BASMAA

Clean Watersheds for a Clean Bay (CW4CB)

Retrofit Pilot Study Plan

October 22, 2013

Contents

Introduction ……………………………………………………………………………………4
Problem Definition 5
Study Goals ……………………………………………………………………………………6
Primary Research Question... 6
Secondary Research Questions 6
Level of Confidence 7
Catchment Efficiency Calculation 7
Overview of Performance Measurement Options 9
Influent versus Effluent Sampling 9
Sediment Sampling 11
Before and After Construction Sampling.. 12
Paired Watershed Concentration Measurements 12
Volume Measurement.. 12
Study Design Options for Primary Data Objectives 14
Data Objective 1: Mass Removal Estimates... 14
BMP Processes and Potential Water and Sediment Data Collection............................. 15
Bioretention (with underdrain)... 16
Permeable Pavement 16
Flow-through Biotreatment (Swale) 17
Tree Wells 17
Catch Basin Media Insert... 18
Sand Filter 18
Hydrodynamic Separators 19
Variables Affecting Water and Sediment Data... 20
BMP Siting, Land Use, and Land Cover 20
BMP Design and Hydraulic Loading 20
Operation and Maintenance Parameters 21
Pollutant Behavior 21
Storm Intensity.. 22
Soil Permeability... 23
Uncontrolled Variables and Study Assumptions 23
Final Study Plan Design 25
Statistical Testing Considerations... 25
Selected Sample Size.. 25
Selected Constituents for Water Quality Analysis.. 26
Selected Constituents for Sediment Analysis 29
Budget and Schedule 30
Cost and Schedule-Limited Study Plan 30
Adequacy of Study Plan to Answer the Primary Research Question....................... 35
Reducing Analytes and Monitoring Alternative Locations....................................... 35
Data Analysis ……………………………………………………………………………………..37
Catchment Efficiency... 37
Cost-Benefit Ratios... 37

BMP Comparisons..37
Applicability to Future Retrofits..37
Recommendations for Sampling and Analysis Plans..39
Flow Measurement Devices 39
Storm Event Mobilization Criteria 40
Data Quality Objectives..41
Event-Specific Reporting...41
Field Guide Sampling and Analysis 41
Recommendations for Cost Tracking and Analysis 43
References ……………………………………………………………………………………..44
Appendix A ……………………………………………………………………………………..46
Statistically Derived Sample Size 46

Introduction

The Clean Watersheds for a Clean Bay (CW4CB) project is funded through a U.S. EPA grant and focuses on evaluating pilot stormwater best management practices (BMPs) for the control of polychlorinated biphenyls (PCBs) and mercury in stormwater runoff from urban areas in the San Francisco Bay Area (BASMAA 2009). The Bay Area Stormwater Management Agencies Association (BASMAA) is implementing the project. The CW4CB project consists of several tasks. This study plan pertains to the monitoring program required to evaluate pilot stormwater treatment retrofits (CW4CB Task 5).

The report is organized by discussion of the problem, followed by sections on study goals, options for performance monitoring, a review of BMP processes, and other variables affecting performance. This is followed by a description of the selected study design in “Final Study Plan Design”. Additional sections address suggestions for analysis, sampling and analysis plans, and cost tracking.

Problem Definition

Discharges of PCBs and mercury in stormwater cause impairment to the San Francisco Bay, which have resulted in adoption of total maximum daily loads (TMDLs) (SFBRWQCB 2012). Some of the areas with elevated sediment concentrations for these pollutants have been identified in areas with past and current industrial activity (Yee and McKee 2010), but these areas have a fairly low redevelopment rates. Consequently, these areas will not benefit from post-construction standards until they are redeveloped (SFBRWQCB 2011). Instead, these areas can be targeted by stormwater BMP retrofit projects.

There are a number of retrofit options that have the potential to control PCB and mercury discharges in areas with elevated concentrations, but their relative pollutant reduction per catchment area (catchment efficiency) and cost-effectiveness is unknown.

Study Goals

The goal of this pilot study is to develop estimates of load reduction effectiveness at the catchment scale. To reach this goal, the following research questions are prioritized as primary or secondary research questions. This prioritization will be useful in budgeting different elements of the study.

Primary Research Question

A properly conceived study will address the study goal in a manner that supports management decisions. For future retrofit technology selection decisions within the TMDL construct, designers will want to know the PCB and mercury removal in the context of the catchments hydraulic and pollutant load. Site constraints will define the particular size of the retrofit, as experienced in this study. Costs, while important, often are the outcome of site feasibility and area constraints and can be estimated off of the quantities of the site-specific designs. So the resulting primary research question focuses on performance and is:

1. What is the PCB and mercury load reduction catchment efficiency of retrofits?

Secondary Research Questions

Secondary research questions allow implementing agencies to make more informed decisions regarding both retrofit and other sediment control options. Answering these questions is helpful, but they are not critical to the usefulness of the study. Study scope, budget, and schedule limit how well they can be answered. The secondary research questions are:

2. What is the cost-effectiveness of each retrofit type in units of cost per mass of pollutant removed over the project life-cycle?[1]
3. Can PCB and mercury removal effectiveness be tied to a better-studied surrogate constituent, such as turbidity, so that performance is better understood?
4. Can site-specific knowledge of the association of PCB and mercury to specific particle sizes be used to better understand performance?
5. Is the measured level of mass removal within the range of pollutant removal that is expected from a conceptual understanding of removal mechanisms?
6. What is the level of confidence in cost-benefit ratio comparisons among the treatment types? (Though this is a very desirable outcome, Task 5 does not have a sufficient number of scenarios affecting cost to answer this question. There are insufficient installations of each type to develop confidence intervals for costs. Since size is a strong indicator of cost, at least three installations at different sizes would be needed of each BMP type. Regardless, this question is included here for completeness, and site-specific cost comparison among BMP types is the assumed outcome of this study.)
7. How does fate and transport of methyl mercury and total mercury affect the design of the BMP? (For example, elevating discharge pipes within the gravel underdrains of

[1] Due to the study duration, maintenance cost and pollutant removal performance will need to be extrapolated to an estimated life of the retrofit.

bioretention BMPs encourages infiltration and pollutant load reduction. But this may also encourage methylation.)

8. What is the maintenance frequency necessary to maintain measured performance levels? (The limited study period will not be sufficient to establish a relationship between maintenance and performance.)
9. What is the maintenance frequency necessary to avoid expensive sediment disposal costs? (The limited study period will not be sufficient to establish maintenance frequencies necessary to avoid additional sediment handling and disposal costs.)

Level of Confidence

The level of confidence in the answers to the above research questions depends on sample size. A power analysis is traditionally used to select the proper sample size to draw statistical conclusions for BMPs that show adequate performance and reasonable variability (See Appendix A). In this study, the Technical Advisory Committee and the study team have noted that performance can be estimated by application of removal mechanisms and corroboration of these known removal mechanisms would establish a fairly high level of confidence in the research results. So instead of traditional tests of significance, this study plan focuses most of the monitoring effort on data gaps and applies a minimal effort to corroborate better-understood performance characteristics.

For example, treatment of solids by many BMPs has been described in the International BMP Database. However, the ability to use that information is limited because of a lack of knowledge of the occurrence of PCBs within solid size fractions and how that relationship varies among candidate watersheds for retrofit (watersheds with elevated pollutant concentrations). The use of the International BMP Database is also hindered by a lack of understanding of hydraulic performance of the retrofit pilots that have been custom-designed to fit into available areas. Consequently, a portion of the budget is allocated to analysis of PCB for different size fractions and monitoring of bypass.

Catchment Efficiency Calculation

In calculating the catchment efficiency, the concentration of bypassed volume is conservatively assumed to equal the concentration of the total volume from the catchment. The total catchment load reduction removal fraction can then be expressed as the product of the fraction treated and the BMP load reduction fraction for the water treated, as derived in Equation 1. It is helpful to consider Figure 1 to visualize pollutant pathways.

Equation 1. Derivation of load reduction catchment efficiency assuming bypass water quality equals inlet water quality

$$catchment\ load\ reduction\ fraction = \frac{\Delta Lc}{Lc} = \frac{Qi}{Lc} * \frac{\Delta Lc}{Qi} = \frac{Qi}{Qc * Cc} * \frac{Lc - (Lo + Lb)}{Qi} = \frac{Qi}{Qt} * \frac{Lc - (Lo + Lb)}{Qi * Cc}$$

$$= \frac{Qi}{Qc} * \frac{Lc - (Lo + Lb)}{Qi * Ci} = \frac{Qi}{Qt} * \frac{QcCc - (QoCo + QbCb)}{Qi * Ci} = \frac{Qi}{Qt} * \frac{QcCi - (QoCo + QbCi)}{Qi * Ci}$$

$$= \frac{Qi}{Qt} * \frac{QcCi - QbCi - QoCo}{Qi * Ci} = \frac{Qi}{Qt} * \frac{QiCi - QoCo}{QiCi} = \frac{Qi}{Qc}\left(1 - \frac{Lo}{Li}\right)$$

where
L_c is pollutant load from the catchment
L_b is the bypassed load
L_o is BMP outlet load that is treated
L_i is inlet load (no bypass)
Q_i is inlet volume (no bypass)
Q_c is the total volume from the catchment
C_c is the concentration from the catchment
C_i is the concentration to the BMP, $C_i = C_c$
C_b is the concentration bypassed, $C_b = C_c$

By similar deduction used , the load reduction within the BMP can be shown to equal the sum of the fraction of concentration reduction to the fraction of volume reduction, and then subtracting the product of the two as shown in Equation 2.

Equation 2. Treated (no bypass) load reduction based on concentration reduction and volume reduction

$$treated\ load\ reduction\ fraction\ without\ bypass = 1 - \frac{Lo}{Li} = \frac{Ci - Co}{Ci} + \frac{Qr}{Qi} - \frac{(Ci - Co)}{Ci} * \frac{Qr}{Qi}$$

where
Q_r is the retained volume via infiltration and evapotranspiration, $Q_r = Q_i - Q_o$

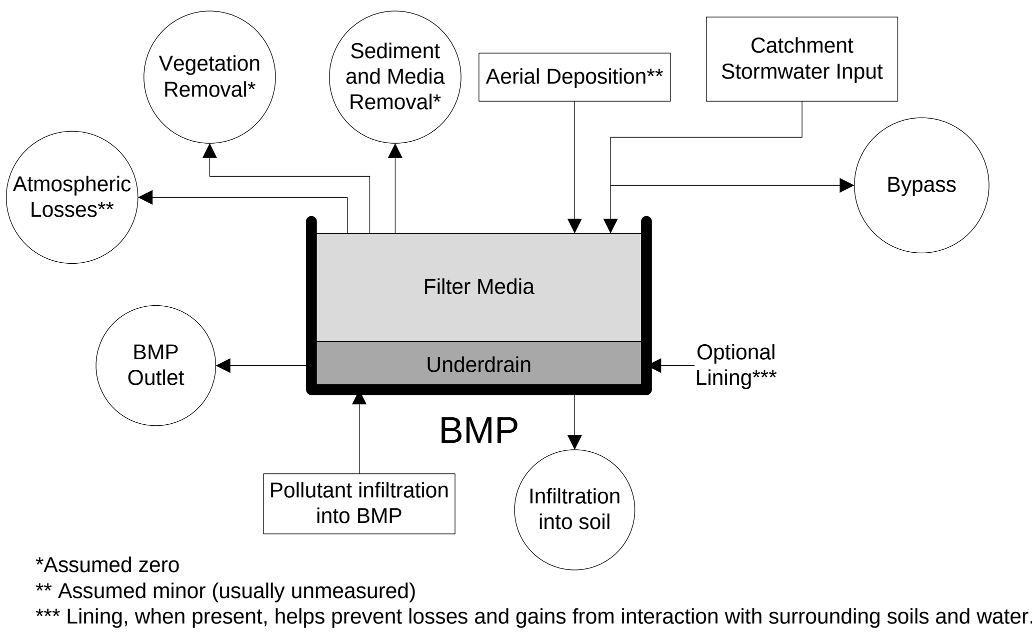


Figure 1. Typical BMP system and pollutant pathways

Overview of Performance Measurement Options

A review of the available study designs is presented here to understand the methods, value, and constraints of each design. This information is helpful in identifying which study designs are appropriate for the various research questions (addressed in the following section). This overview of study designs describes how data are collected and how the data are used to answer the primary study questions. Each monitoring scenario will seek to measure the major inputs, outputs, and losses within an open system, as illustrated in Figure 1. The following section will explore characteristics of each BMP type and potential monitoring techniques.

