STATE DEPARTMENTS OF TRANSPORTATION STORMWATER MONITORING PROGRAM GOALS, OBJECTIVES, AND PROTOCOLS

Prepared for:

National Cooperative Highway Research Program

Transportation Research Board of

The National Academies of Sciences, Engineering and Medicine

Prepared by:

WSP USA
1 East Pratt Street
Suite 300
Baltimore, MD 21202

The information contained in this report was prepared as part of NCHRP Project 25-25, Task 101, National Cooperative Highway Research Program.

NCHRP 25-25B
(Task 101)
FINAL REPORT

Acknowledgements

This study was conducted for the AASHTO Standing Committee on the Environment (SCOE), with funding provided through the National Cooperative Highway Research Program (NCHRP) Project 25-25, Task 101, *State Departments of Transportation Stormwater Monitoring Program Goals, Objectives, and Protocols*. The NCHRP is supported by annual voluntary contributions from the state Departments of Transportation. Project 25-25 is intended to fund quick response studies on behalf of SCOE. The report was prepared by WSP USA. The work was guided by a technical working group that included:

- G. Scott McGowen, Michael Baker International (Chair)
- Henry Barbaro, Massachusetts DOT
- William B. Fletcher, II, Oregon DOT
- Kenneth M. Stone, Washington State DOT
- Nicklas Tiedeken, Minnesota DOT
- Susan Jones, Federal Highway Administration (Liaison)

The project was managed by Ann Hartell, NCHRP Senior Program Officer.

TABLE OF CONTENTS

Table of Contents

TABLES

FIGURES

GLOSSARY

Best Management Practice - Per 40 CFR § 122.2, a Best Management Practice is a schedule of activities, prohibitions of practices, maintenance procedures, and other management practices to prevent or reduce the pollution of 'waters of the Unites States'.

Coefficient of Variation – a descriptive statistic used to express the variability in the data. It is calculated as the standard deviation divided by the mean.

Composite Sample - a mixed or combined sample that consists of a series of individual sample aliquots (sub-samples) that are collected at either specific times, specific flow volumes, or specific flow rates. Composites are typically referred to as “time-weighted” or “flow-weighted” depending on how they are collected.

Continuous Measurements – measurements that are made at a field monitoring station using an in-situ monitoring device (such as a water quality sonde) that is mounted permanently in the flow path. These measurements do not require physically collecting a sample and no laboratory analysis is required. Parameters that are typically measured continuously with an in-situ device include water temperature, specific conductivity, dissolved oxygen, and pH.

Discrete Sample - an individual sample collected by hand or by automated equipment at a single point in time during a storm.

Flow-weighted Composite – a mixture of multiple sample aliquots that accounts for variations in flow during a storm event. In a flow-weighted sample, peak flows are more heavily weighted than low flows.

Load Allocation – as part of a Total Maximum Daily Load calculation, it represents the maximum amount of a pollutant allowed to enter a waterbody from nonpoint sources that would still meet water quality standards for that pollutant. Load allocations include all remaining sources of the pollutants not covered under wasteload allocations, as well as natural background. See also definition of wasteload allocation.

Margin of Safety –as part of a Total Maximum Daily Load (TMDL) calculation, it represents the amount of pollutants attributed to uncertainty in the TMDL analysis.

Pollutant – a pollutant is considered a water quality constituent that is present in excess, or is inherently toxic to aquatic biota.

QA/QC – quality assurance (QA) is the process or set of processes used to measure and assure the quality of a product, and quality control (QC) is the process of ensuring products and services meet project expectations.

Reserve Capacity – as part of a Total Maximum Daily Load calculation, it represents the amount of pollutants set aside for future development.

Sample Aliquot – a small sub-sample that represents part of a composite sample, usually about 50 ml to 500 ml in volume. Sample aliquot is distinct from a “sample” which is the total sample delivered to the laboratory for analysis.

Sample Type - refers to the kind of storm sample to be collected - either a discrete sample or a composite sample.

Sampling Frequency– the target for how often storms are to be sampled over the total duration of monitoring, e.g. 1 storm/month. Sampling frequency does not refer to how often samples are to be collected within a given storm.

Sampling Technique - the method by which a sample is physically collected - either manually (by hand or using an extension pole) or by automated sampler.

Sheetflow – diffuse, non-channelized storm runoff generated from impervious surfaces such as roadways.

Total Maximum Daily Load - a calculation of the maximum load of a pollutant a water body can receive and still meet local water quality standards. A Total Maximum Daily Load is the sum of all point source loads (waste load allocations) and nonpoint source loads (load allocations), plus a Margin of Safety, and in certain areas a reserve capacity.

Treatment Train – A series of individual or discrete BMPs designed to remove the pollutants in stormwater runoff. The runoff flows from one BMP to the next in the sequence, and is treated multiple times. Each BMP generally has its own inflow and outflow point.

