
  - Managers could just continue what they are doing since the BMP is performing as planned.
- Evaluate maintenance if the BMP is underperforming or failing.
  - Maintenance is critical for keeping a BMP operating at its maximum capacity for its design lifetime if the BMP is underperforming or failing.
- Modify design specifications for construction or retrofit.
  - If a BMP is underperforming or failing, it is not operating as designed. First, managers will want to ensure the BMP was constructed to its specifications. Next, managers will want to conduct maintenance to ensure the BMP is operating as designed. If construction meets design specifications and the BMP is operating as designed, then modifying the BMP design specifications makes sense to improve performance for this (and future) BMPs. Additional AWM response options can include changing the type of BMP to be used or adding a BMP treatment train.

Table 5. Defining the four attributes of the BMP effectiveness monitoring question.

Management Question: What is the treatment effectiveness of my BMP?		
Space and time scale:	Space: Any BMP with and inlet and outlet	
	Time: Wet or dry weather, typically event-based condition	
Indicator(s) to be measured:	Various based on local regulatory issues:	
	Bacteria, Metals, Nutrients, Sediment, Pesticides, PCBs	
Benchmark for comparison:	Achieving defined effluent concentration such as a water quality	
	threshold	
	Comparison to other BMPs statewide	
Action to be taken once you have an	Nothing if achieving benchmark	
answer:	Alter maintenance strategy or schedule to enhance performance	
	Modify design specs on construction or retrofit	

## **BMP Effectiveness Monitoring Guidance**

Assessing if structural BMPs are performing to expectations is a challenge because there is no universal standard for making these assessments. Several approaches can be found in the literature including percent reduction, linear regression, and effluent probability, amongst others (Barrett 2008; Afrooz et al. 2019). Each has their own sets of advantages and disadvantages, which influences estimates of performance accuracy, precision, and bias.

In this document, monitoring guidance for stormwater managers incorporates both percent reduction and linear regression approaches, maximizing the advantages and minimizing the disadvantages of each approach alone. Ultimately, the effectiveness of a BMP for pollutant reduction is whether it is achieving the water quality threshold such as a standard, objective, action level, or TMDL target.

Figure 6 identifies the assessment framework for BMP effectiveness. BMP effectiveness is a function of concentration reduction from influent to effluent relative to the water quality threshold trying to be achieved. It is comprised of three components.

The first component of the framework is the x and y axes in Figure 6, which are BMP influent and effluent, respectively. To collect these monitoring data, event mean concentrations (EMCs) are recommended (Leecaster et al. 2002), and both the influent and effluent must be collected during the same storm.

The second component of the assessment framework is the 1:1 line, shown as a diagonal line in Figure 6. This indicates where influent and effluent concentrations are identical. Data points to the right of the 1:1 line represent a reduction of pollutant concentrations from influent to effluent, indicating some amount of BMP pollutant treatment is working. Data points to the left of the 1:1 line represent an increase of pollutant concentrations from influent to effluent, indicating the BMP is exporting pollutants and is not providing treatment.

The third component of the assessment framework is the water quality threshold, shown as vertical and horizontal lines for influent and effluent, respectively. Data points above the threshold require treatment and data points below the threshold are meeting water quality goals. Water quality benchmarks can be water quality objectives, TMDL targets, or action levels.

Combining all three components yields the different narrative categories defined in Figure 6. A BMP is "Performing Well" when influent is above the water quality threshold and effluent is

below the water quality threshold (lower right). A BMP is "Failing" when the pollutant concentration in effluent is greater than influent, and it is also above the water quality threshold (upper left). A BMP is "Underperforming" when the pollutant concentration in effluent is less than influent, but not below water quality thresholds. Similarly, a BMP is "Underperforming" when the pollutant concentration in effluent is greater than influent, but still below water quality thresholds.

Category	Definition
Performing well	effluent concentration less than influent concentration; influent concentration greater than water quality threshold and effluent concentration less than water quality threshold.
Performing	effluent concentration less than influent concentration; influent concentration less than water quality threshold and effluent concentration less than water quality threshold.
Underperforming	effluent concentration greater than influent concentration; influent concentration less than water quality threshold and effluent concentration less than water quality threshold OR effluent concentration less than influent concentration; influent concentration greater than water quality threshold and effluent concentration greater than water quality threshold
Failing	effluent concentration greater than influent concentration; effluent concentration greater than water quality threshold

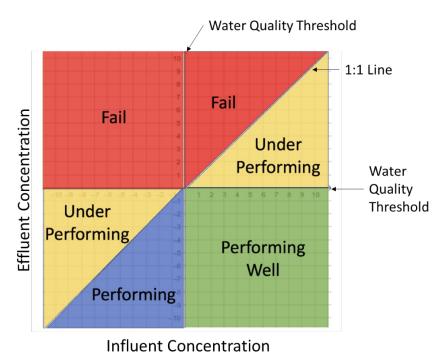


Figure 6. Assessment framework for determining BMP effectiveness.

There is still no currently accepted BMP effectiveness monitoring protocol for the State of California (Afrooz et. al. 2019; Schiff and Brown 2015). This is an area of active development and a standardized monitoring protocol should be available in the near future. Thus, we will refrain from adding detail in this guidance other than managers should follow the development of this valuable resource when it becomes available.

Given that no standardized BMP effectiveness monitoring exists, this guidance document provides some mandatory monitoring design elements for the AWM decision making. These requirements are critical for using the BMP effectiveness web application.

1. Both the influent and the effluent should be monitored in the same storm event.

Since the assessment framework compares influent to effluent to determine pollutant removal, both influent and effluent are required.

2. Multiple storm events are preferred.

Storms and stormwater quality are notoriously variable, as is BMP performance over time. Thus, multiple events are preferred over single events.

3. Event mean concentrations are the preferred sampling design.

Because of the variability in stormwater runoff concentrations, composite samples are preferred over instantaneous grab samples. Single grab samples are likely to provide erroneous and misleading results. Flow-based composite samples are optimal.

4. Particular attention should be directed at flow monitoring.

Flow monitoring for BMPs is challenging. Oftentimes, inlet or outlet structures are difficult to hydrologically rate. To compound this challenge, flows are typically small and even minor inaccuracies in flow monitoring can lead to large bias in results. The best monitoring programs have BMPs with design specifications for installation of flow monitoring equipment.

## The BMP Web App

A web application was created to assist regulated and regulatory managers when designing their BMP effectiveness monitoring program for supporting AWM decisions (BMP App). The BMP app can support monitoring designs for most any indicator (chemical, bacteria, trash, volume) as long as it is continuous data. The BMP app can be used for a variety of BMPs (bioretention, biofilter, dry wells, detention basins, baffle boxes, and infiltration galleries), but is designed to be used one location at a time. Finally, the BMP app can be used for dry or wet weather, but should be used separately in each of the weather conditions.

The BMP app has three elements: 1) upload your data and water quality thresholds, 2) evaluate your BMP effectiveness relative to water quality thresholds, and 3) evaluate your BMP effectiveness relative to other BMPs of the same type.

The data upload can be done manually (i.e., one storm and one parameter at a time) or as a batch file upload using a downloadable data template. The data template will allow users a nearly limitless number of pollutant-site event combinations. The data upload also requires manual entry of the water quality threshold, enabling users to select the thresholds most applicable to their site.

Evaluating your BMP effectiveness relative to water quality thresholds is the prime feature of the BMP app. This information gives the manager instantaneous information on the effectiveness of

their BMP ranging from failing to performing well (Figure 6). If multiple storm events are used, each storm is assessed individually and a summary of all storms combined is provided. This enables managers to evaluate their BMP effectiveness and its variability.

Evaluating your BMP effectiveness relative to other BMPs of the same type is the last piece of information managers need to determine what AWM actions they should take to improve BMP performance. If a manager determines their BMP is underperforming or failing, they want to know if it is just their BMP or if all BMPs of this type underperform or fail. This will guide managers on whether more effort should be spent on maintenance or if a retrofit is in order. For this element, the web app utilizes monitoring data from the California and International BMP databases as BMP comparators.

## Case Study: Biofilter

There are a large variety of biofilter or bioretention cell BMPs (i.e., planter boxes, bioswales, rain gardens, etc.). While each has unique design characteristics specific to their location, they all have a similar basic design concept: inflow (influent), ponding zone and engineered media (for capture and treatment), and outflow (effluent) from emergency overflow with or without underdrain (Figure 7).

For this case study, we utilized biofilter performance monitoring for total copper taken from the California BMP database (Afrooz et al. 2019). Samples were collected during four wet weather events in 2016 from the Los Angeles region. Influent EMC concentrations ranged from 6 to 37  $\mu$ g/L and effluent ranged from 6 to 21  $\mu$ g/L. The water quality threshold selected was 17  $\mu$ g/L, a default TMDL target for wet weather runoff in the Los Angeles River.