To answer many of the primary research questions, the mass of pollutants that is prevented from discharging from the sites must be quantified. This is accomplished by accounting for the pollutant inputs and losses of each BMP type. For example, the mass balance of a filtration BMP is illustrated in Figure 1 along with the mass balance prior to BMP installation.

An ideal study design quantifies all the inputs and losses of the system. The study designs discussed here address major inputs and losses, but not all. The study designs discussed, along with their strengths and weaknesses, are

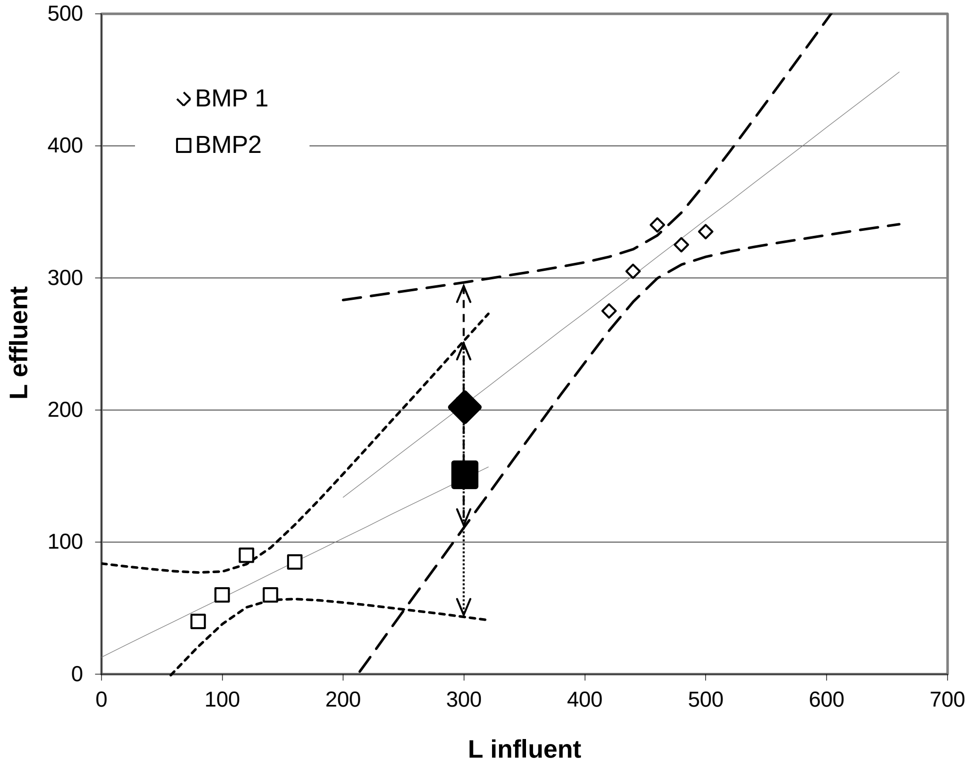
- Influent and effluent sampling:
  Influent and effluent sampling tests water going into and discharging from a particular BMP for a particular storm event.
- Sediment sampling
  Sediment sampling occurs within the BMP and is used to estimate cumulative load removed over several storms. Sediment sampling can occur in dry periods.
- Before and after construction sampling
  Before and after construction concentration measurement occurs at the same discharge location. This introduces variability because the storms monitored before construction will not have the same characteristics as those after construction.
- Paired watershed sampling
  Paired watershed concentrations measure the discharge locations at two watersheds that are as similar as possible, except one has the retrofit pilot. While the storms monitored are the same, inevitable differences in the watersheds often lead to unexplainable variability.
- Volume measurement
  Volume measurement is used for load analysis and the measurement can occur at influent and effluent locations, and within the BMP for individual storms.

The following subsections provide more detail on each method.

Influent versus Effluent Sampling

This type of study results in paired samples. Paired samples are beneficial because fewer samples are needed to establish predetermined levels of pollutant reduction, compared to unpaired samples. This can result in substantial cost savings for sample collection and sample analysis. The drawback, however, is that proper representation of influent and effluent concentrations requires special monitoring equipment consisting of flumes or weirs and monitoring cabinets. Though the equipment may not be very expensive, the proper installation of flumes is infeasible for many space-constrained retrofit applications. Weirs are a good alternative for many situations (WERF 2009). Common limitations to using flumes and weirs include sheet flow influent and subsurface underdrains that connect the effluent directly to existing storm drains. In these cases, installing flumes or weirs can be infeasible or very expensive.

A limitation of comparing BMP performance among a limited number of locations is that the BMP loading scenarios are likely to be different. This makes comparison of BMP performance among BMP types difficult. In regression analysis, the confidence intervals of the predicted values increase as *x*-values at the point of comparison deviate from the mean *x*-value of the data. This is illustrated in Figure 2 for two non-overlapping data sets. In the figure, prediction of BMP performance (y-values) is being attempted at a common influent value. Because of the extrapolation involved in making this point-estimate, the confidence intervals for the two regression lines are wide enough to overlap. This means the two BMP performances cannot be said to be statistically distinguishable.



Note: Confidence intervals are shown as the vertical dashed and dotted lines with arrows. The solid black square and diamond are estimates of effluent at a hypothetical influent value of 300. Confidence intervals for effluent estimates calculated from the regression line expand as the distance between the hypothetical influent *x*-value and the mean *x*-value of the data increases. Although the estimates may reflect the different treatment efficiencies of the BMPs, the expanded confidence intervals overlap, which prevents finding a statistical difference between the two BMP performances.

Figure 2. Comparison of two hypothetical non-overlapping BMP regressions (Caltrans 2010a)

The end result is that the purely statistical approach does not generate useful results that can be used to rank BMPs. Sediment analysis that complements water quality analysis may help in making distinctions that a strictly statistical analysis cannot make.

Sediment Sampling

Sediment sampling can be used in lieu of influent and effluent monitoring for estimating the total mass of pollutant removed. Analysis of accumulated sediment and filter media can be used to estimate the total mass of more conservative pollutants.[2]

A limitation to the sediment sampling approach is that it does not generate as many data points to statistically compare among BMP types. Another limitation is that influent monitoring data are not available to describe how the mass removal estimates may be sensitive to influent loading. Influent monitoring will help characterize pollutant loading.

[2] In the context of the sediment analysis study design option, "conservative pollutants" are those that are not substantially lost to volatilization or plant uptake in between periods of sediment analysis. Sediment analysis underestimates performance where volatilization or plant uptake is substantial.

Another limitation of sediment sampling is the potential error resulting in non-homogeneous pollutant distribution within the sediment. Compositing multiple samples will better characterize the sediment, much as the collection of several aliquots throughout a stormwater runoff event can better represent the total volume of water. For hydrodynamic separators (HDS), mixing the removed sediment before compositing can further help achieve more homogeneous samples. Yet another limitation, for bioretention, is the presence of mulch and grasses that both obstruct access to underlying sediment and retain pollutants. Tree well filters are also expected to have a mulch layer, though the tree is not expected to interfere with sediment collection.

Consequently, the effectiveness of sediment sampling depends on the type of retrofit BMP. HDS are the best candidates for sediment sampling. This is because at the start of the study the sumps are cleaned and therefore empty and the entire mass of retained sediment is removed at each maintenance event (sump cleanout). Conversely, sand filters, bioretention, and tree well filters have background sediment (planting media or filter media) that obscure pollutant accumulation. Since pollutants tend to accumulate on the surface of media (Li and Davis 2008), surface sediments should be targeted when sampling these systems. Coring these systems and compositing the core sediments will most likely result in further dilution of the PCBs retained in the media, making quantification more difficult. For all systems, larger pieces of litter and vegetation may be difficult to include in the analysis. A conservative approach is to exclude larger material and assume these have little association with PCBs. If the budget allows, this material can also be analyzed, but this analysis is of lower priority.

Before and After Construction Sampling

Pollutant removal can also be estimated by monitoring discharge quality before and after the construction of the retrofits. This may be attractive for green street projects that have multiple BMPs with multiple influent and effluent locations. Monitoring all of these individual systems is almost impossible because of space constraints. Note that since the data from before/after implementation are unpaired, variability is larger and the number of samples required to show significance is expected to be much higher than for paired samples.

Paired Watershed Concentration Measurements

The paired watershed approach is used when monitoring the influent of the BMP is infeasible. A paired watershed is monitored in lieu of the actual watershed contributing to the BMP. For example, sheet flow into a BMP is difficult to monitor. A paired watershed needs to have similar pollutant-generating characteristics as the actual watershed contributing to the BMP. The paired watershed also needs to be located nearby so that the same storm patterns influence both the test and paired watersheds. A limitation to paired watersheds is that there is typically no way to be sure that the influent to the BMP is the same as what is measured in the paired watershed. If the runoff is different, then the data analysis should assume unpaired data.

Volume Measurement

Volume measurement is critical to estimating catchment efficiency. Volume measurement can be obtained by totalizing influent, effluent, and bypass flow measurements. Influent flow measurement can be difficult. Alternatively, several hydrologic methods are available to estimate influent volume using watershed size, land cover (imperviousness), and storm characteristics. An alternative to effluent flow monitoring is to estimate the amount of water that is infiltrated into underlying soils or that is retained within planting media and transpired through vegetation. To provide data to estimate infiltration, observation wells can be placed within the BMP and fitted with depth sensors for continuous monitoring.

Study Design Options for Primary Data Objectives

The study design options discussed previously are matched to the primary research question. The primary research question requires two data objectives: mass removal data and water volume data. The selected combinations of these options are presented in the "Final Study Plan Design" section.

Data Objective 1: Mass Removal Estimates

Filter media or soil sediment sampling. Sediment sampling has a high value for estimating mass removal because a single composite sample of retained sediment over a period of several storms can yield an estimate of load removal for the constituents analyzed. However, some influent characterization would also help explain mass removal performance. Another complicating factor is the uneven distribution of pollutants in filter media and soils. Multiple inlets also reduce the accuracy and increase the cost of this sampling method. Consequently, this method has better value when applied to systems that isolate retained sediment (e.g., HDS).

Options:

- ❖ Influent and effluent water quality monitoring for statistical analysis.
- ❖ Drainage area discharge concentration analysis before and after BMP installation. This study design has low value because the statistically derived sample size is too large in comparison to other study designs.

Data Objective 2: Volume Measurements

Volume measurement is a critical component to estimating the volume treated, volume retained, and volume bypassed. Treated volumes are compared to understand removal mechanisms. Retained volumes are assumed to reduce all associated pollutant load. Bypassed volumes are assumed to be untreated.

Options:

- ❖ Water depth within the BMP to estimate bypass and retention. The depth to an overflow riser is used to track bypass. A v-notch in the overflow riser or spillway may increase the accuracy of the overflow discharge equation. Inter-event drawdown of retained water can be estimated by tracking depth through the BMP as water infiltrates into underlying soils.
- ❖ Flow measurement at inlet, outlet, and bypass to estimate treated, retained, and bypassed volumes. (see Equation 1).

BMP Processes and Potential Water and Sediment Data Collection

The treatment mechanisms that occur in the BMP help inform selection and control of the study variables. These treatment mechanisms, also called unit processes, may include physical, chemical, or biological processes. The primary physical and chemical processes that are responsible for removing contaminants in the five types of retrofit pilots include the following:

- Sedimentation – The physical process by which suspended solids and other particulate matter are removed by gravity settling. Sedimentation is highly sensitive to flow, particle size, and influent concentration. Effluent quality is less consistent compared to other mechanisms due to high dependence on flow regime, particle characteristics, and scour potential.
- Filtration – The physical process by which suspended solids and other particulate matter are removed by means of porous media. Compared to sedimentation, filtration provides more consistent effluent quality over a wider range of influent concentrations.
- Sorption – The chemical process of adsorption and absorption that occurs on the surfaces of substrate media, plant roots, and sediments, resulting in short-term retention or long-term immobilization of contaminants.
- Chemical Precipitation – The conversion of contaminants in the influent stream, through contact with the substrate or root zone, to an insoluble solid form that settles out. Consistent performance often depends on controlling other parameters such as pH.
- Infiltration – The physical process of pollutants moving into underlying soils or groundwater due to their presence in the water column. Infiltration is really a flow regime rather than a removal mechanism. As flows enter and pass through subsurface soils, filtration and sorption are the actual removal mechanisms by which pollutants are removed from the infiltrated water. However, from a surface water perspective, it can be considered a removal mechanism since it transfers the associated pollutant load of the infiltrated surface water flows to underlying soils or groundwater. Infiltration is a very reliable removal mechanism when BMPs are designed to soils properties. Without verification of permeable soils and separation to groundwater, infiltration is not a reliable mechanism – particularly in the Bay Area, where less permeable soils dominate.