Waste Load Allocation - as part of a Total Maximum Daily Load calculation, it represents the maximum load of a pollutant a water body can receive from point sources and still meet water quality standards for that pollutant. Waste Load allocations include all point sources subject to regulations under the National Pollutant Discharge Elimination System program.

Water Quality Constituent – a compound in water that can be quantified, either as a concentration or other measure of magnitude. A water quality constituent is not necessarily a pollutant source, unless it is present in excess or is inherently toxic to aquatic biota. See also pollutant.

ABBREVIATIONS

AADT	Annual Average Daily Traffic
AASHTO	American Association of State Highway and Transportation
AC	Alternating Current
ADT	Average Daily Traffic
ANOVA	Analysis of Variance
BMP	Best Management Practice
BOD	Biological Oxygen Demand
COC	Chain of Custody
COD	Chemical Oxygen Demand
COV	Coefficient of Variation
CFR	Code of Federal Regulations
Cu	Copper
DA	Drainage Area
DC	Direct Current
DO	Dissolved Oxygen
DOT	Department of Transportation
EMC	Event Mean Concentration
FHWA	Federal Highway Administration
GIS	Geographic Information System
GOES	Geostationary Operational Environmental
HDPE	High-Density Polyethylene
HSP	Health and Safety Plan
HSPF	Hydrologic Simulation Program-Fortran
ISB	International Stormwater BMP
IQR	Interquartile Range
LA	Load Allocation
LID	Low Impact Development
MASSDOT	Massachusettes Department of Transportation
MOS	Margin of Safety
MS4	Municipal Sepapearate Storm Sewer System
NCEI	National Centers for Environmental Information
NCHRP	National Cooperative Highway Research Program
NCDOT	North Carolina Department of Transportation
ND	Non-Detects
NH3-N	Ammonia
NHR	National Highway Runoff
NPDES	National Pollutant Discharge Elimination System
NO3-2	Nitrate-Nitrite as N
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
ODOT	Oregon Department of Transportation
O&G	Oil and Grease
O & M	Operations and Maintenance
PAH	Polycyclic Aromatic Hydrocarbon
Pb	Lead

PCB	Polychlorinated biphenyl
PFC	Permeable Friction Course
PO4	Orthophosphate
PVC	Poly(vinyl chloride)
QA	Quality Assurance
QAPP	Quality Assurance Project Plan
QC	Quality Control
QPF	Quantitative Precipitation Forecast
RVTS	Roadside Vegetated Treatment Sites
SELDM	Stochastic Empirical Loading and Dilution Model
TDS	Total Dissolved Solids
TKN	Total Kjeldahl Nitrogen
TMDL	Total Maximum Daily Load
TN	Total Nitrogen
TOC	Total Organic Carbon
TP	Total Phosphorus
TPH	Total Petroleum Hydrocarbons
TSS	Total Suspended Solids
UHF	Ultra High Frequency
USEPA	United States Environmental Protection Agency
USGS	United States Geological Survey
VHF	Very High Freqeuncy
VSS	Volatile Suspended Solids
WET	Whole Effluent Toxicity
WLA	Waste Load Allocation
WMO	World Meteorological Organization
WSDOE	Washington State Department of Ecology
WSDOT	Washington State Department of Transportation
Zn	Zinc

EXECUTIVE SUMMARY

An effective and successful stormwater monitoring program yields useful data and information that meet the specific goals and objectives of the monitoring program. Given the wide variability in practices that can be applied to collect stormwater monitoring data, a well-defined monitoring program is essential to enable transferability of the data and to ensure that specific objectives of the monitoring program are met. The purpose of this report is to provide guidance to State Departments of Transportation (DOTs) on how to design an effective monitoring program that is proportionate and appropriate to the purpose of stormwater sampling in any state.

Water quality data obtained from a sampling program may be used for a wide range of purposes: to describe the health condition of a waterbody; to provide a general understanding of the sources of impairment, such as stormwater runoff; to identify patterns and trends of water quality condition; to identify potential problems and major or new pollutants (constituents) of concern; to evaluate the effectiveness of pollution control measures, as well as to provide guidance to designing mitigation practices. Knowing why and how the water quality data will be used to achieve each of these diverse objectives is crucial to the development and implementation of an effective stormwater monitoring program. Chapter 1 of this report identifies and describes the monitoring goals applicable to State DOTs.

A broad-based search and literature review was conducted of existing stormwater monitoring and overall stormwater management programs at DOTs. This effort provided a better understanding of the framework and level of existing monitoring programs. The literature review covered monitoring programs that were developed and implemented for discretionary purposes and examined National Pollutant Discharge Elimination System/Municipal Separate Storm Sewer System (NPDES/MS4) permit water quality monitoring requirements and specific Total Maximum Daily Load (TMDL) obligations. DOT stormwater programs and monitoring plans, as well as available studies and/or published reports identified to be relevant and potentially beneficial to the advancement of this research, were thoroughly reviewed, synthesized, and annotated.