The results of the BMP effectiveness assessment relative to water quality thresholds indicated this BMP was "performing" for total copper (Figure 8). BMP performance ranged from "performing well" to underperforming" on a storm-by-storm basis. In every storm event, effluent had the same or lower concentration of total copper compared to influent, indicating that some treatment was occurring in the BMP. Two of the storm events (50%) had influent that exceeded the water quality threshold. One of these storms had effluent below the water quality threshold, which is the ideal condition managers want and consistent with water quality goals. The second storm event had effluent concentrations lower than influent, but the total copper concentration was still above the water quality threshold. While treatment was occurring, this is not meeting the water quality goal. In this instance, managers might trigger an inspection and/or maintenance visit to ensure optimal BMP performance in future storm events.

Although the local BMP performance was deemed "performing", when comparing the performance to similar BMP types in the International BMP Database, the local BMP could be doing better (Figure 9). In this case, the local BMP was performing near the 95<sup>th</sup> percentile of biofilters from the International BMP Database at this same influent concentration. Watershed managers would see this as 95% of other BMPs across the country perform better than their local BMP. There may be some very good design reasons for this lack of comparability. Then again, there might be design aspects that could be improved in this BMP – either through retrofit or maintenance – or design specifications can be upgraded in future BMPs scheduled in the Watershed Management Plan.

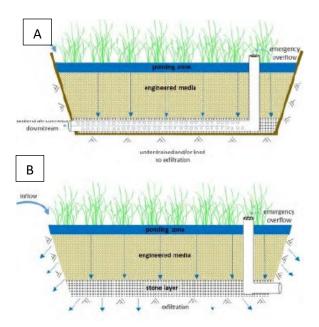
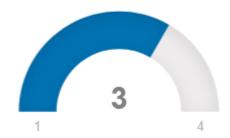


Figure 7. Example of a biofilter with typical design characteristics including inflow (influent), ponding zone and engineered media (for capture and treatment), and outflow (effluent) from (a) emergency overflow with underdrain, or (b) emergency overflow without underdrain. (adapted from Liu and Fassman-Beck 2017).

# A Performing



There are four performance outcomes, your BMP's performance can be **Performing Well**, **Performing**, **Under Performing**, or **Failing**.

	Performance of Individual Storm Data Points	Percentage of Data 🖣
1	1 Performing	
2	Performing Well	25
3 Under Performing		25
CSV	Excel	

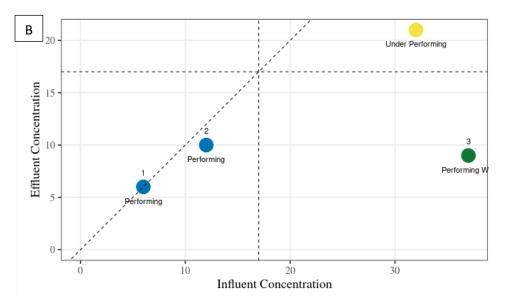


Figure 8. BMP Performance Assessment for the biofilter case study. (A) The overall performance assessment was "Performing" and (B) BMP performance ranged from "Performing Well" to "Under Performing" across the four storm events monitored.

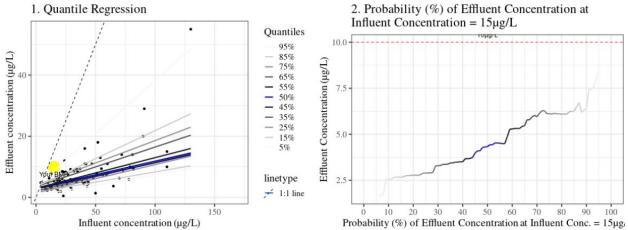


Figure 9. Assessing local BMP performance relative to other BMPs for the biofilter case study. Although the local BMP was "performing" (see Figure 8), other BMPs of a similar type were having better performance indicating managers might want to take adaptive watershed management actions into the future to improve performance.

## **SPATIAL HOT SPOT**

When approaching AWM decision-making, most managers want to know where their best and worst sites are located. The best sites are worthy of protection to ensure they don't degrade. The worst sites are prime candidates for where the management actions should be focused in the upcoming adaptive cycle to improve water quality and achieve water quality goals. For this chapter, we focus on monitoring guidance for "hot spots" to define a watershed manager's worst sites, but the approach simultaneously derives "cold spots" to define the watershed manager's best sites.

## **Spatial Hot Spot Monitoring Question**

The four attributes of a spatial hot spot monitoring question are shown in Table 6. This monitoring design guidance can apply to virtually any site, such as outfalls or receiving waters, and can cover large spatial expanses from city jurisdictions to entire counties or from catchments to entire watersheds. Regional Water Quality Control Boards may want their entire region. Hot spot analysis can be a one-time snapshot, but multiple samples per site are recommended to confirm a hot spot.

The design guidance can apply to most any pollutant including chemistry, toxicity, microbiology, instream biology or habitat, trash, or hydrology. However, the indicator(s) must have an acceptable benchmark or threshold to judge exceedances of water quality goals. Example benchmarks include water quality objectives or standards, TMDL targets, action levels, and alike.

The advisory committee provided these recommendations for actions associated at sites with the greatest frequency of water quality threshold exceedances:

- Additional follow up sampling to confirm the timing and magnitude of water quality threshold exceedances
- Plan and implement source tracking and identification monitoring to pinpoint remediation
- Targeted inspection and enforcement activities
- Targeted non-structural BMP implementation
- Prioritized catchment for future structural BMP implementation

Table 6. Defining the four attributes of the spatial hot spot monitoring question.

Management Question: Which site is my worst site?		
Space and time	Space: User defined site selection; outfalls or receiving waters at watershed to county to	
scale:	regional scales	
	Time: Multiple site visits required, user defined time span. Can be either wet or dry	
	weather, but both combined not recommended.	
Indicator(s) to be	Chemistry, toxicity, biology, microbiology, trash; can be continuous or categorical data as	
measured:	long as it has a threshold	
Benchmark for	Watershed management goal, TMDL target, water quality standard or objective, action	
comparison:	level	
Action to be taken	Identify sites with greatest frequency of threshold exceedance for follow up site-specific	
once you have an	actions: additional sampling for refining timing and magnitude of exceedances, source	
answer:	tracking for remediation, inspections and enforcement, structural and/or non-structural	
	BMP targeting	

## **Hot Spot Monitoring Guidance**

Assessing hot spots is a relatively commonplace monitoring design. Some municipalities monitor (or are required to monitor) many dozens of sampling sites across their jurisdiction, often more than once per year. Thus, threshold exceedance frequency becomes the primary tool for judging hot spots. The higher the exceedance frequency, the "hotter" the site.

The primary challenge to Advisory Committee members was visualization of this monitoring data for decision making. Therefore, this document recommends map-based visualizations. Creating these map-based visualization are straight-forward for some larger municipalities with dedicated GIS staff, but many municipalities do not have this capability. To help with statewide guidance, the web app for this question revolves around map-based features.

The monitoring guidance in this chapter remains silent on exactly where sites should be located. This is a local issue and local managers know their watershed best. There is no minimum or maximum number of sites for this question. Instead, the number of sites is dependent on a manager's need for spatial resolution. However, phased monitoring designs could be utilized whereby less dense resolution is used initially, then more sites are added in hotspot areas as more information is gleaned.

## The Hot Spot Web App

A web application was created to assist regulated and regulatory managers when analyzing their hot spot monitoring program for supporting AWM decisions (Hot spot app). The hot spot app can support monitoring designs for most any indicator (chemical, bacteria, trash, biology, volume) and can accommodate either continuous or discontinuous data. However, multiple sites need to be sampled multiple times, and a water quality threshold must exist. Finally, the BMP app can be used for dry or wet weather, but should probably be used separately in each of the weather conditions.

The web app for mapping hot and cold spots (Hot Spot app) has three elements: 1) upload your data and water quality thresholds, 2) map of threshold exceedances, 3) box plots of threshold exceedances, and 4) bar plot of sample sizes.

The data upload must be accomplished using the downloadable data template. An alternative is to download data directly from CEDEN. Either approach allows users to input as much as 30,000 rows of data. The data upload tab is also the location where the water quality threshold input is required.

The map of threshold exceedances is a thematic map. The larger and warmer the colors are in the map symbol, the greater the frequency of threshold exceedances occurring at this site. In contrast, the smaller and cooler the colors are in the map symbols, the lesser the frequency of threshold exceedances occurring at this site.

The box plots provide managers follow up information on magnitude of threshold exceedances on a site-by-site basis. Managers may wish to re-prioritize two sites with similar exceedance frequencies, but one site exceeds the water quality threshold by a much greater concentration than the other.

The bar chart of sample size provides managers information about sample size for each site. Exceedance frequency is a function of sample size and this visualization allows managers to evaluate confidence in exceedance frequency. Managers may re-prioritize two sites; one with a higher frequency but few samples versus a site with marginally lower exceedance frequencies but many samples.

## **Case Study: Orange County Dry Weather Outfall Monitoring**

Orange County Public Works monitors over 100 sites during dry weather for illicit discharge detection and elimination (IDDE). These sites are monitored monthly during the dry season for hundreds of pollutants and flow. We utilized this data set for the web app constraining it to the years 2012-2018 for one commonly detected pollutant – total copper. Orange County Public Works applies a locally-derived water quality target called the "tolerance interval" (Bernstein et al. 2009). The tolerance interval for copper is  $14 \mu g/L$ .