Contaminant removal is also accomplished through biological processes such as biodegradation and phytoremediation (plant uptake), which are:

- Aerobic/Anaerobic Biodegradation – The metabolic processes of microorganisms, which play a significant role in removing organic compounds and nitrogen in filters.
- Phytoremediation – The uptake, accumulation, and transpiration of organic and inorganic contaminants, especially nutrients, by plants.

The following subsections provide a brief description of the seven retrofit types being evaluated in this pilot study, a review of the unit processes involved in each device, and possible monitoring techniques that might be appropriate for each BMP type. The final selection of the quantity and type of data to collect is presented in the “Study Optimization” section. Each type of data is characterized by how it will be analyzed statistically. Generally, paired data require fewer samples than unpaired data so there is more value in collecting this type of data. Directly measured variables do not require comparison to other variables. Sediment analysis of disposed sediment and cost are both directly measured variables. These data typically result in one data point per pilot BMP. Because mass removal by sediment sampling is directly measured, the value is high, but the ability for statistical comparison is limited by the number of pilot locations rather than by the number of samples taken.

Bioretention (with underdrain)

Bioretention is a slow-rate filter bed system. It is planted with macrophytes (typically shrubs and smaller non-woody vegetation). The major sediment removal mechanism is physical filtration through the planting media. Some minor sedimentation can be expected at the inlet to the bioretention bed. Dissolved constituents can be removed by sorption to plant roots in the planting media, which typically contains clays and organics to enhance sorption. The effluent quality from bioretention BMPs is expected to be consistently lower than for sedimentation-type BMPs. The number of influent and effluent samples required to establish reduction of sediment concentrations is expected to be relatively low because of consistent performance. However, release of fines in the first few events is common and should be considered in targeting storms and evaluating the data for long-term performance. The bioretention systems with underdrains are lined, so infiltration is assumed to be minimal. This factor also removes the influence of subsurface flow and allows the bioretention media to be analyzed over time to quantify the removal of more conservative pollutants such as PCBs.

The possible monitoring techniques for bioretention are:

- Influent sampling to characterize load (unpaired data)
- Influent sampling to pair to effluent (paired data)
- Sediment sampling to quantify removal (directly measured variable)
- Discharge monitoring before/after installation (unpaired data)
- Bypass flow monitoring to calculate catchment efficiency

Permeable Pavement

Permeable pavement is a subsurface storage, detention, and infiltration system. Storage is accomplished within a stone matrix below the permeable pavement. Sediment removal mechanisms include physical filtration through the porous pavement, choker course (if used), and flow through the stone matrix. If the entire storage volume is infiltrated into underlying soils then all influent sediments are removed via surface filtration at the surface of the underlying soil, while the interface with the underlying soil removes the remaining sediment. Many permeable pavement systems allow water in excess of the storage volume to discharge to the stormwater drainage system.

Because of their design, permeable pavements are difficult to monitor. Influent monitoring of the system is nearly impossible so the paired watersheds approach can be used instead of directly monitoring influent water. Effluent monitoring is also difficult because of subsurface construction. Analysis of retained sediment proves difficult because of limited access to subsurface storage. Pre-and post-project analysis may be feasible. For permeable pavements that are designed to infiltrate, observation wells can be used to measure infiltration.

The possible monitoring techniques for permeable pavement are:

- Sediment sampling to quantify removal (directly measured variable). Sediment that accumulates in the subsurface storage will be difficult to monitor. Uneven pollutant loading to the surface of the pavement and the underlying storage volume will require many more samples than filters with more evenly distributed loading.
- Volume reduction via observation wells (assuming there is some dead storage rather than completely free draining to the storm drainage system).
- Paired watershed discharge monitoring and porous pavement discharge monitoring (assumed paired data).[3]
- Discharge monitoring before/after installation (unpaired data).
- Bypass flow monitoring to calculated catchment efficiency

Flow-through Biotreatment (Swale)

Swales are grassy vegetated channels that convey concentrated stormwater flows at fairly shallow depths relative to the height of the vegetation. Primary removal mechanisms include sedimentation and infiltration where soils are sufficiently permeable. Physical filtration also occurs via plant material. Because sedimentation processes dominate, sediment and removal performance is inconsistent. This means that a relatively large number of influent and effluent samples is expected to be required to show a statistically significant removal. .

The possible monitoring techniques for swales are:

- Influent sampling to characterize load (unpaired data).
- Influent sampling to pair to effluent (paired data).
- Sediment sampling to quantify removal (directly measured variable). Uneven pollutant loading to the surface of the swale will require many more samples than are needed for filters with more evenly distributed loading. Because flow through the vegetation is a primary removal mechanism, sediment sampling should include vegetation. The distribution of pollutants across soil and vegetation surfaces makes this sediment sampling particularly difficult and expensive for swales.
- Discharge monitoring before/after installation (unpaired data).
- Design flow exceedances to explain variability in performance since BMP is inline (no bypass features).

Tree Wells

Tree wells are similar to the bioretention systems described previously, except that a small tree is used in place of shrubs and smaller non-woody vegetation. Removal mechanisms and monitoring approaches are the same. Since filtration is the primary removal mechanism, the

[3] For highly variable PCB concentrations, it is very unlikely that a paired watershed could serve as a control for permeable pavement systems.

number of storms required to characterize the effluent is expected to be relatively small. If the schedule allows, BASMAA may consider testing some tree wells with open-bottom vaults that are installed on a course of gravel overlying compacted soils. USDA research has shown that these open systems allow for the development of a much larger urban tree that has other beneficial effects, such as reducing the urban heat island and improving rainfall interception. Observation wells, similar to those proposed for permeable pavement, could also be used for infiltrating tree wells.

The possible monitoring techniques for tree wells are:

- Influent sampling to characterize load (unpaired data)
- Influent sampling to pair to effluent (paired data)
- Sediment sampling of planting media to quantify removal (directly measured variable)
- Discharge monitoring before/after installation (unpaired data)
- Bypass flow monitoring to calculated catchment efficiency

Catch Basin Media Insert

Catch basin media inserts are canisters, bags, or bags of filter media that are retrofitted into a catch basin. Stormwater typically flows directly from a curb cut or drop inlet to the filter, making influent monitoring difficult. This also greatly restricts the amount of sedimentation that occurs prior to the filter. Since filtration is the primary removal mechanism, the number of storms required to characterize the effluent is expected to be relatively small. Influent monitoring, however, is challenging because there is little space to locate influent flow monitoring equipment.

The monitoring approach used by Caltrans for similar systems was a mass balance based on analysis of the filter media and effluent flow monitoring (Caltrans 2004). If effluent flow monitoring is infeasible, calculation of removal efficiency will not be possible, but total mass removed can still be determined from the filter media analysis.

The possible monitoring techniques for catch basin media inserts are:

- Influent sampling to characterize load (unpaired data). This is not likely a viable option due to the configuration of this BMP.
- Influent sampling to pair to effluent (paired data). This is not likely a viable option due to the configuration of this BMP.
- Filter media analysis to quantify removal (directly measured variable).
- Discharge monitoring before/after installation (unpaired data).
- Bypass flow monitoring to calculate catchment efficiency

Sand Filter

Open-bed filters rely on filtration as the primary removal mechanism for particulate pollutants. The addition of other types of media may enhance the sorption removal mechanism. As with filter-based BMPs, consistent effluent quality is anticipated. Since filtration is the primary removal mechanism, the number of storms required to characterize the effluent is expected to be relatively small. Influent and effluent monitoring is a good candidate for this type of BMP, and bioretention media analysis will help corroborate load removal for fairly conservative pollutants such as PCBs.

The possible monitoring techniques for sand filters are:

- Influent sampling to characterize load (unpaired data)
- Influent sampling to pair to effluent (paired data)
- Sediment sampling to quantify removal (directly measured variable)
- Discharge monitoring before/after installation (unpaired data)
- Bypass flow monitoring to calculate catchment efficiency; bypass could be overtopping of the media bed or via pumps that bypass the filters

Hydrodynamic Separators

Hydrodynamic separators rely on sedimentation as the primary removal mechanisms for sediment and particulate pollutants. Performance is highly dependent on flow rate, particle size, and BMP geometry and size. Since effluent quality is highly variable, these systems are expected to require a relatively large number of samples to demonstrate statistically significant reduction. Because of budget constraints, analysis of retained sediment may be an appropriate alternative. However, particles size distribution analysis of the influent water is needed to understand the size fraction that has been retained compared to the influent size fraction. Effluent analysis provides the particle size fractions that are escaping the HDS and provides insight into particle breakdown in the sump by comparing effluent results to what is expected from the influent and sump mass balance. Sediment should be analyzed when the device is cleaned. The total flow or storm sizes between maintenance intervals should be recorded. It is important that the study period include high-flow events to test whether sediment will scour from the BMP.

The possible monitoring techniques for hydrodynamic separators are:

- Influent sampling to characterize load (unpaired data)
- Influent sampling to pair to effluent (paired data)
- Sediment sampling to quantify removal (directly measured variable)
- Discharge monitoring before/after installation (unpaired data)
- Bypass flow monitoring to calculate catchment efficiency

Variables Affecting Water and Sediment Data

Variables that affect water quality and sediment quality data are discussed here. Some variables can be controlled and others should be measured to determine their effect on water quality and sediment quality. Inevitably, some factors will be beyond the control of the study. The impact of these factors should be discussed by the project team. In some cases, models or observations from other studies can be used to explain the effect of unmeasured or uncontrolled factors.

BMP Siting, Land Use, and Land Cover

The location of the BMP can affect water quality because land use and land cover affect sediment mobilization and pollutant concentrations within the sediments. The sites in this study were previously selected based on the feasibility of retrofits within watersheds known to have elevated PCB concentrations, as described in the project proposal (BASMAA 2009). This is not a random design. This means the results of the study may not be applicable to other Bay Area locations where pollutant loading is not as high. Early data analysis may justify switching the monitored BMPs to those BMPs that are in areas with higher loading (discussed in the "Final Study Plan Design" section).

Land use is often used as an indicator of pollutant loading. The land uses of the areas selected for retrofit include industrial, commercial/mixed use, roads/rail, institutional, and residential. Because past use of PCB and past PCB and mercury handling practices, age of the land use is also important. However, PCB analysis by the San Francisco Estuary Institute (SFEI) showed that PCB concentration patterns were patchy within larger urban watersheds that were known to have elevated PCB concentrations . This finding indicates that mass reductions of PCBs will require site-specific sampling of influent loads or site-specific quantification of mass removed from the retrofit. In contrast to PCBs, mercury has contemporary sources and data suggests areas with elevated concentrations are not as pronounced. Also, there was little evidence suggesting a correlation between PCB and mercury concentrations. (Yee and McKee 2010)

The impact of differing influent concentrations on BMP performance can be anticipated in future retrofit scenarios by measuring influent loading to the BMPs. This adds substantially to the cost of the study, and may not be possible at all locations due to budget constraints. While determining point estimates of mass removed is possible without knowing influent loading, it makes comparison of performance among BMPs nearly impossible for scenarios where a limited number of BMPs are distributed in a highly heterogeneous watershed.

BMP Design and Hydraulic Loading

BMP design can have a substantial impact on effluent water quality and the quantity of sediment retained in the BMP. Consequently, the full range of BMP sizes that would be used in future retrofits should be studied. Due to unknown site constraints at future locations, this is impossible. Instead, best judgment will be used to pick typical BMP sizes that would be used in future retrofits. If future retrofit locations have space constraints that result in BMP sizes much smaller than those tested, then a reduction in the estimated load removal may be appropriate.