From the literature review, the research team identified and compiled four (4) primary stormwater monitoring goals that are applicable, useful, and relevant to State DOTs. The stormwater monitoring goals include:

- Goal 1 - Characterize Stormwater Runoff Quality
- Goal 2 - Assess BMP Performance
- Goal 3 – TMDL Effectiveness Monitoring
- Goal 4 – Evaluate Efficiency of Routine Maintenance Activities

DOT runoff has complex dynamics and therefore requires a clear understanding of the governing key elements in order to develop effective monitoring protocols. The condition of DOT stormwater runoff, its constituents, and the level of pollution are highly influenced by the type of the roadway (interstates, local roads, etc.), traffic volumes, maintenance and operation conditions, and the placement of stormwater best management practices (BMPs) and/or control devices. The impact on water quality can further be characterized based on precipitation and storm patterns, surrounding land uses, topography, and the population density of the geographical region. Chapter 2 describes the significant key monitoring

elements needed to develop an effective State DOT stormwater monitoring program. The following key elements were identified based on literature review:

- Monitoring Station Network
- Water Quality Pollutants
- Duration of Monitoring
- Number of Samples to Sample
- Sampling Methods
- Monitoring Equipment
- Storm Event Criteria
- Health and Safety Plan
- Data Analysis Technique
- Quality Control (QC) Procedures

With key elements appropriately identified, analyzed, and linked, the research team developed a set of protocols to create a monitoring plan for each of the monitoring goals. The term “protocols” refers to a set of comprehensive key elements essential to establishing an efficient way to monitor stormwater quality suited for specific monitoring goals and objectives. Chapter 3 presents monitoring protocols for the stormwater monitoring goals. The monitoring protocols address key elements for an effective DOT monitoring program, including physical, chemical, and biological water quality pollutants; sampling techniques; location, number, and geographic distribution of monitoring sites; sampling equipment; frequency and duration of monitoring; data analysis methods; and seasonal, physical, and safety constraints. The recommendations may differ depending on the DOT’s monitoring objective, or may apply universally to all a DOT’s monitoring goals.

Figure 1 displays the relationship between the major components of a monitoring program, their associated key elements, and the corresponding sections in this report.