The map-based visualization illustrates the frequency of copper exceedances of the tolerance interval (Figure 11). Most sites have few water quality threshold exceedances as shown in the numerous small blue symbols. However, there are four sites with larger yellow to red symbols showing a 50% or greater exceedance frequency.

Sample size plays a role in determining what a manager might do with the thematic map information. For example, the site with the greatest exceedance frequency – site BPARA01 in the San Gabriel watershed – had a 100% exceedance frequency but only a single sample collected. However, a site with a lower exceedance frequency – site TTF12@VANLN – had a 72% exceedance frequency of the copper water quality threshold but a sample size of 39 observations. In this case, managers may choose to prioritize the site with a lower frequency but a much greater confidence that it will exceed water quality thresholds into the future. Management actions at this site could include source identification to determine where the copper is coming from so it can be remediated. Site BPARA01 is still worthy of management focus, but the management action might be more measured, including additional monitoring to see if the exceedance frequency continues.

The magnitude of exceedance frequency also plays a role in AWM decision making. In this case study, site TTF12@VANLN had the three greatest copper values in the entire data set, exceeding 150  $\mu$ g/L total copper. In addition, site TTF12@VANLN was the site with the greatest median concentration of copper. The median concentration was 22  $\mu$ g/L, 57% greater than the water quality threshold of 14  $\mu$ g/L. Clearly, this puts site TTF12@VANLN as the hottest of hot spots. The County of Orange worked with their co-permittee municipality to conduct a special study at TTF12@VANLN focused on source tracking.

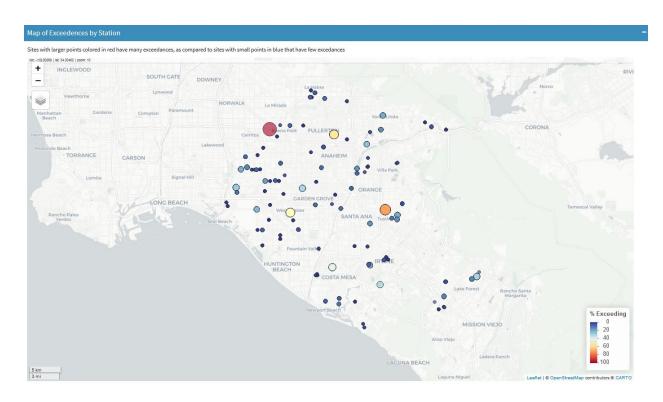


Figure 11. Thematic map of total copper water quality threshold exceedances in Orange County during dry weather 2012-18. This map quickly illustrates which sites have the greatest exceedance frequencies and could be a focal point for future AWM decision-making.

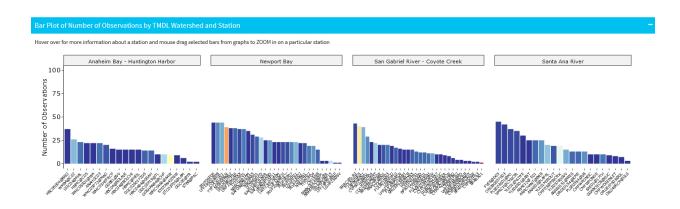


Figure 12. Bar chart illustrating sample size for each site. Fill colors correspond to exceedance frequency with warmer colors have greater exceedance frequencies and cooler colors having lesser exceedance frequencies.

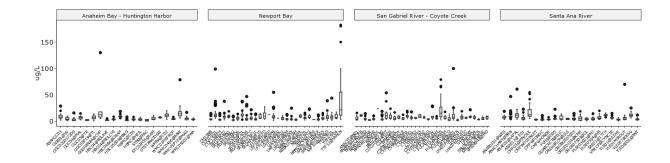


Figure 13. Box plot of total copper concentrations by site within each of the four Orange County watersheds. Boxes represent the 25<sup>th</sup>, 50<sup>th</sup>, and 75<sup>th</sup> percentiles of the data. Whiskers represent 3 standard deviations from the mean. Individual symbols represent outlier data more than three standard deviations from the mean.

### SUMMARY

Adaptive watershed management (AWM) has tremendous value linking watershed planning and implementation with iterative evaluations of progress to achieve desired water quality goals over decadal time periods. But to achieve the desired water quality goals, effective monitoring programs are necessary to provide managers the key pieces of information they need to know on how well the implementation is proceeding. Not many examples exist for this monitoring-informed iterative part of the AWM process. This document advances the guidance in AWM monitoring, and was designed to accommodate the diverse types of watershed landscapes and climates that exist throughout California.

While monitoring guidance was the focus of this document, perhaps the most valuable aspect of this work was developing prioritized monitoring questions, and actions managers might take based on answers to these questions. The questions were developed and prioritized using the Advisory Committee, a committee with some of the most experienced stormwater managers in California. While the committee members had more questions for AWM than could have been addressed in this document, the monitoring questions in this document represent a shared foundation of what most every watershed manager statewide should want to know for AWM regardless of the pollutant, climate, or regulatory status.

The questions ultimately fell into a logical process: 1) Where are my hot spots? 2) Is my hot spot getting better or worse? and 3) Are the BMPs at my hot spot working?

The monitoring designs to answer these questions are challenging, but not incomprehensible or unachievable. The monitoring design tools, in combination with local knowledge, should help any watershed manager create an effective and efficient monitoring program. There is no excuse for not having a monitoring program that can help managers – including both regulated and regulators - make smart and mutually agreeable AWM decisions.

The web applications designed for this project are free and publicly available (<u>Trends App</u>, <u>hot spot app</u>, <u>BMP app</u>). What makes them unique is that users – the managers trying to design their AWM monitoring programs – can utilize their own data to drive them. While case studies are useful illustrations, each watershed is unique and decisions should be made based on local data.

While this document started out as guidance for monitoring design, the Advisory Committee made it clear that data visualization was equally important. What use is a well-designed monitoring program if managers cannot understand and agree upon the output? So, the web applications were created to also produce easily understandable data visualizations. Of course, users can create similar visualizations independent of the web apps.

The monitoring guidance in this document should be updated on a periodic basis as managers iterate through their AWM processes. Important questions now may not be the same questions years from now. The updated guidance should once again re-prioritize critical monitoring questions, upgrade the free and accessible data tools, and ensure easily understandable data visualizations. And, finally, what AWM actions managers will take once they have the answers to their monitoring questions.

#### REFERENCES

Afrooz, A.R.M.N., K.C. Schiff. 2018. <u>Southern California Stormwater Monitoring Coalition Unified Approach to Stormwater Monitoring</u>. Technical Report 1059. Southern California Coastal Water Research Project. Costa Mesa, CA.

Afrooz, N., M. Beck, T. Hale, L. McKee, <u>K.C. Schiff</u>. 2019. <u>BMP Performance Monitoring Data Compilation to Support Reasonable Assurance Analysis</u>. Technical Report 1081. Southern California Coastal Water Research Project. Costa Mesa, CA.

Barrett, M.E. 2008. Comparison of BMP Performance Using the International BMP Database. Journal of Irrigation and Drainage Engineering. 134(5):556-xxx. <a href="https://doi.org/10.1061/(ASCE)0733-9437(2008)134:5(556">https://doi.org/10.1061/(ASCE)0733-9437(2008)134:5(556)</a>

Bernstein, B., B. Moore, G. Sharp, R. Smith. 2009. Assessing urban runoff program progress through a dry weather hybrid reconnaissance monitoring design. Environmental Monitoring and Assessment volume 157: 287–304

Leecaster, M.K., <u>K.C. Schiff</u>, <u>L.L. Tiefenthaler</u>. 2002. <u>Assessment of efficient sampling designs for urban stormwater monitoring</u>. *Water Research* 36:1556-1564.

OWP. 2018. Quantitative Methods that Support Reasonable Assurance Analysis for California's Alternative Compliance Framework. Report to: Strategy to Optimize Resource Management of Storm Water (STORMS), Division Of Water Quality, State Water Resources Control Board. Prepared by: Office of Water Programs, California State University, Sacramento. 40 pp.

Paradigm. 2017. Developing Reasonable Assurance: A Guide to Performing Model Based Analysis to Support Municipal Stormwater Program Planning. Report to: US Environmental Protection Agency, San Francisco, CA. Report from: Paradigm Environmental, San Diego, CA. 52 pp. <a href="https://www.epa.gov/system/files/documents/2021-12/dev-reasonable-assur-guide-model-base-analys-munic-stormw-prog-plan-2017-02.pdf">https://www.epa.gov/system/files/documents/2021-12/dev-reasonable-assur-guide-model-base-analys-munic-stormw-prog-plan-2017-02.pdf</a>

PG Environmental. 2018. Improving Stormwater Program Monitoring, Evaluation, Tracking, and Reporting. Workshop Report and Recommendations. Prepared for: U.S. Environmental Protection Agency Region 9, San Francisco, CA. October 12, 2018. PG Environmental, Chantilly, VA EPA Contract No. EP-C-16-003. 57 pp.