Another aspect of the BMP design that could affect performance is the design of bioretention systems and tree wells with inverts that allow infiltration. Though this type of design variation may not be monitored in the study, the impact on overall performance can be estimated (Equation 1).

Properly designed, BMPs may, depending on rainfall events and antecedent conditions, exceed their design capacity a few times a year. But this may never occur throughout a study period–even in a year with typical rainfall depth. This results in very little data that can be used to corroborate hydraulic design assumptions for filtration BMPs. However, in the retrofit environment, typical hydraulic loading is much higher due to space constraints, thus providing an opportunity to collect more data near design capacity (Brian not sure what you mean here, how do we know that operating near design capacity might affect effluent quality? peter). Therefore, monitoring site selection for filtration BMPs of the same type should give preference to those BMPs with higher hydraulic loading. Other BMPs, e.g., swales, will have effluent quality that is extremely sensitive to BMP size because they rely on sedimentation, so monitoring of sedimentation BMPs should be close to "typical" retrofit sizing practices.

Operation and Maintenance Parameters

Maintenance frequency can impact the performance data. Maintenance that is more frequent than standard practice may enhance removal for sedimentation BMPs by reducing scour and maintaining larger sedimentation volumes. The reverse could be true for filtration-type BMPs where the accumulation of sediment on the surface of the filter can actually improve performance. For sedimentation BMPs, less frequent maintenance may have a negative impact on performance. Sediment levels may exceed the sediment capacity of the BMP, decreasing the volume for sedimentation and increasing scour. Vegetation management is another maintenance practice that may affect BMP performance. Cutting and leaving vegetation can cause pollutant cycling within the BMP. Cutting too much vegetation decreases the infiltration and evapotranspiration capabilities of the BMP. Soil conditioners may also affect performance. Aeration and compost amendments may improve the soil structure and increase infiltration, but they may also contain pollutants.

Maintenance during the study period should be sustainable over the life-cycle of the BMP. Overly aggressive maintenance can be a result of the special attention that a study brings, but this can skew cost estimates (Caltrans 2004). Due to the schedule constraints of this study, optimization of maintenance practices to observe longer-term effects of changes in practices is not possible. Optimizing maintenance practices could be addressed in future study efforts.

Long-term maintenance frequencies, such as those needed to restore filtration capacity, might be addressed by measuring infiltration rates, water levels, or drain times. Determining the frequency of major maintenance activities is important in calculating the life-cycle cost.

Pollutant Behavior

Pollutant behavior also impacts performance data because it is not usually feasible for data collection methods to account for all the pollutant paths within a BMP as illustrated in Figure 1.

Relatively conservative pollutants (e.g., PCB) are more readily monitored because they do not typical change phase in the limited time of a storm treatment episode. Since PCB are strongly associated with sediments, particle size is a top priority. Thorough study of these associations could reduce future cost of monitoring and load estimation.

Non-conservative pollutants will have an impact on sediment-based analysis if conditions are present that would cause their transformation. Losses that affect mass balance approaches (e.g., volatility of Hg) are difficult to control in an in-situ study. It is also difficult to consistently address all inputs and losses among all BMP types when the site constraints and budget require different monitoring approaches for different BMP types. Without accounting for additional losses, sediment analysis will be conservative. Concurrent water quality analysis of total and dissolved fractions may result in a more complete understanding of pollutant behavior within the BMP.

A more thorough discussion of sediment, PCB, and mercury transport is provided in McKee et al. (2006). This report summarizes data from various sources. The 90th percentile PCB concentration in sediment among over 100 sites was approximately 0.5 mg/kg. This statistic may be a useful cutoff point for differentiating BMP performance. Similarly, the 90th percentile mercury concentration in sediment was approximately 0.7 mg/kg. The total mass discharge will depend on sediment concentrations in water. Since the retrofits are designed to target areas with unusual (i.e., elevated) concentrations compared to what is typical in the watershed, it will be important to describe the influent conditions of each monitored BMP in relation to historic sediment analysis as provided in McKee et al. (2006).

Another confounding factor is that the concentration of pollutants of concern (e.g., PCBs) within the sediments in a drainage area may decrease with time. Such a change would likely affect the accuracy of the long-term estimate of mass removed. If PCB or mercury concentration in sediments decreases, pollutant mass removed should also decrease. This factor is being partially addressed outside the retrofit study, where the Ettie Street pump station area was being resampled in 2012 and compared to previous monitoring in 2006 (Mangarella, personal communication, 21 August, 2012).

The size and type of particle that PCBs are attached to also affects BMP performance. Larger particles are more easily filtered. Inorganic particles are more easily settled. To better understand measured performance and to better predict future performance, particle size distribution and associated contaminant concentrations could be measured. Wet sieving techniques work well for collecting this type of data. Based on site investigation of future retrofit locations, wet sieving may also help determine the effectiveness of these technologies. Research on the approximate limits of effective particle removal for sedimentation BMPs, suggest sieve sizes of around 63 and 25 microns for wet sieving. This is consistent with the ranges suggested by Mangarella et al. (2012). This will result in PCB and total Hg relationships to solids for three particle size ranges: greater than 63 microns, 25 to 63 microns, and less than 25 microns. Analysis such as suspended solids and volatile suspended solids can help estimate the amount of lighter, organic particles relative denser, inorganic particles.

Storm Intensity

PCB and to a lesser extent mercury loading is largely a result of historic activities so the accumulation in sediments of pervious areas has been observed. Mobilization of these sediments may require exceeding site-specific intensity and volume thresholds. Intensity is critical to detach and mobilize particles. SFEI research reported that climate and local conditions can impact the dispersion of pollutants from areas with elevated concentrations of pollutants. In particular, storm intensity was noted as being a factor that would cause differences to be observed within a single watershed (Yee and McKee 2010). Storm volume may also be critical to exceed any depression storage within the pervious areas, and provided sustained transport in the storm drain system. However, the precise effect of storm intensity and volume on the mobilization of PCB-contaminated and mercury-contaminated sediments has not been established. Since storm intensity and volume can impact influent water quality, monitoring could target storms suspected of contributing more of the PCB loading. Also, during-storm observations of pavement run-on from erodible surfaces may help explain storm-to-storm differences in solids loading.

Soil Permeability

Soil permeability may be an important factor for unlined BMPs because more permeable soils increase pollutant load reduction via infiltration. Observation wells may be useful for observing actual infiltration rates of bioretention and permeable pavement systems. These systems have substantial storage areas overlying native soil and can be designed with elevated outlets that cause a zone of standing water. Observation wells do not work well on swales if the planting media does not become saturated. Instead, infiltrometers may be more appropriate for swales, but measurement during the storm event is more desirable to account for soil saturation and temperature effects.

Uncontrolled Variables and Study Assumptions

The following assumptions were adapted from the PSGM (Caltrans 2009). These assumptions should be revisited during the BMP design phase and after the BMP is constructed.

- Site Assumptions
  - The range of observed runoff concentrations is representative of candidate sites for future retrofit.
  - There are no unaccounted for external sources. For instance, illicit discharge of pollutants directly to the BMP surface does not occur.
  - There is no base flow or groundwater intrusion into the BMP.
  - Differences in storm patterns throughout the Bay Area are not sufficient to change the performance measurements.
- BMP Operation Assumptions
  - Installation of an impermeable liner or a concrete wall will prevent infiltration losses from the base and sides of the BMP. This may not be the case if the liner is installed improperly or has significant tears that occur during installation.
  - Volatilization of pollutants within the BMPs is negligible. This assumption only applies to sediment sampling, because influent and effluent water samples account for all losses within a BMP.
  - Bacterial growth in soil or media does not have a significant impact on the hydraulic and treatment performance of the BMP.

- There is no short-circuiting of flows within the BMP.
- Maintenance practices are carried out as needed and do not have an adverse impact on BMP operation. An example of when this assumption may not be the case is nutrient addition due to decay of unmaintained vegetation at the surface of the BMP.

- Monitoring Assumptions
  - Data collected from a few sites and over a relatively short time span will accurately represent how the BMP works over wider distribution and over longer time frames.
  - Monitoring that is carried out accounts for all significant inflows and outflows from the BMP.
  - Storm event characteristics are independent of previous storm events.

Final Study Plan Design

The monitoring approach for each retrofit is optimized to answer the primary research questions for each type of BMPs and then focus on areas where existing knowledge is most lacking. This approach allocates the monitoring budget to as many BMPs as possible, as long as the minimum sample sizes can be obtained. Where possible, sediment sampling methods are used in lieu of more expensive water quality monitoring efforts. Monitoring methods that generate more variable data or that do not relate directly to the primary research questions are lower priorities. As the monitoring team fine tunes their sampling and analysis plans (SAPs) and submits budgets, adjustments to this study plan will be inevitable.

Statistical Testing Considerations

In a traditional test of a treatment, the null hypothesis is that there is no difference between the treated and untreated samples (i.e., a BMP does not work). The TAC has observed that tests have been performed for many of these technologies in previous studies for the treatments (pilot BMPs) used in this study. However, the tests were not necessarily performed for mercury and PCBs. So to establish confidence in the observed level of treatment of PCB and mercury, two pieces of evidence are needed. The first is to validate (rather than statistically substantiate) performance of the BMPs to one of the common constituents (e.g., suspended solids). The second is to establish the relationship of PCB and mercury to constituents tested in previous studies (solids) so that BMP performance can be estimated. To the extent that this relationship is site-specific , the application to other locations may require additional analysis to develop, if possible, an appropriate regional or non-site specific ?performance translator.

For the constituents measured here and in previous studies (e.g., solids), statistical comparisons can be made between the measured performance in this study to known performance ranges found previously. When constructing such a test, the null hypothesis is that there is no difference between BMPs of the same type. Unfortunately the small differences in performance that are expected will drive sample size to unreasonable levels, given typical Type I and Type II errors (0.1 and 0.2, respectively). So considering the differences that could be detected given traditional Type I and Type II error values and reasonable sample size (e.g. <10), the BMP performance would be quite near zero, if not below. Hence, it is more practical to simply observe the summary statistics of the pilots and compare these to the range of data in sources such as the International BMP Database. See the subsection “BMP Comparisons” under section “Reducing Analytes and Monitoring Alternative Locations” for a more thorough discussion of approaches to analysis. Examples of traditional statistical designs are included in Appendix A.

Selected Sample Size

Though not supporting a primary research objective, it is instructive to consider the sample size requirements for establishing statistical significance of treatment of suspended solids for the traditional test of statistically significant treatment. The application to PCB is expected to modestly underestimate sample size, while application to mercury may more drastically underestimate sample size.

Generally, the sediment-based power analysis shows that fairly large sample sizes (greater than 30) are needed to establish statistically significant removal rates for unpaired data (Appendix A). The cost of such a monitoring program is fairly high and limits the primary research questions that are being answered within the allocated budget. For influent and effluent sampling that is paired by storm events, the required number of samples for all retrofit types is generally less than 8, but that assumes a pollutant reduction of 60%. This may be low for HDS, where data from Yee and McGee (2010) suggest reductions of 50% for PCB and 20% for mercury. Filtration-type BMPs generally require half as many of paired samples as sedimentation-type BMPs. If paired samples are not feasible, the number of data points for unpaired analysis is expected to be more than four times higher than the number required for paired analysis.

Instead of reliance on typical statistically-based study design, the primary research question is best addressed by collecting sufficient sample to develop an average estimate influent loading, effluent loading, and bypass loading. Where removal mechanisms are well-known and consistent (filters and bioretention), a sample size of 4 is selected for water samples. For other BMPs, sample size could be much higher, but schedule limits the maximum amount of targeted storms at 6. For bypass monitoring, at least 2 storms are selected and up to the number of storms monitored for water sampling are preferred.

Selected Constituents for Water Quality Analysis

Constituents for analysis of water samples are selected to obtain the data objectives discussed previously in "Study Design Options for Primary Data Objectives". The constituents and analysis methods should be consistent with Table 13-2 of the Quality Assurance Project Plan (QAPP) (BASMAA 2013).

Table 1 lists the possible constituents for two water quality monitoring scenarios (sediment is discussed in the following section). The first two columns indicate by a "Yes" or "No" which of the corresponding constituents have been selected for analysis for the respective monitoring scenario represented by the column headings. The laboratory analysis costs used in later cost optimization exercises are also displayed. The first column in

Table 1 is the constituent list for paired influent and effluent water sampling and it addresses the primary research question and both performance data objectives. The second column is the constituent list for treatment mechanism checks via effluent water sampling. Because this only addresses one secondary research objective – the corroboration of treatment mechanism performance – this approach is presented as a cost-savings option in case the monitoring budget is reduced.