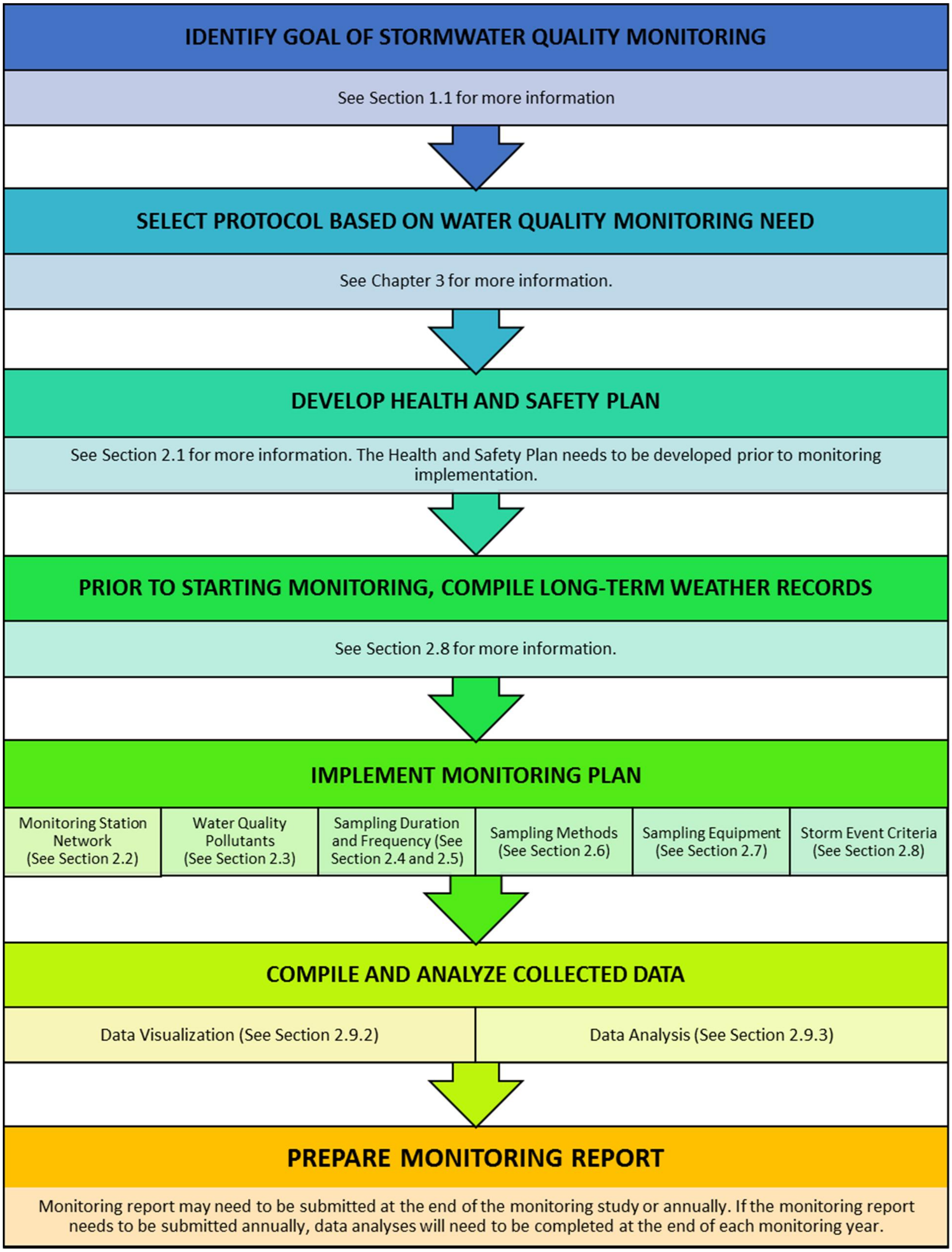

Figure 1. Monitoring Program Process Flow Diagram

CHAPTER 1 – MONITORING GOALS

1. Introduction

The purpose of this report is to provide state Departments of Transportation (DOTs) with guidance on how to design an effective stormwater monitoring program to evaluate storm runoff from DOT highways and facilities. A successful program must be appropriate to the purpose of stormwater sampling in any given state. Guidance is provided for all aspects of program development from the initial development of the study design, to data collection methods, equipment needs, quality control (QC) measures, weather tracking, health and safety plans, data management and analysis, and ultimately interpretation of the results.

Water quality data obtained from a sampling program may be used for a wide range of purposes. For example, the data may be used to describe the health condition of a waterbody, to characterize stormwater runoff quality and quantity, to identify patterns and trends in water quality over time, to identify major pollutants of concern, to evaluate the effectiveness of pollution control measures, and to provide guidance on designing mitigation practices. Knowing how the water quality data will be used to achieve each of these diverse objectives is crucial to the development and implementation of an effective stormwater monitoring program.

The guidance and recommended monitoring protocols provided in this report are based on the findings of an extensive literature review of stormwater monitoring-related publications. The review included many different types of studies, guidance manuals, and other reports by state DOTs, the Federal Highway Administration (FHWA), the American Association of State Highway and Transportation Officials (AASHTO), the National Cooperative Highway Research Program (NCHRP), the U.S. Environmental Protection Agency (USEPA), state regulatory agencies, universities, and many others. Those studies that were deemed relevant and beneficial to the advancement of this research were carefully reviewed, synthesized, and annotated. The literature review covered monitoring programs that were developed for discretionary purposes, as well as for National Pollutant Discharge Elimination System (NPDES) Municipal Separate Storm Sewer System (MS4) permit requirements and specific Total Maximum Daily Load (TMDL) obligations. Based on this review, four monitoring goals were identified as being applicable to state DOTs, where achieving these goals would provide the DOT with valuable information to support their stormwater management program while meeting regulatory requirements. The four goals are listed as follows and will be discussed further below.

- Goal 1 - Characterize Stormwater Runoff Quality
- Goal 2 - Assess Best Management Performance (BMP) Performance
- Goal 3 – TMDL Effectiveness Monitoring
- Goal 4 - Evaluate Efficiency of Routine Maintenance Activities

The literature review also elucidated that DOT stormwater monitoring programs across the U.S. vary widely, from non-existent to very comprehensive. The variation appears to be due to several factors such as NPDES permit and TMDL requirements (where applicable, not every DOT has storm monitoring requirements or is subject to a TMDL agreement), as well as the nature and extent of local pollution sources, funding availability, existing data gaps, and DOT stormwater management program priorities. Therefore, the protocols included herein will not apply everywhere.

There are instances when monitoring by the DOT would not be warranted. Some examples of this include: 1) if the DOT's contributing area of highway runoff to an impaired receiving water is negligible, 2) if the pollutant of concern is unrelated to stormwater runoff, and/or 3) if the runoff is known to come from off-site (i.e. outside the DOT's jurisdiction). In all these cases, monitoring would be of little value and would be a drain on scarce resources. However, for the purposes of this report, it is assumed that monitoring is in fact necessary, and that the DOT is building the monitoring program from the ground up.