Schiff, K.C., J.S. Brown. 2015. North Coast Areas of Special Biological Significance Regional Monitoring Program: First Year Results. Technical Report 856. Southern California Coastal Water Research Project Authority. Costa Mesa, CA.

# APPENDIX A: TRANSLATING PRIORITY ISSUES INTO MONITORING QUESTIONS

Big Issue	Focal Management Questions	Monitoring Question Translation
Receiving Water Health	Where are my healthy streams located that need to	What is the spatial extent and magnitude of
Healthy stream ecosystems	be protected?	[healthy stream indicators] in [my jurisdiction]
Public health protected		during [pick time span]?
Trash minimized		
Are receiving waters healthy/supporting BU	Is stream health getting better or worse based on	What is the trend in [healthy stream indicators] at
Improvements in water quality	my previous action/inaction?	[pick location] between [pick time span]?
		Are [pick indicators] below [pick threshold] during
	Are body contact areas safe to swim?	[define swimming time period}?
	Are my streams meeting water quality objectives?	
Outfall Monitoring	Which outfall is my worst outfall?	What is the relative frequency of [pick indicator]
Spatial explicitness – loads, concentrations		exceeding [water quality threshold] of storm drain
Outfall concentrations compared to numeric		outfalls in [pick jurisdiction] during [pick time
thresholds		period]?
Effluent relative to receiving water concentrations		
Effluent numerics	Is my outfall meeting watershed plan goals?	Are [pick indicators] below [pick threshold] during
Load reduction		[define time period}?
Outfall impacts on receiving waters	1	NA/
	Is my outfall getting better or worse?	What is the trend in [pick indicators] at [pick outfall]
		between [pick time span]?
		NA/bet is the veletive fleeding or consentration! for
	Is my outfall responsible for receiving water	What is the relative [loading or concentration] for [pick indicator] of outfall XX compared to sources
	impacts?	XYZ to [pick waterbody] during [pick time period]?
BMPs and Load Reduction	What is the treatment effectiveness of my BMP(s)?	What is the [pick performance metric] in [pick
BMP effectiveness	Which is the best performing BMP?	indicator] during [pick time period] at [locations]?
Multi-benefit improvement projects (i.e., WQ and	William is the best performing biving	indicator during tpick time period at flocations!
flow modification)	Is the BMP improving my outfall?	See outfall trends question
Project implementation with multiple goals and	is the bivin improving my outlain:	Oee outrail trends question
partnership with multiple agencies		
Source control		
Post construction O&M effectiveness		
Volume capture		
Pounds of trash removal – increasing over time		
Program Administration	Am I completing the required number of	Is [pick inspection type] frequency above [pick
Maximum return on investment (activities for	inspections?	target] in [pick jurisdiction] during [pick time
achieving receiving water objectives)		period]?
Pollution prevention – IC/ID, housekeeping		
Inspections response	Are inspections resulting in violations that require	What is the relative frequency of [pick violation
Action implementation completion and response –	enforcement?	type] in [pick jurisdiction] during [pick time period]?
trust		,, , , , , , , , , , , , , , , , , , ,

Big Issue	Focal Management Questions	Monitoring Question Translation
Monitoring	Are enforcement actions leading to corrective	What is the frequency of compliance with [pick
Enforcement actions	actions?	violation type] in [pick jurisdiction] during [pick time
Reporting		period]?
Leadership – position of authority/chain of		
command/experience & knowledge	Am I maximizing integration across my different	What frequency per [pick time period] do I have
Account for activities	programs?	partnered projects with [choose partner type] on
How effective at integrating program pieces,		[choose project type]?
synergy		\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
Accomplishments relative to planned activities		What frequency per [pick time period] do I have
Management Reactions to persistent problems Relative effectiveness of different		leverage-funded projects with [choose partner type] on [choose project type]?
approaches/strategies		on [choose project type]?
Training and qualifications of staff		
Project implementation with multiple goals and		
partnership with multiple agencies		
Watershed area treated – asset management		
Implementation of monitoring activities		
Amount of green infrastructure implemented or		
planned		
Cost effectiveness of monitoring relative to		
information		
Cost of implementation activities relative to		
effectiveness		

# APPENDIX B: TREND ANALYSIS CALCULATIONS

The web applications focused on an assessment of trends over time to understand long-term changes at specific locations for parameters of interest, and if management actions to reduce pollutant concentrations or loads have been successful. Trends were evaluated visually using plots of concentration or loading of observed data over time, including an evaluation of monthly and annual trends (McGill et al. 1978). Formal hypothesis tests for trends also included linear regression analyses. Linear regression trends were evaluated based on deviations of the annual average of a parameter from the grand mean at an individual monitoring station. The regression analysis reported an estimate of slope (change in concentration per year) and overall significance of the regressions.

# **Power analysis**

A critical question addressed during monitoring program development was how well the current sampling design was able to detect trends of interest. In particular, questions were evaluated regarding the ability to detect a specified magnitude of change (e.g., 30% decrease per year over ten years). Power analyses were conducted for specific parameters and locations where sufficient data were available. In essence, power describes the probability of observing a true event in a population, based on a sample of the population and if the true event actually occurred. This is analogous to observing an actual change in water quality conditions for a given sample design with the knowledge that sampling is discontinuous over time and at varying time intervals depending on location.

The first power analysis estimated the ability to detect a specific trend for a desired sampling frequency. For a chosen parameter and location, the observed time series was first detrended by taking the residuals of a regression of concentration or load vs time. From the residuals, the variance of the dataset around the mean was estimated and used to simulate new time series from which power was evaluated. For example, if a 50% change (increase or decrease) was considered the true change, a simulated time series was created by first estimating the linear change over the length of time that the true time series was observed (e.g., ten years) and then imposing uncertainty in the linear estimate by adding variance from the residuals to the linear trend. The observed level of sample effort was considered 100% of the current effort if the number of observations in the simulated time series was the same as the observed. Evaluating power at different levels of effort required subsampling of each simulated time series for the selected level of effort. For a large number of simulated time series (n = 1000), power was estimated as the percentage of simulations where the change was significant based on linear regression. This was repeated for varying sample effort from 10% to 200% of the current for a given time series.

Power analysis was constructed by comparing two linear regressions and determining if the slopes were statistically different from each other using a t-test for a fixed number of data points in both regressions (reviewed in Andrade and Estévez-Pérez 2014). With this approach, a simple linear least-squares regression was created from the temporally sampled water quality data, where data of sampling was the predictor variable and concentration of the a given analyte was the response variable. From this regression, the number of data points (n), the variance of the data (i.e., the sample variance in the response variable  $S^2$ ) and the slope of the regression (i.e.,  $\beta$ , the magnitude of the temporal trend) were extracted. The slope term was standardized by the mean value of the analyte concentrations in the initial dataset, transforming the slope from a

value representing the change in absolute concentration over time (e.g., mg L<sup>-1</sup> year<sup>-1</sup>) to a value representing the relative magnitude of change in concentration over time (e.g., % change year<sup>-1</sup>). A control regression was created from the initial regression that had a slope of zero but retained the sample variance of the initial regression. A test regression was created from the initial regression with a 5% per year slope but retaining the sample variance of the initial regression. A Student's t-test was then used to determine if the slope of the **test** regression was significantly different ( $\alpha$ =0.05) from the slope of the **control** regression for a sample size of 3 data points. This comparison between **test** and **control** regression slopes was repeated iteratively from a sample size of 3 to a maximum of 10n. From this iteration, the smallest sample size that produced a significant difference between the test and control slopes, with the test slope set to 5% per year was retained as the optimal sample size to detect a 5% per year difference in concentration for the analyte of interest. This iterative application of t-tests with different sample sizes was then repeated for a 10% per year slope in the test regression, a 15% per year slope, and so forth at increments of 5% per year of effect until 95% per year. For each change in slope, a new optimal sample size was retained that contained the minimal number of samples to detect the given difference in the two slopes. These values were subsequently plotted against each other to illustrate the trade-off between minimal number of samples and the magnitude of effect for which a statistical difference could be determined, given the variance structure of the data submitted to the application. From this curve, optimal sampling effort could be determined.

# **Optimal sample effort**

An optimal level of sample effort was derived from the power analyses to describe the balance between over- and under-sampling. From a programmatic perspective, the optimal level of effort minimizes sampling cost by identifying the level of effort where any additional samples do not substantially increase the ability to detect a trend, whereas reductions in sample effort cause a disproportionate increase in the magnitude of the trend to be detected for a desired level of power. Graphically, the optimal level of effort is the inflection point on a power curve where the y-axis shows the magnitude of the trend to detect, and the x-axis is the level of sampling effort. This inflection point was determined quantitatively for each water quality parameter as the point in a monotonic power curve where the slope of y versus x (i.e., trend to detect vs sample effort) exceeded that of x versus y (i.e., sample effort vs. trend to detect). Given that sample effort and variance of each time series differed considerably among the observed time series, optimal effort was identified for power curves only where sufficient data were available. A power estimate of 80% was considered a sufficient target for optimal effort.

#### Citations

Andrade, J. M., and M. G. Estévez-Pérez 2014. Statistical comparison of the slopes of two regression lines: A tutorial. *Analytica chimica acta* 838: 1-12.