Water quality constituents were first selected based on the primary research question to quantify PCB and mercury reduction, so PCBs and mercury are measured directly. Secondary research questions relate to understanding removal performance for PCB and mercury and discovering less costly surrogates for PCB and mercury.

In addition to PCB and mercury, the other constituents selected for influent and effluent analysis are TSS, SSC, volatile solids, settleable solids, total organic carbon, lead, particle size distribution, and wet sieving. Though not listed in the QAPP, TSS was selected for the sake of comparing performance to other studies that predominantly analyze for TSS. SSC was also selected because it more accurately characterizes larger size fractions within the water column.

SSC is more critical for influent monitoring; because larger particles are not expected in effluent monitoring, TSS should be adequate. Volatile solids was selected to indicate how much vegetal debris contributes to suspended solids compared to erosion of soil minerals. Though not listed in the QAPP, settleable solids was selected because it indicates the amount of solids that can be removed by traditional sedimentation BMPs without the use of filtration. The difference between settleable solids and SSC also can indicate the fraction of solids that may require filtration. The presence of substantial amounts of settleable solids in the effluent would indicate a design problem or flaw in the BMP. Turbidity was selected because it is an inexpensive and quick field test that can be used to describe treatment efficiency where strong correlation to other pollutants has been established. Turbidity is also measured at all influent sampling events since it can be used to reduce the cost of future site screening for retrofits. Total organic carbon was selected because it can affect the fate and transport of other pollutants such as mercury. Lead was selected because it is a legacy pollutant that may corroborate where PCBs are originating within each BMP drainage area. Sometimes lead can also be a time marker for the age of the eroding sediment since newer sediment sources have not been exposed to the effects of leaded gasoline.

To complement past studies on particle size correlations, three size fractions are suggested for analysis: less than 25 µm, 25 to 63 µm, and greater than 63 µm. PCB, mercury, and TOC are suggested to be analyzed for these three particle size ranges.

Table 1 contains analysis of PCB for four sediment size fractions: total, dissolved, < 25 microns, and < 63 microns. The remaining fraction greater than 63 microns is calculated by subtraction. In many cases low concentration of PCB in water samples negate the value of quantifying PCB for different size fractions. Similar to the approach in Task 4, a screening methodology will be implemented to identify samples that will have PCB and mercury analyses performed for two grain size fractions (less than and equal to, and greater than 63 microns [µm]), based on the analysis of the whole sediment sample. Data analyzed by Geosyntec show that the 61 percent of sediment samples had PCB concentrations < 100 µg/Kg and 54 percent of samples had mercury concentrations < 150 µg/Kg (Figure 3). To target samples with high concentrations relative to the existing data, PCB and mercury levels within these grain size fractions will only be be analyzed if the whole sample PCB concentration is ≥ 100 µg/Kg, or the mercury concentration is ≥ 150 µg/Kg, and the solids concentration exceeds 100 mg/L. Recommended initial thresholds (which may be adjusted during the studies) for water sample concentrations are ≥ 10 ng/L PCB, ≥ 15 ng/L mercury, and >20 mg/L SSC

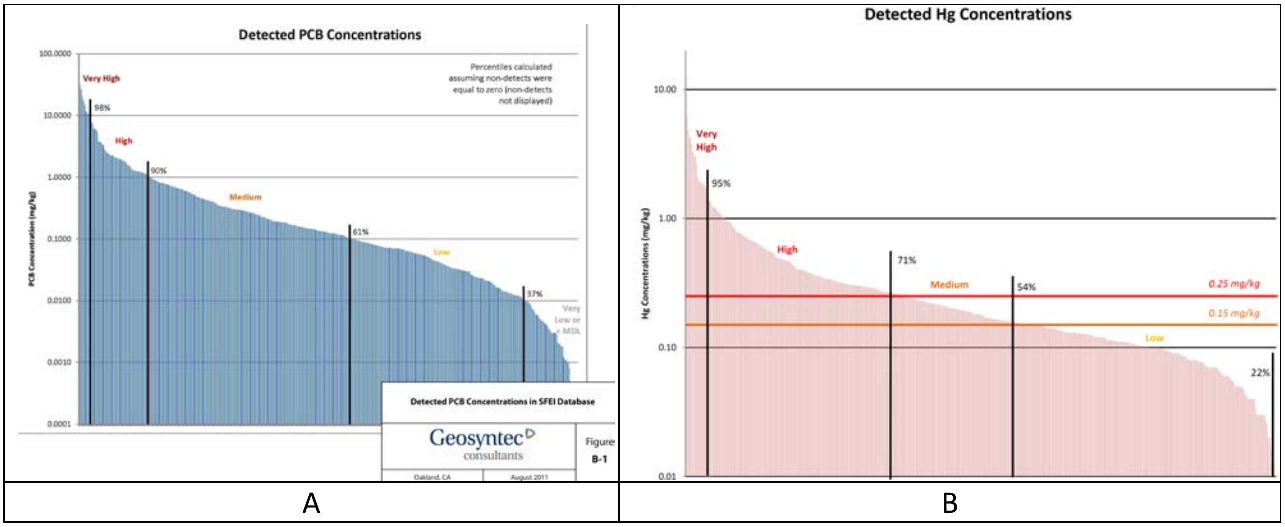


Figure 3. Ranked Sediment Concentrations for PCB (A) and Hg (B)

The quantity of water needed for the analysis recommended in this section may not be available from each sampling event. In that case, PCBs, and mercury should be the top priorities for analysis because to address the primary research question. Further prioritization is discussed in the "Cost and Schedule-Limited Study Plan" subsection.

Table 1. Costs of Selected Constituents for Water Monitoring Scenarios **

Influent and Effluent Paired Samples	Effluent Treatment Mechanism Check	Potential Water Analytes	
Yes	Yes	TSS	$25
Yes	No	SSC	$80
Yes	Yes	Volatile Solids	$50
Yes	Yes	Settleable Solids	$15
Yes	Yes	Turbidity	$8
Yes	No	Total Organic Carbon	$45
Yes	No	Lead	$25
Yes	No	Mercury (Method 1631)	$125
Yes	Yes	PCBs	$725
Yes	No	PCB Dissolved	$725
Yes	Yes	Particle Size Distribution	$75
yes	Yes	25 micron Wet Sieve and Analysis of PCB and mercury*	$900
Yes	Yes	63 micron Wet Sieve and Analysis of PCB and mercury*	$900
Total cost per sample, as selected:			
$3,698	**$2,698**		

* analyze fractions only if PCB concentration is ≥ 10 ng/L, or the mercury concentration is ≥ 15 ng/L nd solids concentration exceeds 20 mg/L.

Selected Constituents for Sediment Analysis

Constituents selected for sediment analysis are based on the type of BMP being monitored and whether media within the BMP could interfere with the accumulated sediment and pollutant loading. Table 2 lists the possible constituents for two sediment quality monitoring scenarios. The first two columns indicate by a “Yes” or “No” the corresponding constituents listed in the third column that have been selected for analysis under the respective sediment quality monitoring scenario represented by the column headings. The laboratory analysis costs are presented in the final column.

The primary objective of sediment analysis is quantification of the mass of PCBs and mercury accumulating within the BMP. Consequently, PCBs and mercury should be analyzed for all sediment analysis.

The secondary objective is to establish a relationship between PCB, mercury, and particle size. Filters, bioretention, and tree wells are not good candidates to explore this relationship because the sediment that exists in the planting and filtration media will be difficult to differentiate from influent sediment. In contrast, if an HDS begins each season with an empty sump, then the solids accumulated within the sump are presumably those that entered in the influent stream. Chemical analysis of the solids collected in the sump will help establish the relationship between PCB, mercury, and coarser solids that are removed and retained in a fairly high-energy environment. This limits the application of the relationship outside of HDS units or other high-energy environments. As with water quality analysis, correlating PCB and mercury to particle sizes will complement past studies and provide insight into the type of retrofit BMPs that are appropriate to achieve the most cost-effective mass removal. Three size fractions are suggested for analysis: less than 25 µm, 25 to 63 µm, and greater than 63 µm. PCB, mercury, and TOC are suggested to be analyzed for these three particle size ranges.

Table 2. Costs of Selected Constituents for Sediment Monitoring Scenarios

Full Sediment Analysis (sumps)	Potential Sediment Analytes	
Yes	Lead	$25
Yes	Mercury	$50
Yes	PCBs	$725
Yes	Grain Size	$75
Yes	Bulk Density	$30
Yes	25 micron Wet Sieve and Duplicate Analysis of PCB, and Hg*	$825
Yes	63 micron Wet Sieve and Duplicate Analysis of PCB, and Hg*	$825
Total cost per sample, as selected: $2,555		

* analyze fractions only if PCB concentration is ≥ 100 µg/Kg, or the mercury concentration is ≥ 150 µg/Kg

Budget and Schedule

The monitoring budget for Task 5 of the CW4CB project is approximately $1 million. A contingency of 20 percent of the water quality monitoring budget is recommended to account for false starts on storm events that do not materialize as expected or for missed storms due to equipment failure. Projected cost of the optimized study plan are provided and compared to the budget in the next section.

Another constraint is that all sampling will occur in one wet season. To increase the probability of successfully meeting the target number of storms, no more than six storms will be targeted for any one location.

Cost and Schedule-Limited Study Plan

The cost and schedule-limited study plan is presented in Table 3. Several iterations were analyzed and what is presented in Table 3 is just one option that uses best professional judgment to allocate the budget to the various data collection options for BMPs that will be available for most of the 2013/14 wet season. A guiding principle from the TAC was to increase quality at the expense of replication and to focus on data gaps.

Delays in the construction of some BMP pilots also will affect the cost and schedule for sampling. Unfortunately construction delays at Nevin, and 1st and Cutting has eliminated consideration of paired sampling at these locations for the 2013/14 wet season. A late completion date for the swale at Broadway and Redwood would only allow a couple months of wet weather for sampling. Since the swale requires additional time for vegetation establishment, the swale will not be monitored as part of this study. The PG&E substation in Vallejo is also delayed, but the canister filter installed here does not require as many samples as a swale so that this retrofit is still scheduled for monitoring. Assuming the grant can be extended, funding for these systems is included in the 2014/15 wet season. For Nevin, the funding is only programmed for observation of the hydraulic performance of the porous pavement system since that system exclusive to the Nevin project. If the system is constructed with bypass, overflow, or underdrain features, the visual monitoring could be upgraded to metered monitoring (from monitoring scenario F to C in Table 3) if budget allows. Another budget shift to consider is from Cerritos to the Silva Cell systems at Nevin and the swale at Redwood. But because of the Cerritos has established vegetation and seasoned media, it represents a more sure monitoring location and so it remains in this plan.

If an extension of grant funds is not feasible, the monitoring savings from delayed projects can be allocated for additional 2013/14 monitoring of the BMPs in West Oakland and Bransten Rd. in San Carlos. Monitoring an addition tree well and bioretention without an underdrain would allow a better understanding of site-specific water quality performance and hydraulic performance (i.e., amount of bypass).

For the projects selected for monitoring in Table 3, the recommended sampling method by BMP type is as follows:

- Sediment sampling for mass removal[4]
    - HDS

[4] Based on preliminary results from screening sampling, sediment sampling at the HDS unit at Leo Avenue may also help identify non-stormwater contaminant issues.

- Water quality sampling among BMP types, where construction schedule allows, and focus on lesser-studied BMPs:
    - Pump station filters
    - Canister filters
    - Unique bioretention media (e.g., biochar)
    - HDS
    - Typical bioretention media
    - Inlets to pilots that have not completed construction (e.g., porous pavement); to be used in future mass removal estimates

For water quality analysis, sample volume is not always sufficient to complete all of the analysis. The primary research question and secondary data gaps are considered in setting the following analyte priority for each sample location:

- Effluent samples
    - PCB and mercury
    - particle size distribution
    - PCB within particle size fractions
    - All other constituents (Pb, VS, TOC)
- Influent (untreated)
    - total PCB, SSC, and turbidity
    - PCB within particle size fractions (only analyze if whole sample PCB concentration is greater than 10 ng/L and solids is greater than 10 mg/L, as discussed in the “Selected Constituents for Water Quality Analysis” subsection)
    - Hg
    - All other constituents (Pb, VSS, TOC)

The contingency to cover false starts is 20 percent of the water quality monitoring subcontracted costs (does not include analysis).