This report is focused on stormwater monitoring and specifically on the measurement of stormwater runoff quality and volume from DOT highways and transportation infrastructure. Dry weather monitoring such as the sampling of street sweeping waste, catch basin sediment and debris, and gross solids (trash) are not addressed in this report. Although dry weather monitoring provides valuable information to support a DOT's stormwater management program, it does not directly involve wet-weather measurements of storm runoff, hence it is omitted. Also, this report does not address stormwater monitoring in surface waters (streams, rivers, lakes), as it is generally outside the purview of DOTs. Lastly, this report does not address water quality modeling.

This document is structured as follows. The remainder of this chapter describes the four DOT monitoring goals identified above and outlines the benefits of each. For each goal, a general monitoring strategy is presented. Chapter 2 provides further details on these strategies and recommends specific protocols for designing an effective stormwater monitoring program, including all of the key monitoring elements that must be considered. Throughout Chapter 2, an effort was made to provide several examples of DOT monitoring designs and existing studies from different regions across the U.S. The sources are provided where DOTs may obtain more information. Chapter 3 provides a condensed summary of the recommended protocols from Chapter 2 for each of the identified key monitoring elements. Throughout this report, the monitoring protocols are broken down by goal where different approaches are needed to achieve a particular objective. However, in certain cases, the recommended protocols are the same regardless of the monitoring goal, and therefore a single protocol is described which applies universally to all of the goals. This reflects the fact that some monitoring procedures are widely applicable while others are relatively specialized and are only needed in certain cases to meet a particular monitoring objective.

1.1. Monitoring Goals

This section provides a description of the four state DOT monitoring goals, highlights their benefits to the DOT, and recommends some general monitoring approaches. This includes some basic recommendations for how to design the study in order to maximize its value to the DOT. This section is intended as a broad overview of the types of factors to consider when preparing a monitoring plan for a specific goal. More detailed information will be provided in Chapter 2.

Goal 1: Characterize Stormwater Runoff Quality

Characterizing stormwater runoff quality from DOT highways and transportation infrastructure may be needed to meet NPDES permit requirements, or as part of a broader data collection effort to support DOT stormwater management programs. For DOTs subject to TMDL requirements, characterization monitoring also forms the basis of a TMDL compliance strategy (see Goal 3 below).

Characterizing stormwater runoff quality and volume provides multiple benefits. First, it provides data on the type and magnitude of pollutants[1] present in the stormwater discharged from the DOT right-of-way. Characterization of stormwater runoff identifies the primary pollutants of concern and indicates where (or if) stormwater needs to be treated and managed. This provides a basis for selecting what types of BMPs or source control activities are needed to treat stormwater effectively. Second, characterizing runoff quality may identify the presence (if any) of contaminant "hotspots" where pollutants tend to accumulate. Third, characterization data can be used to understand how traffic volume (annual average daily traffic or AADT), precipitation characteristics, or land use affect stormwater runoff quality. Fourth, measuring the volume of runoff from a contributing drainage area provides information on the water quality volume needed for sizing treatment BMPs. Finally, characterization monitoring data can be used to support research and development needs, and to populate water quality models for conducting pollutant load assessments.

To characterize stormwater runoff from DOT highways and facilities, data must be gathered from representative areas before the runoff discharges to a receiving water or flows outside the DOT's right-of-way. It is generally easier to monitor where the runoff naturally collects and can thus be measured and sampled, for example along a concrete road-side ditch or in a storm drain pipe. Alternatively, a containment device may be constructed (such as a flume, weir, or piped conduit) to channelize and direct the runoff to the point of measurement, although this is more labor- and materials-intensive. It is important that the samples are not influenced by any external factors from outside the DOT's right-of-way, for example runoff from the surrounding watershed. In addition, areas with atypical conditions should be avoided, for example highways with construction projects or highways with an abnormal amount of erosion of the surface material.

It is generally recommended to sample at the base of a highway section at a pipe outfall. This provides safer working conditions for field personnel since it is away from vehicle traffic. Monitoring larger pipes is generally better than monitoring smaller ones, since they have larger contributing drainage areas and therefore allow a broader area to be characterized. Any spatial variation in runoff quality across the drainage area will be averaged at the sampling point. Monitoring larger pipes also helps ensure there will be sufficient runoff present to collect samples and make flow measurements. While it is possible to monitor smaller pipes, there is a risk that the runoff draining into the pipe will not be typical of the larger highway section.

When selecting a monitoring location, careful planning is needed. It is preferable to monitor relatively homogeneous highways in terms of design, road surface material, traffic, adjacent soils, surrounding land use, and other factors that can influence stormwater runoff quality. This provides a level of experimental control to ensure that the monitoring data are representative of the area of interest. Monitoring homogeneous highways also facilitates extrapolation of the monitoring results to other highways with similar conditions, and thus larger areas can be characterized without the need for additional monitoring. On the other hand, monitoring large complex highways of varying design and characteristics should be

[1] For the purposes of this report, a "pollutant" is defined as a water quality constituent that is present in excess, or is inherently toxic to aquatic biota. See also Footnote 7.

avoided. Other areas to avoid include eroded road shoulders, steep slopes, cut and fill areas, and areas that have the potential to be developed or redeveloped.