# APPENDIX C: WEB APP TUTORIALS

# App 1 Tutorial: Adaptive Management Application for Stormwater Managers

#### **Table of Contents**

1 . 4	
Introc	luction
HILLOU	luction

II Application Layout

III Application Home Page

IV Select My Data

V Data Selection Panel

VI Input Data

A <u>Existing Data</u>

B <u>Upload Data</u>

CEDEN Direct Download CEDEN Modified Download CEDEN Submission Template

General Template

C <u>Filter Data Panel</u>

D Summary Information for Currently Selected Dataset

VII View Data Results

A Trends Plot

B Regression Results

VIII How is My Sampling Frequency

A <u>Effort Gauge Charts</u>

B Effect Size Chart

#### I. Introduction

Trends monitoring – assessing changes in stormwater runoff and water quality over time with the implementation of management actions – is one of the key monitoring designs for Adaptive Watershed Management.

The purpose of this shiny application is to evaluate trends in your own data, and to design a monitoring program to effectively and efficiently quantify trends. This information helps managers determine if changes have occurred with specified levels of confidence, if managers can confidently determine that changes have not occurred and further adaptive action is necessary, or that it is too soon to tell if a change has occurred and how much more monitoring is needed.

This tutorial describes how to use the shiny application for designing a temporal trends monitoring program for Adaptive Watershed Management.

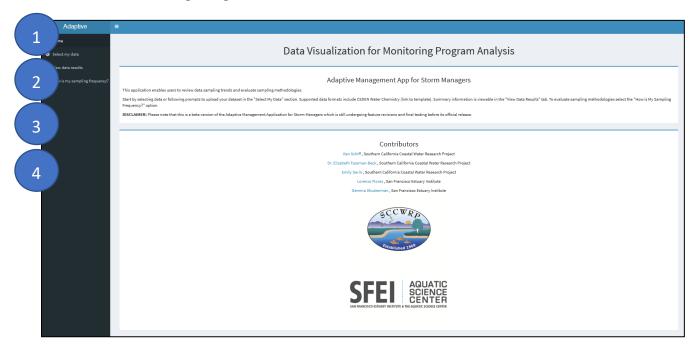
This application utilizes a user's local data entered using a template or downloaded from California Environmental Data Exchange Network (<u>CEDEN</u>). The data can be collected from receiving waters or storm drain outfalls. The data types can include almost any parameter including chemistry, toxicity, bacteria, volume, or trash. The data can be collected either during wet or dry weather, but combining dry and wet weather is not recommended.

For more information about this application, please go to the project Final Report or contact Ken Schiff (KenS@sccwrp.org) and Gemma Shusterman (gemmas@sfei.org). You can visit the website for this application by clicking this link.

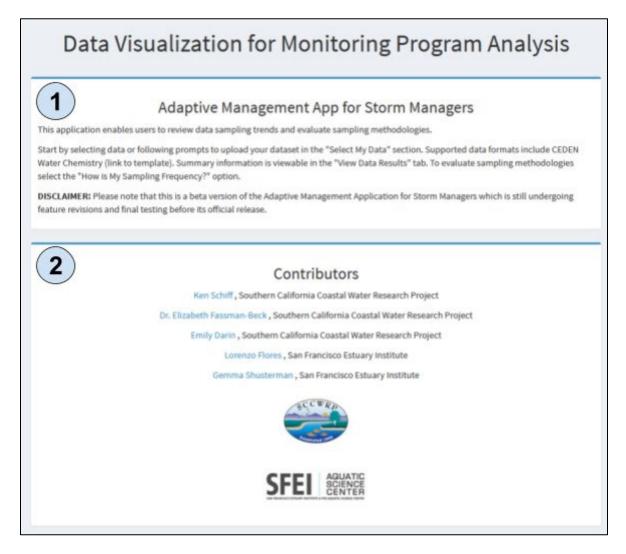
# II. Application Layout

This application has four main tabs accessed by using the left-hand side navigation pane:

- 1. 'Home' Application introduction text and contributors list
- 2. 'Select my data' Use to select or upload data, then filter data for use in results and sampling frequency sections
- 3. 'View data results' Trend plot for user-selected data and linear model regression results
- 4. 'How is my sampling frequency?' Gauge and effect size power curves for an optimally efficient monitoring design



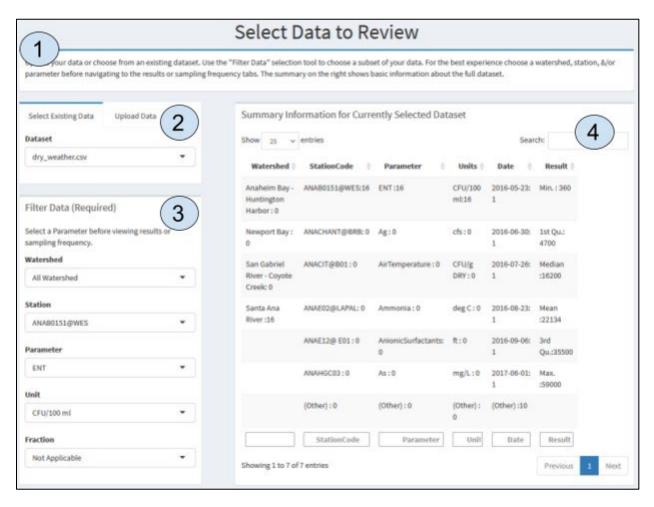
# III. Application Home Page



The Adaptive Management App for Storm Managers Shiny Application landing page contains:

- 1. Introductory text and disclaimer
- 2. Contributor listing containing links to profiles for individual contributors to the project

# IV. Select My Data



The "Select my data" tab contains the following sections:

- 1. Introduction / overview
- 2. Data selection panel (required)
- 3. Data filter panel (required)
- 4. Summary information table for currently selected data

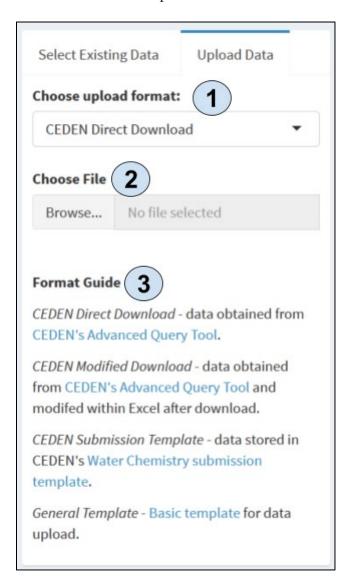
### V. Data Selection Panel

The Data selection panel has two tabs: "Select Existing Data" or "Upload Data."

The "Select Existing Data" tab contains a dropdown list which allows you to select from 1 of 3 pre-loaded datasets. See Existing Data for more detail regarding these datasets.

The "Upload Data" tab allows you to upload data in four different formats, see Uploading Data for more information.

- 1. Use the "Choose upload format" dropdown to select between data upload formats, described above.
- 2. Click "Browse..." to select a file to upload.
- 3. The "Format Guide" offers links to further information or templates which correspond to the described upload formats.



# VI. Input Data

The application requires an initial dataset to drive analyses and visualizations. There are three sample datasets available for demonstration purposes or the user may upload data formatted to specifications detailed below.

# A. Existing Data

The application includes the following sample datasets:

- dry\_weather.csv Data gathered during dry weather conditions (default)
- SMC-SCCWRP\_trash.csv Trash collection dataset
- Wet weather.csv Data gathered during wet weather conditions

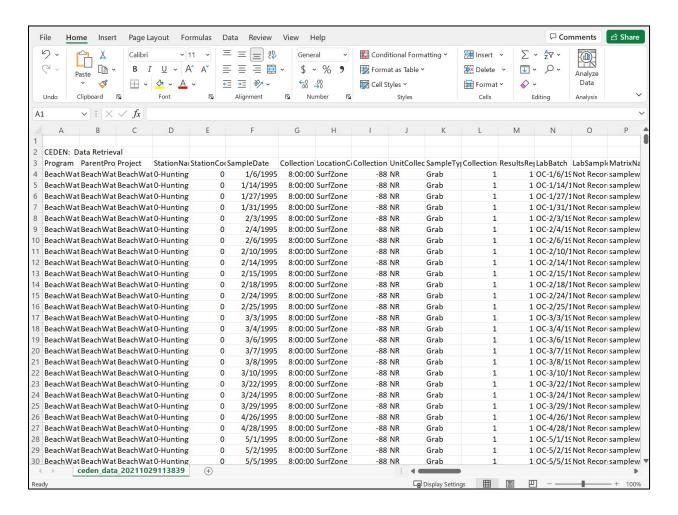
These datasets were chosen to represent a variety of data requiring differing monitoring designs. One data set consists of multiple parameters collected from storm drain outfalls during dry weather, The second data set consists of a different set of parameters collected from receiving waters as event mean concentrations during wet weather. The third data set consists of trash collected from storm drains. The application will display the first dry weather dataset by default. Use these existing data sets for learning how to use the application. For best results, use your own local data using the 'upload data' tab.

# B. Upload Data

There are four options for uploading data. Each data upload option corresponds to a specific template, all of which are available or linked to through the application. The next section provides the user details on how to use each of the four options.