The results of the prioritization strategy will yield approximately 2 sediment samples and over 80 water samples. The BMPs, monitoring location types, monitoring scenarios, and approximate costs are presented in Table 3. Note that some monitoring locations types in Table 3 are not designated for monitoring as indicated by a “0” for the number of stations selected. These locations types remain in the table in case further adjustments are needed (e.g., if cost cannot be shifted to the 2014/15 season). Storm tracking, precipitation monitoring, laboratory coordination, and QA/QC review and reporting were included in the cost estimates, though not listed as line items in Table 3. Table 4 presents sediment and water quality data collected for each BMP type.

Table 3. Cost-Optimized Study Plan and Cost with Selected Monitoring Effort for Each Location

Stormwater Program	Project Number	Project Name	Number of Each Treatment Type – LID: Bioretention w/ underdrain	LID: Bioretention w/out underdrain	LID: Permeable Pavement/ Subsurface Storage	LID: Flow through Biotreatment (Swale)	LID: Silva Cell	LID: Tree Well	Other: Catch Basin Media Filter	Other: Sand Filter	Other: Hydrodynamic Separators	Date Available for Monitoring (month/year)***	Monitoring Location Type	Number of Stations Selected per Location	Monitoring Scenario*	Analytical Scenario**	Total Samples per Station (Water, Sediment, or Flow)***	Typical Setup Cost	Approximate Analysis Cost - All Samples	Approximate Labor - All Events	Approximate Cost per Monitoring Location	2013/14 Wet Season Cost***	2014/15 Wet Season Cost***	Approximate Cost per Project
ACWP	1	Ettie St. Pump Station								2		12/13	Post-settling tank influent	1	B	B	5	$ 10,000	$ 18,490	$ 20,000	$ 48,490	$ 149,225		$ 149,225
													Effluent	2	B	B	5	$ 20,000	$ 36,980	$ 40,000	$ 96,980			
													Filter sediments	0	A	D	0	$ -	$ -	$ -	$ -			
													Settling tank	1	A	A	1	$ 200	$ 2,555	$ 1,000	$ 3,755			
	2	Alameda and High St. HDS									1	12/12	HDS Unit Sump	0	A	A	0	$ -	$ -	$ -	$ -	$ -		$ -
													HDS Unit Influent	0	D	B	0	$ -	$ -	$ -	$ -			
													HDS Unit Effluent	0	D	B	0	$ -	$ -	$ -	$ -			
	3	West Oakland Industrial Area						6				End of 10/13	(blank)	0	A	A	0	$ -	$ -	$ -	$ -	$ 160,768		$ 160,768
													Bypass Flow Monitoring	2	C	F	0	$ 16,000	$ -	$ -	$ 16,000			
													Influent monitoring	2	D	B	4	$ 6,000	$ 29,584	$ 36,800	$ 72,384			
													Effluent monitoring	2	D	B	4	$ 6,000	$ 29,584	$ 36,800	$ 72,384			
													Planting media sediments	0	A	D	1	$ -	$ -	$ -	$ -			
CCCWP	4	Nevin Avenue Improvements (Green Streets)	10		1		2	0				7/14	Multiple porous pavement and underground storage influent	1	F	F	6	$ 200	$ 1,200	$ 6,000	$ 7,400		$ 28,800	$ 28,800
													Porous Pavement Bypass, Underdrain, or Overflow	1	F	F	6	$ 200	$ 1,200	$ 6,000	$ 7,400			
													Influent for Silva Cell bioretention	0	D	B	0	$ -	$ -	$ -	$ -			
													Silva Cell Bypass, Underdrain, or Overflow	1	C	F	4	$ 8,000	$ 800	$ 5,200	$ 14,000			

(Table 3 *continued*)

Stormwater Program	Project Number	Project Name	Number of Each Treatment Type: LID: Bioretention w/ underdrain	LID: Bioretention w/out underdrain	LID: Permeable Pavement/ Subsurface Storage	LID: Flow through Biotreatment (Swale)	LID: Silva Cell	LID: Tree Well	Other: Catch Basin Media Filter	Other: Sand Filter	Other: Hydrodynamic Separators	Date Available for Monitoring (month/year)***	Monitoring Location Type	Number of Stations Selected per Location	Monitoring Scenario*	Analytical Scenario**	Total Samples per Station (Water, Sediment, or Flow)***	Typical Setup Cost	Approximate Analysis Cost - All Samples	Approximate Labor - All Events	Approximate Cost per Monitoring Location	2013/14 Wet Season Cost***	2014/15 Wet Season Cost***	Approximate Cost per Project
CCCWP (*continued*)	5	PG&E Substation 1st and Cutting	1	1								7/14	Influent Cutting, Media A	1	D	B	4	$ 3,000	$ 14,792	$ 18,400	$ 36,192		$ 172,768	$ 172,768
													Influent Cutting, Media B	1	D	B	4	$ 3,000	$ 14,792	$ 18,400	$ 36,192			
													Effluent	2	D	B	4	$ 6,000	$ 29,584	$ 36,800	$ 72,384			
													Bypass Flow Monitoring	2	C	F	4	$ 16,000	$ 1,600	$ 10,400	$ 28,000			
													Filter sediments	0	A	D	0	$ -	$ -	$ -	$ -			
	6	El Cerrito Green Street	2									previously constructed	Influent	1	D	E	4	$ 3,000	$ 3,812	$ 18,400	$ 25,212	$ 71,404		$ 71,404
													Effluent	1	D	C	4	$ 3,000	$ 10,792	$ 18,400	$ 32,192			
													Bypass Flow Monitoring (may bypass at curb cut)	1	C	F	4	$ 8,000	$ 800	$ 5,200	$ 14,000			
													Filter core composite sediments	0	A	D	2	$ -	$ -	$ -	$ -			
SCVU-RPPP	7	Leo Avenue HDS									1	10/12	HDS Unit Sump	1	A	A	1	$ 200	$ 2,555	$ 1,000	$ 3,755	$ 76,139		$ 76,139
													HDS Influent	1	D	B	4	$ 3,000	$ 14,792	$ 18,400	$ 36,192			
													HDS Effluent	1	D	B	4	$ 3,000	$ 14,792	$ 18,400	$ 36,192			
SMC-WPPP	8	Bransten Road Green Street	4	6								End of 10/13	Effluent from underdrain	1	D	B	4	$ 3,000	$ 14,792	$ 18,400	$ 36,192	$ 136,576		$ 136,576
													Influent with underdrain	1	D	B	4	$ 3,000	$ 14,792	$ 18,400	$ 36,192			
													Influent without underdrain	1	D	B	4	$ 3,000	$ 14,792	$ 18,400	$ 36,192			
													Bypass Flow Monitoring	2	C	F	4	$ 16,000	$ 1,600	$ 10,400	$ 28,000			
													Observation well (system with underdrain)	0	C	F	0	$ -	$ -	$ -	$ -			

(Table 3 *continued*)

Stormwater Program	Project Number	Project Name	Number of Each Treatment Type									Date Available for Monitoring in 2013/14 Wet Season	Monitoring Location Type	Number of Stations Selected per Location	Monitoring Scenario*	Analytical Scenario**	Total Samples per Station (Water, Sediment, or Flow)***	Typical Setup Cost	Approximate Analysis Cost - All Samples	Approximate Labor - All Events	Approximate Cost per Monitoring Location	2013/14 Wet Season Cost***	2014/15 Wet Season Cost***	Approximate Cost per Project
			LID						Other															
			Bioretention w/ underdrain	Bioretention w/out underdrain	Permeable Pavement/ Subsurface Storage	Flow through Biotreatment (Swale)	Silva Cell	Tree Well	Catch Basin Media Filter	Sand Filter	Hydrodynamic Separators													
FSURMP/VSFCD	9	Broadway and Redwood				1						Mar-14	Multi-influent locations	0	D	B	6	$ -	$ -	$ -	$ -			
													Effluent	1	C	F	6	$ 8,000	$ 1,200	$ 7,800	$ 17,000		$ 17,000	$ 17,000
													Observation well or infiltrometer	0	C	F	1	$ -	$ -	$ -	$ -			
													(placeholder)	0	A	D	0	$ -	$ -	$ -	$ -			
	10	PG&E Substation							1			Mar-14	Influent	1	B	B	3	$ 10,000	$ 11,094	$ 12,000	$ 33,094			
													Effluent	1	B	B	3	$ 10,000	$ 11,094	$ 12,000	$ 33,094	$ 78,688		$ 78,688
													Bypass flow over weir	1	C	F	3	$ 8,000	$ 600	$ 3,900	$ 12,500			
													(placeholder)	0	A	D	0	$ -	$ -	$ -	$ -			
		Total	17	7	1	1	2	6	1	2	2							$ 175,800	$ 294,668	$ 416,900	$ 887,368	$ 668,800	$ 218,568	$ 887,368

* Monitoring Scenario Description:
A. Grab sample of sediment
B. Automatic flow-weighted composite water samples
C. Automatic flow-only monitoring (at least throughout sampled events)
D. Time-weight grab, auto flow
E. Grab water quality and visual flow read or estimation
F. Visual flow monitoring

** Analytical Scenario Description:
A. Full Sediment
B. Full Water Quality
C. Treatment mechanism check: TSS, PSD, VSS
D. Filter or planting media
E. Drainage Loading Characterization and PCB/solids
F. No laboratory analysis: Reporting Only

***For the 2013/14 wet season, consider deferring 1 targeted storm event to the 2014/15 wet season. The purpose of this deferment is to monitor the early-season first flush effects in the 2014/15 wet season that were missed due to construction delays in the 2013/14 wet season.

20% contingency on mobilization for false-starts	$ 83,380
Total Project Cost Estimate	**$ 970,748**
Budget	**$ 1,000,000**
Total Sediment Samples	2
Total Water Samples	81

Adequacy of Study Plan to Answer the Primary Research Question

The primary research question is reviewed in this section in light of the budgeted data collection efforts. The primary research question is restated and followed by an analysis of the adequacy of the data collection effort. Table 4 shows the types of data collected for each BMP type.

1. *What is the catchment efficiency of retrofits*?

Table 4 displays the number of data points that are anticipated for the retrofits. Bioretention and sand filters are monitored at multiple installations because each BMP type have been designed with different media. The sampling effort is sufficient to support catchment efficiency calculation for the following retrofits:

- Bioretention with underdrain
- Bioretention without underdrain
- Tree Well
- Catch basin media filter
- HDS

Due to construction schedule issues, the following systems will require catchment efficiency calculation based on the PCB/SSC correlation data collected at other locations, solids performance data from other studies (e.g., International BMP database), and estimated hydraulic performance:

- Permeable pavement and subsurface storage/retention
- Silva Cells
- Bioswale

All three of the above systems are supported by flow monitoring observations in the 2014/15 wet season as shown in Table 3, which will assist in estimates of pollutant removal.

Table 4. Data Collection by BMP Type

	Bioretention w/ underdrain	Bioretention w/out underdrain	Permeable Pavement/ Subsurface Storage	Flow through Biotreatment (Swale)	Silva Cell	Tree Well	Catch Basin Media Filter	Sand Filter	Hydrodynamic Separators	Total
No. of Installations	17	7	1	1	2	6	1	2	2	39
No. Monitored for Water Quality	3	2	0	0	0	2	1	2	1	11
No. Monitored for Flow Only	0	0	1	1	1	0	0	0	0	3
No. of Sediment Samples	0	0	0	0	0	0	0	0	1	1
No. of Water Quality Influent Samples	8	4	0	0	0	8	3	5	4	32
No. of Water Quality Effluent Samples	8	4	0	0	0	8	3	10	4	37
No. of Paired Influent/Effluent Samples	8	4	0	0	0	8	3	10	4	37

Reducing Analytes and Monitoring Alternative Locations

As the study progresses and data are collected, certain constituents can be dropped from further analysis at particular locations. There are two primary criteria for dropping the constituent. First, the constituent is consistently measured at a level so low that the data for the site would not be representative of for future candidate sites for retrofit. Second, there are no alternative locations of that BMP type that would be likely to have higher influent concentrations. In that case, the project management team can consider dropping the constituent or terminating monitoring for that BMP. As a general guideline, influent suspended solids concentrations of less than 20 mg/L are fairly low and approach the level at which many BMPs would not be able to substantially reduce them.