Finally, DOTs may wish to conduct characterization monitoring specifically to detect or evaluate suspected or known hotspots. Hotspots are localized areas where contaminants are abnormally high, reflecting the presence of some activity or circumstance which has caused an atypical condition that is not representative of the surrounding area. In the transportation environment, hotspots may potentially occur at, but are not limited to, vehicle fueling stations, along road shoulders where herbicides have been sprayed, inside equipment storage areas, adjacent to industrial areas, in homeless encampments, on highways where a fuel spill or vehicle accident has occurred, or at trash dumping spots. For the purposes of this report, hotspot monitoring is considered a specialized type of characterization monitoring. The monitoring strategy is the same as described above (and further detailed in subsequent sections) except the targeted area of interest is intentionally *not* typical of other areas and a wider or alternative list of pollutants may need to be monitored (e.g. PCBs and PAHs, see Table 1 in Section 2.3) to account for the specific condition or activity being evaluated.

Goal 2: Assess BMP Performance

Assessing the performance of existing stormwater treatment BMPs is a key component of a DOT's stormwater management program and is often required by NPDES permits or other regulatory requirements. Assessing a BMP's performance through monitoring helps identify which BMP types and designs (including retrofits) are most cost-effective at addressing water quality impairments. This is important for DOTs for whom BMP implementation is especially challenging due to the lack of space in urban areas, and given that traditional BMPs may not be adaptable to the linear highway environment. Such difficulties drive the demand for innovative technologies to treat stormwater such as permeable friction courses, alum injections, media filter drains, and low impact development (LID) or green infrastructure practices. Some studies have demonstrated that innovative treatment technologies are better at treating pollutants than traditional structural BMPs, particularly for pollutants which are relatively difficult to treat such as nitrogen and dissolved metals.

Another benefit of BMP performance monitoring is that it contributes to a growing body of knowledge on BMP effectiveness. One well-known depository of BMP performance data is the International Stormwater BMP Database, which includes the findings of hundreds of BMP monitoring studies (although not all are highway-related). These studies help provide an understanding of how regional characteristics such as climate, geology, vehicle traffic, surrounding land use, and other factors can potentially affect BMP functioning. DOTs that share BMP monitoring data through such databases, or even informally through DOT workshops or other meetings, promote a wider understanding of various stormwater treatment methods and facilitate technology transfer for the mutual benefit of all DOTs.

The basic monitoring strategy for assessing the performance of structural BMPs (e.g. extended detention ponds, infiltration basins, bioretention areas) is to conduct paired influent-effluent monitoring at the BMP's inlet and outlet[2]. If there are multiple inflow or outflow points, all of them should be monitored to properly assess the BMP's performance. Monitoring stations should be selected carefully to ensure that the BMP's ability to treat pollutants is not affected by outside factors. For example, the influent station

[2] Water quality model simulations and laboratory bench-scale testing to assess BMP performance are not addressed in this report, but may potentially be conducted in lieu of field monitoring to save on resources.

should be located as close as possible to the BMP without being influenced by the BMP. If the water table is high, the influent station should be located above the water table to ensure that groundwater does not mix with stormwater runoff. Locations with backflow, steep slopes, and/or dense vegetation should be avoided (Geosyntec and Wright Water Engineers, 2009).

Typically, the influent samples are collected from the bottom of the flow stream (e.g. the invert of a pipe) since that is the easiest place to install the sample intake line of an automated sampler (see Section 2.6.2). This may introduce some bias in the data, particularly for larger storms with higher flows, because sediment concentrations and particle sizes are often greater at the bottom of the flow stream and therefore are not representative of the entire water column. If this is a concern, the intake can be mounted slightly above the invert (e.g. along the sidewall of a pipe) and/or the sampling program can be set to trigger at a higher level or flow threshold to ensure that the sediment particles are more uniformly mixed when the samples are collected[3]. In addition, this bias may to some extent be mitigated by evaluating the BMP's performance in terms of load reductions, not concentration reductions, since this method accounts for the volume of runoff.

There are cases, however, where paired influent-effluent monitoring may not be feasible due to long lag times between inflows and outflows. For example, wetlands, extended detention ponds, and other larger BMPs tend to have relatively long detention times, on the order of several hours or even days. In this case, it is recommended to conduct paired influent-effluent monitoring for no more than 24 hours, or until 75% of the total storm volume has been captured, whichever comes first (Washington Stormwater Center, 2008). For this reason, automated flow-weighted composite sampling is highly recommended (discussed later). For more information on monitoring BMPs with long detention times, see the Washington Stormwater Center's *Evaluating Stormwater Treatment Technologies with Long Detention Times* (Washington Stormwater Center, 2008).