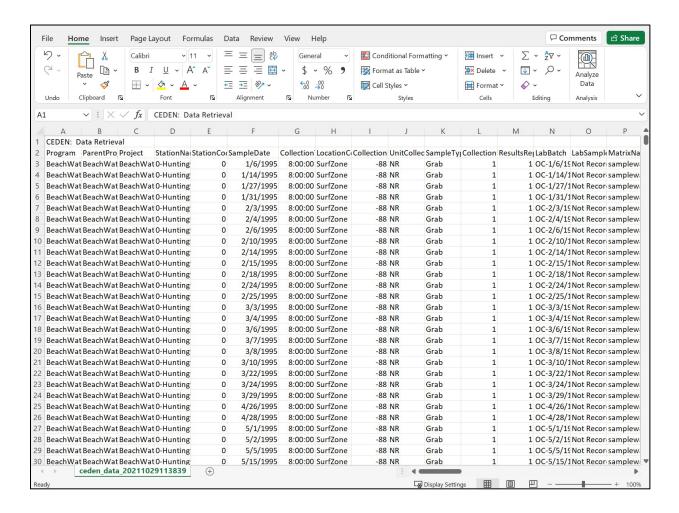
#### CEDEN Direct Download

Input data may be obtained from CEDEN's Advanced Query Tool. This format assumes no modifications have been made to the downloaded file (see image below). Column names should not be modified and any changes may result in errors when uploading to the application.



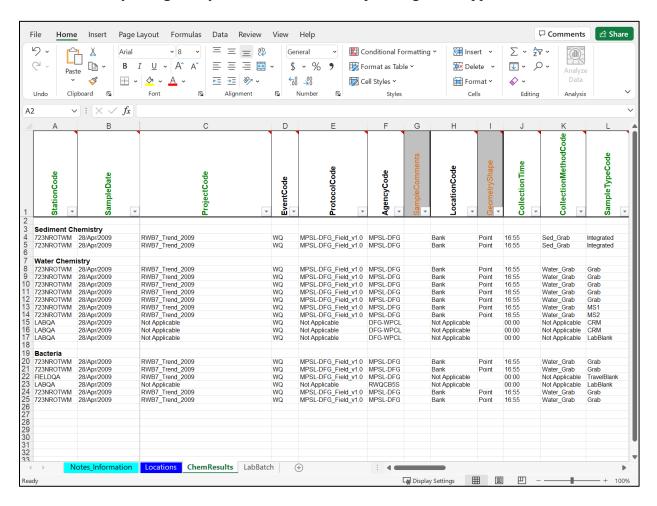
#### CEDEN Modified Download

Data obtained from CEDEN's Advanced Query Tool and modified within Excel after download may also be used. Column names should not be modified and any changes may result in errors when uploading to the application.



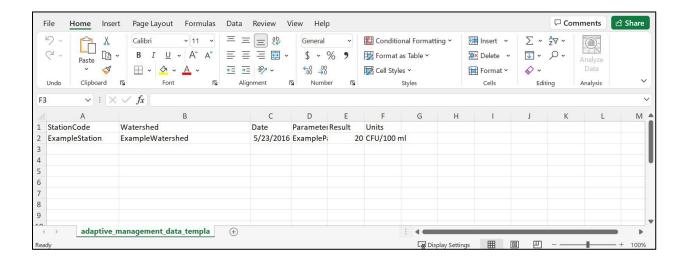
### CEDEN Submission Template

Data stored in CEDEN's Chemistry submission template format (pictured below) is the third option for uploading. This template holds data meant for submission to CEDEN for review/inclusion. Data formatted using the "ChemResults" tab, (selected in the below screenshot) may be used as input to the application. Again, column names should not be modified and any changes may result in errors when uploading to the application.



# General Template

A basic excel template for data upload. This simplified template is for users who are not planning to submit data to CEDEN. Column names should not be modified and any changes may result in errors when uploading to the application.



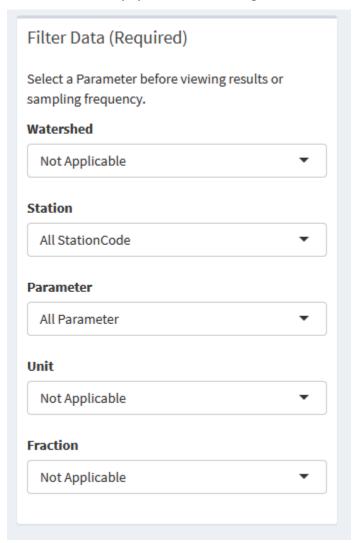
# C. Filter Data Panel

The Filter Data panel allows you to subset the selected or uploaded data. Each dropdown selection box contains unique values automatically generated from the corresponding columns within the selected dataset.

If the value "Not Applicable" appears, that means the data do not have values for that column and it will not be used to filter data.

Users must apply filters from top to bottom. If you select a Watershed, then Station, then Parameter, and then go back and select a different Watershed, the selection for Station and Parameter will revert to the default "All" state.

The data selection must be filtered in order to activate "View data results" or "How is my sampling frequency?" sections. We suggest the data should be subset by watershed or station and, most crucially, you must select a parameter.



# D. <u>Summary Information for Currently Selected Dataset</u>

The "Summary Information for Currently Selected Dataset" table shows unique values and value counts for filterable fields such as Watersheds, Station Codes, Parameters, and Dates. It also shows statistical information for the Results column which are the data used to drive calculations and visualizations in the "View data results" and "How is my sampling frequency?" tabs.

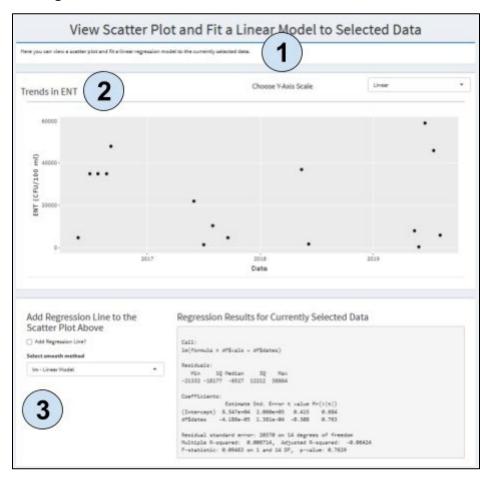
- 1. Use the search bar to filter the table for specific records. You can also filter the table using individual column search boxes located under each column.
- 2. The "Show X entries" dropdown has no effect on the data shown in the summary table.
- 3. The text "Showing Y of X entries" indicates the number of unique records displayed out of the total number or unique record values shown in the table.



### VII. View Data Results

The "View data results" tab contains three panels:

- 1. Summary information
- 2. Trends plot
- 3. Regression results



# A Trends Plot

The trends plot shows numeric results for the currently selected data by date. You can hover the mouse cursor over individual data points to view the specific result value, date, and station.

The "Choose Y-Axis Scale" dropdown allows the user to select between Linear and Logarithmic scales for the Y-axis.

#### B Regression Results

This panel shows linear regression model results for the currently selected data.

Select "Add Regression Line?" to display the computed regression line in the Trends scatter plot. This option will also display confidence bands around the line.

The "Select smooth method" dropdown box allows users to choose between a Linear Model (lm) and Polynomial Fit (loess) model.

# VIII. How is My Sampling Frequency

The "How is my sampling frequency" tab allows you to evaluate the power of current sampling efforts to detect trends and identify an optimal sampling methodology to achieve statistically significant results. This section contains three panels:

- 1. Introduction text
- 2. Gauge charts showing current and potential level of effort
- 3. Power curve relating effect size and sample size



# A. Effort Gauge Charts

The effort gauge charts are a graphical representation of how much change can be detected for a given sampling effort.

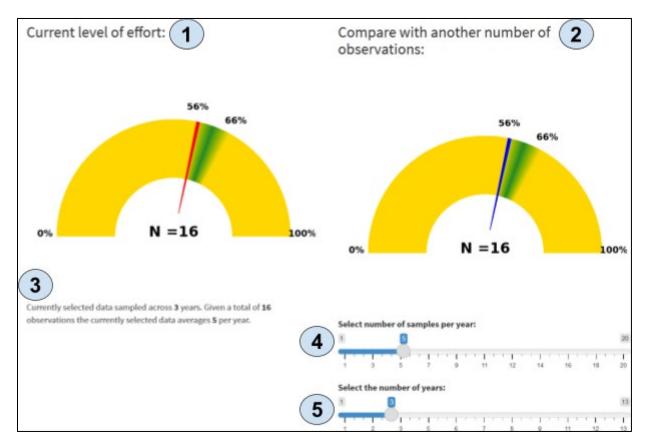
The level of effort gauge chart section is composed of the following parts:

- 1. Gauge chart showing the amount of detectable change based on the current level of effort for selected data. This is static and cannot be modified
- 2. Gauge chart showing the amount of detectable change based on the user defined level of sampling effort using the sliders to select number of samples per year and number of years

for sampling. The amount of detectable change adjusts automatically when sampling effort is modified

- 3. Sampling effort summary information for the currently selected data
- 4. Slider to select number of samples per year in the new monitoring program
- 5. Slider to select the number of years to sample in the new monitoring program

Both charts contain a yellow band representing the amount of detectable change from 0% to 100%. The red needle indicates the amount of detectable change for the current level of effort. The blue needle indicates the amount of detectable change for the user selected level of effort (using the sliders below the gauge chart). The green band indicates the range of optimal sampling effort where reductions in effort have disproportionately large decreases in the amount of detectable change and increases in effort have disproportionately small increases in the amount of detectable change.