Another reason to drop a constituent would be in cases where the desired research question has been sufficiently addressed. For example, if turbidity is shown to have a strong correlation to another constituent that is much more costly to analyze, that constituent could be discontinued from further analysis. This is not expected because of the limited number of samples that can be collected within the term of this study.

Insufficient sample volume will also require a reduction in analytes. In that case, refer to the prioritization of constituents presented previously in the "Cost and Schedule-Limited Study Plan" subsection.

Data Analysis

Analysis of the generated data should result in ranges of costs for BMP types and ranges of pollutant removal for both BMP types and influent conditions. This information can be used to compare BMP alternatives for future retrofit locations.

Catchment Efficiency

See discussion in Study Goals section.

Cost-Benefit Ratios

Cost-benefit ratios should be calculated by dividing the life-cycle costs by the mass of pollutant removed over the life-cycle of the BMP. See "Recommendations for Cost Tracking and Analysis".

BMP Comparisons

Statistical hypothesis testing can be used to compare measured BMP performance to expected performance as documented by other studies. The BMPs must have influent and effluent water quality data. A statistical comparison of slope and intercept or a statistical comparison of effluent quality at a common influent concentration may be useful in discerning differences in treatment effectiveness. The International BMP Database is a helpful source of performance data.

The cost-benefit ratios of all BMP types can also be compared. However, due to the limited number of duplicate BMP types, statistical comparisons of cost-benefit ratios will not be possible among BMPs with fewer than 3 monitored installations. Further, such statistically based cost-benefit comparisons will likely be inconclusive due to extremely small sample size. Instead, a range of cost-benefit values should be developed from the corroborated performance data and the results of the cost analysis.

Applicability to Future Retrofits

Some organization of the results will help estimate cost-benefit ratios to plan future retrofits. One option is to group the mass removal estimates by ranges of influent quality measurements. The data should be examined for grouping of mass removal levels for various combinations of BMP type and influent quality. A theoretical example of how the data might be organized is shown in Table 5. The range of pollutant loading should be based on observation of where meaningful differences occur. Because it is impossible to predict exact influent loading, it should be expected that such a table will have missing values. The range of influent conditions can be broadened to minimize these missing values. This will make the resulting tables easier to use for predicting the cost-benefit ratio of future retrofits. For each influent category, the ranges of mass removal may also be useful. This will allow the proponent of future retrofits to estimate a range of cost-benefit ratios for comparisons with other alternatives.

Table 5. Possible Organization of Annual Mass Removal Results by Site Conditions, for Use as a BMP Selection Tool for a Candidate Retrofit Sites

Typical TSS or SSC concentration, mg/L	PCB hot spot range, mg/kg	BMP A	BMP B	BMP C	BMP D		BMP E	BMP F	BMP G
					Alt. Design 1	Alt. Design 2			
< 20	May not be appropriate data for future retrofit scenarios								
20 to 63*	0.10 to 0.5*								
	> 0.5*								
>63*	0.10 to 0.5*								
	> 0.5*								

*If there is a noticeable change in performance within the range of influent loading values, the value dividing high and medium loading conditions should be selected where the transition in performance occurs. This level of analysis is not likely possible with the limited data collected.

Another analysis approach is to also present cumulative frequency distributions curves of inlet and outlet concentrations. This approach will also be limited by the lack of data, which will result in sparse curves.

Recommendations for Sampling and Analysis Plans

This section presents specific recommendations for the development of SAPs, focusing on flow measurement device selection and sizing, storm mobilization criteria, data quality objectives, and event-specific reporting. These recommendations are followed by suggestions for development of field-level SAPs. More detailed information is available in Section 6 of the *Caltrans Guidance Manual: Stormwater Quality Monitoring Protocols* (2003) and in the *Urban Stormwater BMP Performance Monitoring* (WERF 2009). The specifics of SAPs cannot be developed until the BMP plans are fully developed. Revisions are to be expected once the BMPs are constructed and again after monitoring crews gain some sampling experience with the systems. Analysis of constituents should follow the CW4CB Quality Assurance Project Plan (BASMAA 2013).

Flow Measurement Devices

For flow-weighted composite sampling, flumes, weirs, and tipping buckets are used for the collection of samples that are representative of the entire influent and effluent. In this study, the subset of pilots that was to be monitored was unknown during the design phase, so it was not feasible to design proper space for flumes. Further, it was recognized that previous monitoring at the El Cerrito pilot successfully used time-weighted sampling techniques. Even so, the use of flow measurement is critical to volume estimation.

Traditionally, flumes are sized for peak hydraulic capacity, including bypass flows. This approach results in a flume that is drastically oversized for the more frequent storm events that are monitored for BMP effectiveness. Instead, flumes and weirs should be optimized for accurate measurement of small, more frequently occurring storm events.

Travel time, attenuation, and bypass should also be considered in the selection of flumes sizes. Influent flumes or weirs for longer watersheds, and longer times of concentration (travel time) can be designed based on smaller storm intensities compared to flumes and weirs for shorter watersheds. In many cases in California, influent flows may be as low as those resulting from a storm intensity of 0.2 in/hr. Less frequent flows that exceed the capacity of the flume can be estimated by more crude methods such as pipe flow. For effluent monitoring, the BMP may impede flow or divert higher flows around the primary treatment mechanism. In such cases, flume or weir selection should take advantage of the smaller flows by selecting smaller flumes or weirs compared to the influent. Using this strategy will result in collection of more accurate data for the range of flows expected in the effluent. Researchers at the University of New Hampshire have had success using V-notch weirs for space-constrained locations (WERF 2009). Ongoing bioretention evaluation in Contra Costa County has successfully used tipping bucket rainfall measurement devices, but the space constraints may limit the use of this tool.

Proper installation of automatic flow measurement equipment is critical to the accuracy of the device. Manufacturers' recommendations must be followed, particularly in regards to the unobstructed upstream and downstream distances (in terms of the number of pipe diameters), backwater conditions, slope, and level. Flumes and weirs cast in concrete should be avoided since their position cannot be fine-tuned.

Calibration of the flow or weir is equally important to proper installation. Calibration indicates when adjustments are necessary so that the stage-discharge curves provided by the manufacturers can be used by the automatic samplers. Ideally, a metered water truck offers the best potential to calibrate flumes installed in fairly small watersheds. For extremely small watersheds, a bucket and stopwatch may be sufficient to verify the measurement of lower flows during a storm event. *Standard Operating Procedures for Stormwater Flow Measurement Verification* (Caltrans 2010b) covers several flume and weir scenarios. Calibration methods are also discussed in Section 3 of *Urban Stormwater BMP Performance Monitoring* (WERF 2009).

The inspection and maintenance plan for flumes and weirs should ensure the continual accuracy of flow monitoring equipment.

More information on types of flumes can be found in the *Caltrans Guidance Manual: Stormwater Quality Monitoring Protocols* (2003) and in the *Open Channel Flow Measurement Handbook* (ISCO 2006).

Storm Event Mobilization Criteria

Storm event mobilization criteria for water quality sampling often change throughout a study period in consideration of storm size, probability, antecedent dry period, number of successfully monitored storms, and the expected number of storms for the remaining wet season. The mobilization criteria target storms that best address the objectives of the study. The primary performance data objective is mass of pollutant removed over the life-cycle of the BMP. To compare BMPs, those with water quality monitoring should develop data that can be used to estimate average annual load reductions (e.g., that consider all storm sizes). Though it is ideal that every storm size be monitored, it is fiscally infeasible for most studies. Instead, storms that would represent typical storms seen within the period of interest are targeted. Where only a few storms are monitored throughout the study, it is important to monitor a range of storm sizes so that annual load reductions can be modeled. However, a constraint to targeting storm sizes is that smaller storms are less reliable, and mobilizing to capture small storms will increase the chance of false starts.

Where more storms are monitored during a relatively brief study period, monitoring crews often are not able to be selective in the size of storm monitored because of the limited number of storms throughout the study period. A typical minimum quantified precipitation forecast (QPF) for stormwater monitoring is 0.25 to 0.50 inches of precipitation. Because of the single season of monitoring, crews should more aggressively pursue storms, but this comes at a risk of increased false starts. It is common practice to relax mobilization criteria only if the number of captured storms in the beginning of the wet season is less than expected. To further reduce cost and false starts, crews should consider targeting compact storms where precipitation is concentrated within a fraction of a day instead of drawn out over many days. When monitoring a portion of a storm, the residence time of water within the BMPs can cause an unclear division between storm runoff events. This issue should be considered if a storm cell that occurs adjacent to another cell is targeted. Generally, separation of a few hours may be adequate for the BMPs used in this study, due to short residence times. Effluent rates can be observed to check this assumption. A marked increase in rate after a decrease from the initial peak may be an indication to cease sampling.

Many monitoring protocols suggest minimums for antecedent dry periods between storms. This is usually done for two reasons. The first reason is to focus the monitoring on more contaminated stormwater runoff. This can result in a bias of the data set, which will then be used to model all future storm runoff events. The second reason is more practical; it seeks to prevent the comingling of water from two separate events. A minimum antecedent dry period allows a BMP that retains water for prolonged periods to fully drain before the next storm event. Using drain time as the minimum antecedent dry period gives monitoring crews the fullest flexibility in capturing the desired number of storm events.

Unlike water samples, flow monitoring should be continuous, if at all possible. This is usually only possible with automatic flow monitoring equipment.

Data Quality Objectives

Data quality objectives (DQOs) should follow standard stormwater monitoring protocols and be described in detail in individual SAPs. Both sampling and laboratory data quality objectives should be included. For sampling, the SAP should specify the minimum number of aliquots and percent coverage of the event to assure representative sampling. For laboratories, numeric DQOs are appropriate for sample blanks, duplicates (or field splits), and matrix spike recovery.

The quality of the flow data is also critical to establishing that the analyzed samples represent the concentration and the volume of water that was monitored. Systematic error is not as critical for flow-weighted composite sampling as it is for estimating total volume, because sample pacing is relative to the volume measured, and the correct flow weighting will occur whether or not the flows are systematically high or low. This is not the case for estimating total water volume. To assist in accurate total volume measurements (fractions treated, retained, and bypassed), verify drainage areas and flow paths during storm events.

Event-Specific Reporting

It is highly recommended that field monitoring forms be prepared for pre- and post-storm site visits. These forms should be developed and included in the sampling and analysis plans.

Data collected via automated samplers should be run through a utility to develop a standard hydrograph for each monitored event. This is a crucial step to enable review and validation of representative sampling techniques. The hydrograph should show when all aliquots were triggered, where successful aliquots were taken for each composite sample, and where any triggered samples were missed.

Field Guide Sampling and Analysis

The overall SAP is being prepared concurrent to this document. Much of the SAP is not relevant to field activities. Instead, field-level SAPs that use checklists can be more helpful for field crews to implement the overall SAP. The following should be considered when developing field-level SAPs:

- Description of Site
    - Site Locations and Access
    - BMP Features

  - Monitoring Locations
- Monitoring
  - Analytical Constituents
  - Monitoring Equipment Checklist
  - Monitoring Preparation and Logistics
    - Weather Tracking Coordinator Contact Information
    - Storm Mobilization Criteria (keep updated)
    - Continuous Flow Monitoring Setup Procedures
    - Communication/Notification Procedures
    - Monitoring and Equipment Preparation
  -
  - Sample Collection, Preservation, and Delivery
  - Chain of Custody Forms
  - Photo Log
  - Visual Observation Forms
- Health and Safety (including map with nearest ER)

Recommendations for Cost Tracking and Analysis

Cost tracking is a critical component for comparing the cost-effectiveness of retrofit types. Life-cycle cost is necessary to compare the treatment retrofits (Task 5) with the municipal operations that are enhanced for sediment removal (Task 4). Capital costs must be collected from the contracting agencies. These costs must be carefully reviewed to exclude any costs that are not typical of future retrofit projects. For projects with more than one BMP type, it is also critical to track cost by individual BMP type. Appendix C of the *Caltrans BMP Retrofit Pilot Program Final Report* (2004) contains useful categories to help identify costs that are unique to pilot programs and ancillary costs that are not critical to the installation of the BMP.