Monitoring BMPs that lack a well-defined inlet and/or outlet is generally more difficult because they require the use of constructed devices or more specialized monitoring equipment to channelize the flow for measurement and sampling purposes (see Section 2.7.2). Some examples include vegetated filter strips, swales, or grass channels that run parallel to the highway and receive sheetflow along their entire length. In general, when monitoring BMPs that lack an inlet and/or outlet, a higher labor and time investment should be expected relative to monitoring more traditional structural BMPs with well-defined inflow and outflow points.

In some cases, it may be desirable to monitor within the BMP itself to understand how the BMP is functioning internally. In such cases, the interior station should be in an area that is representative of the BMP's ability to treat pollutants at that particular point. This requires a clear understanding of the flow pathway through the BMP system to avoid "dead zones" that are not representative of the BMP's functioning. For example, monitoring the interior of a wetland or a wet detention pond should be performed in a location (typically in the middle of the wetted area) where the slope, vegetation, and channel width are relatively uniform to avoid any microcosms where these conditions differ. If a BMP has a poorly defined or non-existent outlet, and it is not feasible to artificially channelize the flow at the outlet by building a containment structure, interior monitoring is mandatory to accurately evaluate the BMP's

[3] It is acknowledged that even using a higher trigger will not eliminate all the bias. While there are more complex methods available to help reduce the errors (e.g. use of static mixers, movable sampling arms, and/or sampling from multiple points within the water column), these methods are not included in this report. For more information on these methods, see a USGS study of highway runoff quality by Smith (2002).

performance. Monitoring can be conducted in underdrains, the vadose zone, and/or in areas with groundwater or sheetflow.

Some DOTs may wish to monitor more complex BMP systems with multiple treatment facilities or chambers that are linked in series, commonly known as “treatment trains.” In these types of systems, storm runoff flows from one part of the system to the next and is treated multiple times. Alternatively, it may be useful to monitor each component of the overall system separately, or monitor a sub-set of the components. In this context, a BMP component is defined as having a unique inflow and outflow point, a self-contained pollutant reduction mechanism, and a clear flow pathway from one component to the next. The overall pollutant removal of the entire treatment train may be assessed by monitoring the initial inflow into the first BMP component and the final outflow from the last BMP component. If the goal is to evaluate the performance of the individual BMP components within the treatment train, each component’s inflow and outflow point must be monitored. In this case, it is important to sample just as the water is discharged from one BMP component and just before it enters the next BMP component in the sequence. This avoids any influence from factors unrelated to the BMP component’s performance. Monitoring should not be conducted where any backflow or mixing occurs, which is especially likely during high flows (Geosyntec and Wright Water Engineers, 2009).

Goal 3: TMDL Effectiveness Monitoring

The purpose of TMDL effectiveness monitoring (where applicable, not all DOTs are subject to TMDLs) is to evaluate if the DOT’s TMDL Implementation Plan has resulted in any measurable water quality benefits in the impaired receiving waters. An effectiveness monitoring plan should be developed during, or immediately following, TMDL development (USEPA, 2011). Effectively addressing TMDLs requires monitoring of baseline pollutant loads within, and adjacent to, the DOT’s right-of-way. The monitoring data are then used to populate and calibrate a water quality model to estimate the potential load reductions that can be achieved under various scenarios of BMP implementation.

TMDL monitoring can help DOTs meet several objectives. This includes calculating or validating the assigned Waste Load Allocation (WLA) to allow the DOT to measure progress toward meeting TMDL requirements. TMDL effectiveness monitoring provides data to help determine what level of load reduction is realistically achievable, and to enable calculating the absolute and relative contributions of a highway to the total load in the TMDL water body. TMDL effectiveness monitoring data can also be used prior to, and during, TMDL development to more accurately estimate the DOT’s actual load. Additionally, this type of monitoring enables the DOT to share monitoring data and collaborate with local, state, and federal agencies in developing new TMDLs. Getting the DOT involved early in the TMDL development process helps ensure that the DOT is not disproportionately burdened by TMDL load reduction requirements.

Several DOTs are required to conduct TMDL compliance monitoring. For example, the California Department of Transportation (Caltrans) performs TMDL compliance monitoring in 84 watersheds throughout California encompassing over 300 reaches. One monitoring study addressed the Lake Tahoe TMDL for phosphorus, nitrogen, and fine sediment particle impairments. Based on the findings of the study, Caltrans was able to reduce fine particle (less than 16 micrometers) loading to Lake Tahoe by employing high efficiency sweepers. As another example, the New Hampshire DOT conducted a water quality monitoring study to address a chloride TMDL. The results of the study helped them reduce their road salt usage by 20 percent by upgrading their snow plows with an advanced technology to automatically measure road surface temperature to prioritize spray areas.