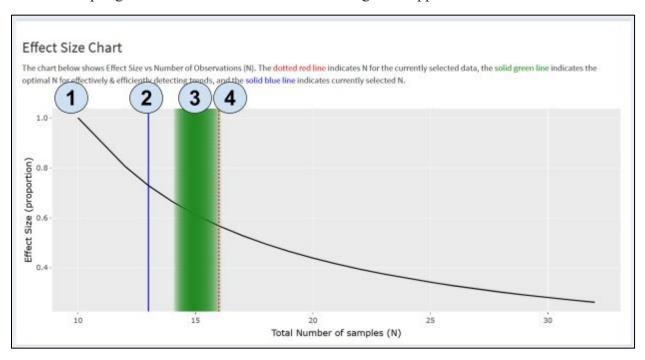
Sampling effort is a function of number of samples per year and number of years. As sampling effort increases, smaller changes can be detected, and the needle will rotate counter-clockwise. As sampling effort decreases, only larger changes can be detected, and the needle will rotate clockwise. For managers designing a monitoring program and the amount of change desired is not *a priori* defined, the green band is the "sweet spot".

### B Effect Size Chart

The effect size chart - a power curve - shows:

- 1. line graph representing the Effect size (proportion) vs Total Number of samples (N)
- 2. blue line indicating the user defined effort via the sliders
- 3. green band representing the optimal sampling frequency for the most efficient monitoring program.
- 4. red line indicating existing effort for the selected data

The blue line will adjust automatically as you adjust the sampling effort in the "Compare with another number of observations" gauge. Hover the mouse curser over the plot elements and the level of sampling effort and amount of detectable change will appear.



# **App Tutorial: Evaluating BMP Performance**

### Table of Contents:

- I. Introduction
- II. Application Layout
- III. How is my BMP performing?
  - A. Manual Upload
  - B. Data Upload
  - C. Examples and Definitions
  - D. Your Data
- IV. How are other BMPs doing?
- V. Glossary

#### I. Introduction

BMPs (Best Management Practices) are designed to reduce runoff volume and or pollutant concentrations in urban stormwater. The purpose of this web application is to quantify how well a BMP is performing either individually compared to water quality goals *or* compared to other BMPs of the same type. This BMP performance information helps watershed managers adaptively manage by determining if their BMP's performance is meeting engineering design goals *or* for selecting the right BMP for meeting watershed management goals.

For more information about this application, please go to the project Final Report or contact Ken Schiff (KenS@sccwrp.org) and Emily Darin (EmilyD@sccwrp.org). You can visit the website for this application by clicking this link.

# II. Application Layout

This application has three main tabs by using the left-hand side navigation pane:

- "Welcome",
- "How is my BMP Performing?"
- "How are other BMP's doing?".

Click on each tab to access its content. The "Welcome" tab has introductory information about the app and its creators. The "How is my BMP performing tab" allows the user to analyze their own BMP monitoring data (using either manual enter or data file upload), and the "How are other BMP's doing" tab allows the user to analyze the performance of their BMP compared to similar BMP types in the international BMP database.

# III. How is my BMP performing?

In this tab, the user can either input their BMP data manually or by uploading a data file.

The user needs five data type: an **Influent** concentration, **Effluent** concentration, a pollutant **Parameter**, concentration **Units**, and a **Water Quality Threshold** for the BMP performance web app to function.

# A. Manual Upload

You can enter the influent and effluent concentrations manually using the "manual upload tab" or you can upload a CSV file using the "data upload" tab. The application works best using Event Mean Concentrations (EMCs) for influent and effluent.

- 1. Select the Manual Data tab
- 2. Enter the Influent EMC, Effluent EMC, and concentration units
- 3. Enter the Water Quality Threshold

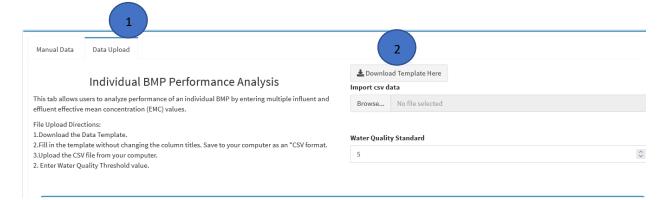


Influent and effluent concentrations must be from the same storm event. You can manually enter up to thirteen storm events. If you have more than thirteen storm events, you should utilize the data file upload procedure in the next section.

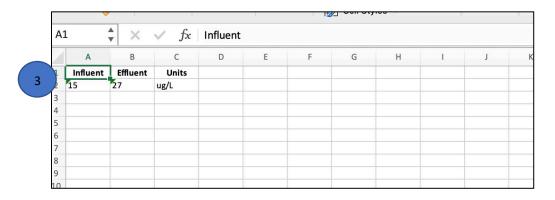
### B. Data Upload

Bulk upload, including multiple parameters from the same storm events can be used to make things easier for larger data sets.

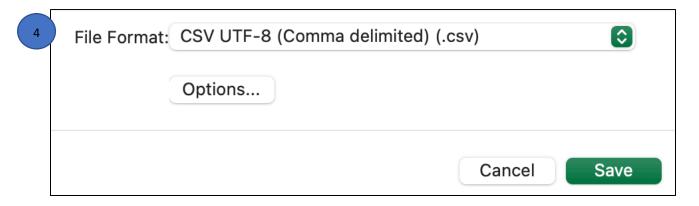
- 1. Select the Data Upload tab
- 2. Download Excel Template



3. Enter the Influent, Effluent and Units into the Excel File. Each row should be a different storm event.



4. Save the data file as a CSV.

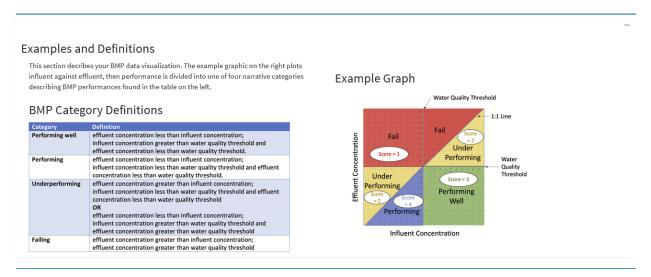


- 5. Upload your data file into the application by clicking "Browse" and selecting your just created CSV file.
- 6. Enter your Water Quality Standard or other meaningful water quality threshold being targeted by your watershed management program. This can only be done manually.



# C. Examples and Definitions

The BMP performance outcome is dependent on both your influent and effluent EMC values as well as your water quality threshold. The graph and table below are found in the app (with a collapsible tab) provide an example and definition of how BMP performance is quantified. Go to the project Final Report to get more detailed interpretation of the performance categories and definitions.

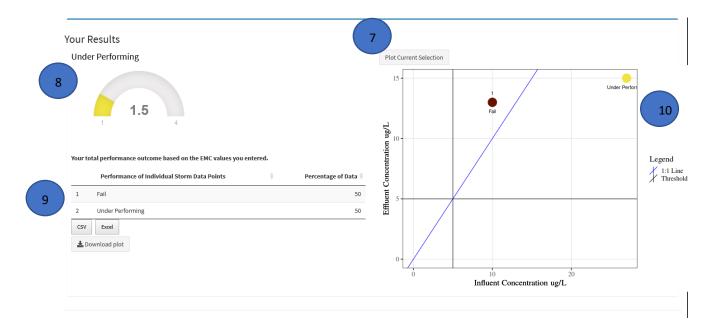


### D. Your Results

Each data point represents an influent-effluent EMC pair and, depending on which quadrant the data point falls in, your BMP will receive a performance score.

Your results are provided three different ways:

- 7. Click on **Plot Current Selection** to get your results.
- 8. Overall summary of your BMP performance using a performance meter
- 9. Distribution **table** of your individual performance data. These data can be downloaded.
- 10. **Graph** of individual storm results. This plot can be downloaded.