Operational costs should be collected by the maintaining consultant or agency. For maintenance and operational costs, considerable effort is needed to separate operational costs from other costs, regardless of the operating entity. The challenge for municipal maintenance crews is that the specific tracking of labor to specific sites and activities is not commonplace. For example, crews may be in an area picking up trash and debris from channels and BMP maintenance may only account for a small portion of the work done in that area. In many municipalities, all the hours would be captured in a general category. The challenge with using research staff to operate and maintain BMPs is that this effort often coincides with monitoring and inspection efforts. An operation and maintenance guide will help track these costs, regardless of the entities maintaining the BMPs. Costs should be reported to BASMAA.

References

BASMAA. 2009. *Clean Watersheds for a Clean Bay* BASMAA 2009 SF WQIF Proposal (v.11-09).

BASMAA. 2012. Clean Watersheds for a Clean Bay (CW4CB) Semi-Annual Progress Report Number Three. February 13, 2012.

BASMAA. 2013. Quality Assurance Project Plan (QAPP). Clean Watersheds for a Clean Bay – Implementing the San Francisco Bay's PCBs and Mercury TMDLs with a Focus on Urban Runoff. August 15, 2013.

Caltrans. 2003. Comprehensive Monitoring Protocols Guidance Manual: Stormwater Quality Monitoring Protocols, Particle/Sediment Monitoring Protocols, Gross Solids Monitoring Protocols, Toxicity Monitoring Protocols, and Caltrans Data Reporting Protocols. CTSW-RT-03-105.51.42. Sacramento: Caltrans.

Caltrans. 2004. *BMP Retrofit Pilot Program*, Final Report. CTSW-RT-01-050. Sacramento: Caltrans.

Caltrans. 2009. *BMP Pilot Study Guidance Manual.* CTSW-RT-06-171.02.1. Sacramento: Caltrans.

Caltrans. 2010a. *Performance-Based Ranking of Treatment BMPs*. CTSW-RT-10-239.01. Sacramento: Caltrans.

Caltrans. 2010b. *Standard Operating Procedures for Stormwater Flow Measurement Verification*. CTSW-RT-10-211.19.01. Sacramento: Caltrans.

ISCO. 2006. *Open Channel Flow Measurement Handbook*. Diane K. Walkowiak (Ed.), John A. Replogle (Foreword). Lincoln, NE: Teledyne ISCO, Inc.

Li, H. and A. Davis. 2008. Heavy Metal Capture and Accumulation in Bioretention Media. *Environmental Science and Technology*, 42: 5247-5253.

Mangarella, P., C. Sommers, K. Havens, and M. Leisenring. 2012. Integrated Monitoring Plan: Data Needs and Guidance. Draft Memorandum. To BASMAA Monitoring and POC Committee. 13 July 2012.

Mangarella, P. Personal communication. 21 August 2012.

McKee, L., P. Mangarella, B. Williamson, J. Hayworth, and L. Austin. 2006. Review of methods used to reduce urban stormwater loads: Task 3.4. A Technical Report of the Regional Watershed Program: SFEI Contribution #429. San Francisco Estuary Institute, Oakland, CA.

Murphy, K., B. Myors, and A. Wolach. 2009. *Statistical Power Analysis*, 3rd ed. New York: Routledge.

San Francisco Bay Regional Water Quality Control Board. 2011. Bay Area Municipal Regional Permit Order No. R2-2011-0083. Amended November 28, 2011.

San Francisco Bay Regional Water Quality Control Board. 2012. Total Maximum Daily Loads (TMDLs) and the 303(d) List of Impaired Water Bodies. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/TMDLs/ Last accessed August 6, 2012.

Yee, D., and L. J. McKee. 2010. Task 3.5: Concentrations of PCBs and Hg in soils, sediments and water in the urbanized Bay Area: Implications for best management. A technical report of the Watershed Program. SFEI Contribution 608. San Francisco Estuary Institute, Oakland CA.

Water Environment Research Foundation (WERF). 2009. Urban Stormwater BMP Performance Monitoring. Prepared by Geosyntec Consultants and Wright Water Engineers, Inc. October 2009. http://www.bmpdatabase.org/MonitoringEval.htm Last accessed August 6, 2012.

Appendix A

Statistically Derived Sample Size

The performance level and variability is carefully selected at meaningful levels. Table A-1 explores some reasons for selection of these two standards statistically-based study design.

Table A-1. Dual Standards for Statistically-Based Study Design

Standards	Reason for Standard	Example Criteria
Performance Result	This standard is selected at a level that will have meaningful impact from the perspective of the discharger. Performance can be adjusted upward if justified by other studies, reducing sample size. Failure of this standard is demonstrated statistically.	• Cumulative mass removed • Removal efficiency • Effluent quality
Performance Variability	This standard is selected at a level based on other similar studies of competitive alternatives so that less reliable BMPs studied here must compensate with a higher level of performance to conclude statistically significant treatment. Failure of the statistical test, when a performance level appears adequate, indicates to the discharger that unexplained variability may negatively affect performance at future retrofit locations.	• Coefficient of variation

Traditionally, answers to statistical questions are generally expressed as whether differences exist between using a BMP and not using a BMP. Where differences exist (any amount of treatment by the BMP), the answers to secondary statistical questions are generally expressed as whether differences exist among BMPs. Proper statistical design and investment of effort will ensure useful answers, even when statistical significance is not concluded. For example, it would be a useful conclusion, given proper sample size, if one or even all retrofits did not show sufficient performance to conclude statistically significant treatment. That is because it is during the study plan development that a minimum level of treatment and variability is defined and used to determine sample size. This is helpful if the level of treatment cannot be estimated from other studies or when those estimates show that removal is much lower than what the project proponents would consider the minimum acceptable. The values chosen for the variability in performance criteria can be estimated from past studies, but using sediment as a surrogate for PCB will underestimate sample size due to additional variability.[5]

[5] The selected performance level in power analysis is often based on performance of similar systems. Where these data are not available, the study plan should pick a minimum level of performance that would make the BMP somewhat competitive with the cost-benefit ratios of other BMPs.

Table A-2. Sediment-based Power Analysis for Determining Minimum Sample Size for Treatment Effectiveness of Individual BMPs

Estimated Bay Area Influent Concentration	79.1 mg/L, from NSQD [a]			
Coefficient of Variation for TSS from NSQD [a]	1.6		Z80	0.84
Selected Delta, Percent Removal of Load [b]	60%		Z95	1.65

Unpaired Sample Size Calculations to Test Statistically Significant Removal [b]

	ISBMPD [c] Mean		ISBMPD Difference in Means (Potential Delta)	For an Influent of 79.1mg/L, Regression[d] Estimate for Effluent	Sample Count to Estimate Composite Variance [e]		Standard Deviation, s [f]		Common Standard Deviation for Influent and Effluent s_p	Difference in Regression Effluent and Influent of 79.1mg/L	**Total Unpaired Samples Needed in Each Group for Expected Delta [i]**	Selected Delta in Influent and Effluent [j]	**Total Unpaired Samples Needed in Each Group for Selected Delta [i]**
	Before	After			Before Installation	After Installation	Before [g]	After [h]					
Bioretention	73	18	55	9.8	62	62	127	3.5	88.1	69.3	**21**	35	**81**
Bioswale	40.8	27.1	13.7	44.2	63	63	127	21.2	89.3	34.9	**82**	35	**83**
Media Filter	96	22.4	73.6	16.2	34	34	127	3.7	86.9	62.9	**25**	47	**43**
Manufactured Device	117	47.5	69.5	64.6	49	49	127	14.4	88.2	14.5	**461**	47	**44**
Permeable Pavement	189.4	29.4	160	29.4	73	73	127	63.1	98.6	49.7	**50**	35	**101**

	ISBMPD Std Dev [k]		
	Before	After	Paired Difference in Means
Bioretention	100.00	30.00	107
Bioswale	53.6	35.2	50
Media Filter	134.9	40.8	123
Manufactured Device	320.4	94.9	267
Permeable Pavement	389.8	63.1	393

Paired Sample Size Calculations to Test Statistically Significant Removal

	Initial Guess on Sample Size for *t* Test, Iterate to Match Calculated Samples	Average Std. Dev. of Percent Removal from PSGM [l]	Estimated Volume Reduction	Estimated Concentration Change to Achieve Load Removal Delta	**Total Paired Samples Needed for Expected Delta, *t* Test [m]**	For Comparison, Output from GPower3	Initial Guess on Sample Size for *t* Test, Iterate to Match Calculated Samples	Selected Delta for Percent Removal	**Total Paired Samples Needed for Selected Delta, t Test [m]**
Bioretention	4	0.26	0.30	0.43	**5**	3	4	0.44	**4**
Bioswale	3	0.2	0.30	0.43	**4**	3	3	0.44	**4**
Media Filter	3	0.26	-	0.60	**3**	3	3	0.60	**3**
Manufactured Device	5	0.42	-	0.60	**5**	4	5	0.60	**5**
Permeable Pavement	29	0.42	0.50	0.20	**29**	6	8	0.44	**8**

a. National Stormwater Quality Database, Version 1.1, Maestre and Pitt (2005) .

b. Unpaired monitoring is before and after BMP installation. Paired monitoring is influent/effluent or paired watersheds that are monitored during the same events.

c. International Stormwater BMP Database.

d. Regressions from the Caltrans BMP Retrofit Pilot Program Final Report (2004); Bioretention is based on the MCTT permeable pavement taken as the mean effluent from the ISBMPD.

e. Iterate for solution by matching Total Samples Needed in Each Group.

f. Unequal variance results in an approximate solution for unpaired samples requirements.

g. Based on 1.6 times the selected mean because variance within the Bay Area retrofits is expected to be less than that from the IBMPD because of the wide variation in land use and topography in the IBMPD. COV for TSS from Maestre and Pitt ranged from 1.6 (Industrial) to 2.0 (Commercial). Std. Dev. for TSS at maintenance stations was reported at 95, and 188 for highways.

h. Based on Caltrans (2004), except for permeable pavement value of 63.1 from the ISBMPD.

i. For a normal distribution and large sample sizes (*Z* test). Groups are before construction and after construction. Formula (from EPA QA/G-95) assumes constant variance, so solution is approximate.

j. This is an estimate of the percent reduction in TSS load that would be meaningful to load reduction goals (e.g., difference required for TMDL compliance or expected mass reducion of other retrofit or muni-operation alternatives). 60% removal is used by Caltrans (PPDG) to differentiate between marginal and good mass removal.

k. Informational for comparison of selected variance to the variance in the International Stormwater BMP Database.

l. Typical values as reported in the Caltrans Pilot Study Guidance Manual, Appendix K. The upper range for sedimentation BMPs was used for manufatured devices and permeable pavement. The filtration value was used for bioretention.

m. Biostatistical Analysis, 4th Ed., Zar, Jerrold. Equation 7.8. p106.

Table A-3. Sediment-based Power Analysis for Determining Minimum Sample Size for Comparison of Effluent among BMP Types

Paired and Unpaired Sample Size Calculations to Compare among BMPs (not exhaustive; other comparisons to explore include slope and intercept differences of influent/effluent regressions).

	Assumed Average Effuluent BMP 1	Assumed Average Effluent BMP 2	Sample Size for both (n,m)	S, BMP 1	S, BMP 2	Common Standard Deviation for Influent and Effluent S_p	Estimated Delta	Total Unpaired Needed to Compare BMPs, Z Test	initial guess on sample size for t test, iterate to match calculated samples	Total Unpaired Needed to Compare BMPs, t Test	For Comparison, Output from GPower3
Bioretention & Bioswale	9.8	44.2	3	3.5	21.2	10.7	-34.4	3	3	**3**	3
Bioretention & Media Filter	9.8	16.2	3	3.5	3.7	2.5	-6.4	3	3	**3**	4
Bioretention & Manufactured Device	9.8	64.6	3	3.5	14.4	7.4	-54.8	3	3	**3**	2
Bioretention & Permeable Pavement	9.8	29.4	46	3.5	63.1	43.7	-19.6	46	48	**use Z**	48
Bioswale & Manufactured Device	44.2	64.6	6	21.2	14.4	15.3	-20.4	6	7	**7**	8

a. Faul, F. et al. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods. 41 (4), 1149-1160.