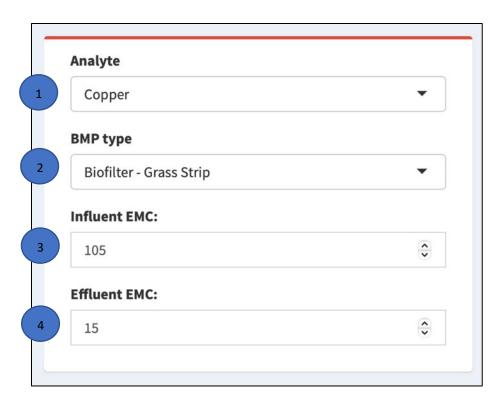


# IV. How are other BMPs doing?

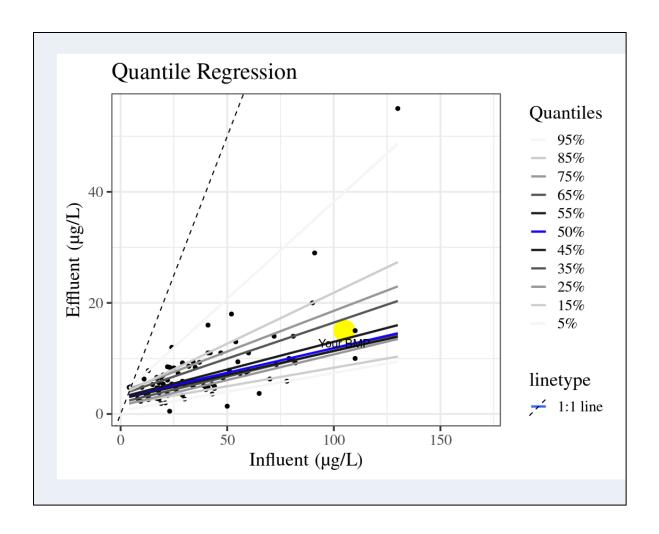
This tab allows the user to quantify how their individual BMP performance compares to similar BMP types using data from the <u>International BMP Database</u>. A full description of the analysis and interpretation is provided in the project Final Report.

To use this tab, you must:

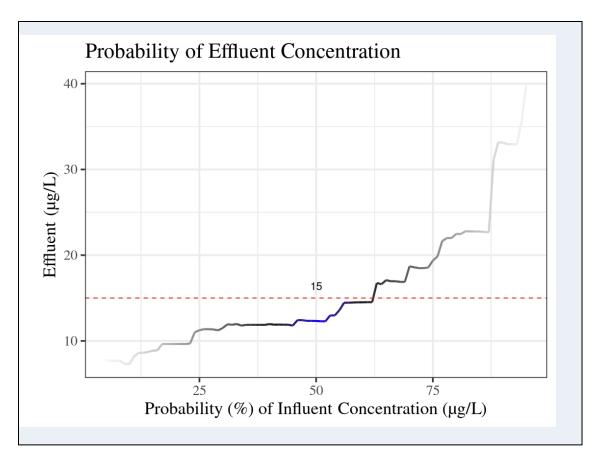
- 1. Select your **Analyte**
- 2. Select your **BMP type**
- 3. Input your average Influent EMC
- 4. Input your average Effluent EMC



Once the necessary data have been entered above, the application will automatically create a quantile regression using data from the <u>International BMP Database</u> on the left side of the pane with BMP influent concentration on the x-axis, BMP effluent concentration on the y-axis, and your BMP data plotted as a yellow symbol.



The probability of effluent concentration is automatically generated by the application and plotted on the right.



This graph is created by taking a vertical cross-section of the quantile regression at the user-defined influent concentration. The plot illustrates the probability of achieving any effluent concentration based on the user's local influent concentration using this BMP type. The red dashed reference line is the user-defined effluent concentration.

# V. Glossary

**BMP: Best Management Practice** 

Effluent: Storm water exiting the BMP after treatment.

Influent: Storm water entering the BMP prior to treatment.

Water Quality Standard or Water Quality Threshold: The "allowable" concentration in stormwater, which is typically the targeted performance goal of the BMP engineering design.

Quantile Regression: quantile regression, an extension of linear regression, estimates the conditional median or other quantiles (as opposed to the conditional *mean* in linear regression).

# **App Tutorial: Hotspot Mapping**

### Table of Contents:

VI. Introduction

VII. Application Layout

VIII. Hotspot Detection

A. Data Upload

B. Map of exceedances by station

C. Box plot of pollutant concentrations

D. Bar plots of sample size

IX. Graphs Glossary

#### VI. Introduction

This application allows managers to detect hot spots of contamination within their region based on frequency of water quality threshold exceedances. This will help watershed managers identify locations where additional adaptive management actions can be prioritized including additional monitoring, source tracking, or enforcement.

For more information about this application, please go to the project Final Report or contact Ken Schiff (KenS@sccwrp.org) and Emily Darin (EmilyD@sccwrp.org). You can visit the website for this application by clicking this <u>link</u>.

# VII. Application Layout

This application has two main tabs by using the left-hand side navigation pane:

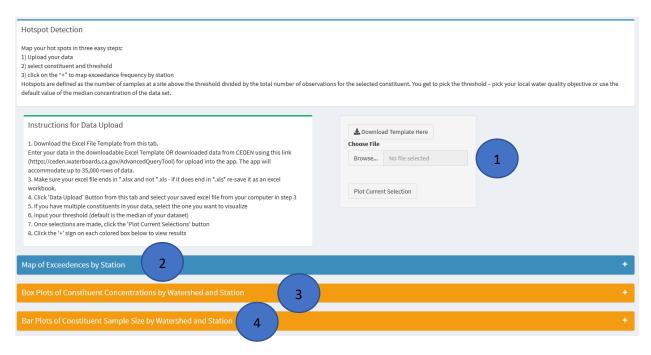
- "Welcome"
- "Hotspot"

You can click on each tab to access its content. The "Welcome" tab has introductory information about the app and its creators. The "Hotspot" tab allows the user to analyze their own monitoring data.

# VIII. Hotspots

This tab has four windows including:

- 1. Data upload
- 2. Map of exceedances by station
- 3. Box plot of pollutant concentrations
- 4. Bar plots of sample size

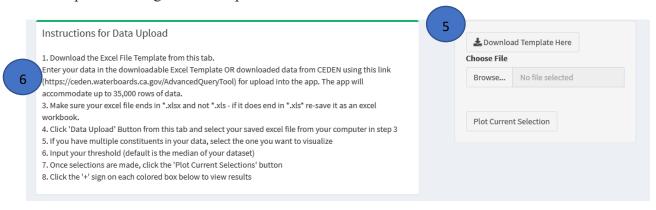


### A. Data Upload

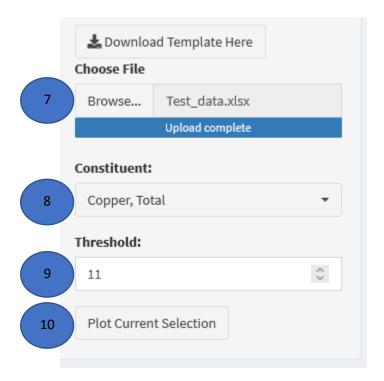
In this tab, the user can input their hotspot monitoring data using the

- 5. Excel template provided, or
- 6. Downloading data from CEDEN.

The minimum data required is **Station Code**, **County**, **Date**, **Analyte Parameter**, **Result**, **Units**, **Longitude**, **Latitude**. You must save your file as a XLSX [Excel Workbook] format to upload it using the "data upload" tab.



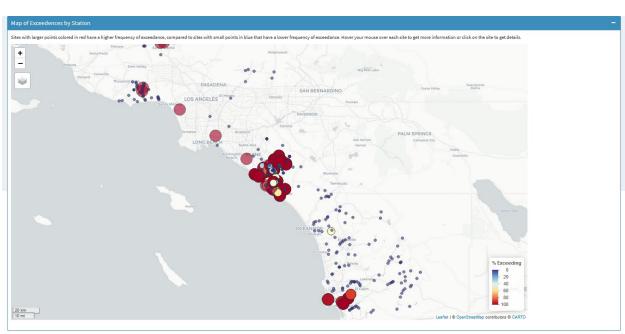
- 7. Upload the saved XLSX on your computer by clicking the "Browse" button.
- 8. Select the **Constituent** (auto-populated from your data set)
- 9. Select your Water Quality Threshold
- 10. Click **Plot Your Selection** to create the map



# B. <u>Map of Exceedances by Station</u>

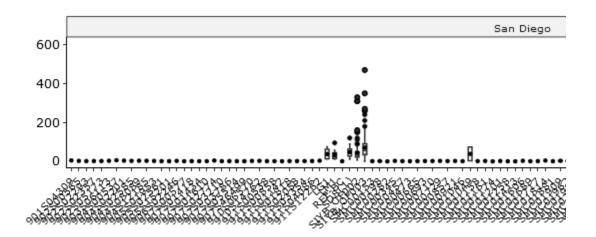
11. Click the blue "+" to view a map of water quality threshold exceedances by station.





- c. Box plot of pollutant concentrations
- d. Bar plots of sample size
- 12. Click the yellow "+" to view a box plot of concentrations or bar plot of sample size by county and station.





Hovering your mouse over the top graph will pop up graph tools including zoom in or out, panning, or resetting axes. Hovering your mouse over the plots will pop up data behind each site's results. Concentration axis can be converted to arithmetic scale to **Log Scale** by clicking the radio button and then clicking **Plot Current Selection** (step #10).

# IX. Glossary

Box Plot: boxes include the 5<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, and 95<sup>th</sup> percentile, plus all individual outliers.

Hot spot: defined as frequency of exceedance of the water quality threshold.

Monitoring site: any site where monitoring occurs including receiving water, storm drain outfall, or source sampling.

CEDEN California Environmental Data Exchange Network