In general, state DOTs face the unique challenge of having linear highway systems that traverse multiple watersheds, making them subject to multiple TMDLs. In many cases, the DOT's load contribution to the impaired receiving water is negligible relative to the overall contribution from the much larger surrounding watershed. Nonetheless, DOTs may be required to reduce loads from many miles of highway. Since it is not feasible to monitor everywhere, TMDL effectiveness monitoring should focus on collecting a limited amount of site-specific data which can then be used to calibrate a water quality model to evaluate the overall pollutant loading of the entire highway system within the TMDL watershed. When using models, it is important to collect high-quality, representative monitoring data so that the model can extrapolate over unmonitored areas with reasonable accuracy. The extent of monitoring will depend on the diversity of highway types (e.g. high traffic versus low traffic) within the TMDL watershed. A greater diversity of roadways would require additional monitoring to quantify any potential differences in pollutant contributions among highways of different types.

TMDL effectiveness monitoring is best conducted in close proximity to the TMDL receiving water in an area where some water quality improvements are expected. An ideal location to monitor is at a pipe outfall that drains a highway section where one or more pollution control strategies have been implemented, for example street sweeping, catch basin clean-outs, or other source control strategies or BMPs. It is important to sample the stormwater just before it discharges into the TMDL receiving water (i.e. samples should not be collected within the receiving water). This ensures that only the DOT's pollutant contributions are captured, not those from the larger watershed which is outside the DOT's jurisdiction.

The recommended study design for TMDL effectiveness monitoring is a before/after study. In this type of study, monitoring is conducted at the same monitoring locations both before and after TMDL implementation. The before and after data are then compared and any differences are attributed to the effect of the pollution control strategies. This requires that either the TMDL implementation has not already begun, or that historical data are available for use as a baseline. Ideally, the before and after datasets should have the same (or approximately the same) variance in pollutant concentrations and flow volumes. This may not always be feasible but can greatly enhance the statistical power of the findings. The power analysis method (see Section 2.5) may be used to calculate the minimum sample size needed to achieve statistical significance. At least 2 years of data are needed to perform a meaningful before and after assessment (see Section 2.4 for a discussion of monitoring duration).

TMDL effectiveness monitoring should also include trend monitoring of water quality pollutants and runoff characteristics. Trends are an important means of showing incremental improvements in water quality over time. For trend monitoring, a long data record is needed consisting of continuous flow measurements and periodic water quality measurements from storm event sampling. Again, at least 2 years of data are needed to detect potential trends in water quality over time. At shorter time scales, detecting a trend can be more difficult due to the potential for various confounding factors such as a protracted TMDL implementation schedule, a lag time in the water quality benefits, natural variations in weather and hydrology over time, and measurement errors. Trend analysis should be conducted using proper statistical techniques to avoid misinterpreting the data (see Section 2.9.4.3 for discussion of trend evaluations for TMDL effectiveness monitoring).

Goal 4: Evaluate Efficiency of Routine Maintenance Activities

Certain routine DOT maintenance activities[4], such as street sweeping, catch basin clean-outs, trash pick-ups, and ditch cleaning provide pollutant source control -- that is, they remove pollutants from the road surface and drainage system before they are mobilized by storm runoff. These practices can potentially improve stormwater quality if they are performed on a regular basis (and with the right technology in the case of sweeping). However, it is also possible that these practices could increase pollutant loads, which would warrant modifying the practice to achieve some water quality benefit. At a minimum, monitoring of maintenance activities helps identify an appropriate maintenance interval. For example, monitoring data can indicate how often drain and ditch cleaning is needed to ensure proper drainage and avoid the risk of flooding or damage to drainage structures. Ditch monitoring may also help quantify any decreases in pollutant loads discharged to receiving waters, which can be important for TMDL compliance purposes (see Goal 3 above). It is important to emphasize that monitoring the water quality benefits of maintenance activities can be used for water quality "credits" in a permitting and/or TMDL context.

Monitoring of street sweeping as it relates to improvements in stormwater runoff quality is particularly important. Street sweeping is widely recognized as a means of water quality control, and advances in sweeper technology such as regenerative air or vacuum trucks has increased the efficiency of waste collection from road surfaces. Stormwater monitoring data on the efficiency of sweeping would help support the DOT's stormwater management program, and potentially provide significant cost savings by avoiding the need for large structural BMPs. Stormwater monitoring data would also indicate what frequency of sweeping is needed to achieve a water quality benefit[5].

Other routine DOT maintenance includes road-side vegetation control using herbicides and the application of de-icing materials to the road surface during winter storms (in cold climates). There are benefits to monitoring runoff from an herbicide application, as well as monitoring during, or after, snow storms to measure de-icing agents. Monitoring is directed toward reducing the use of these chemicals (herbicides and de-icing materials) to save on resources and minimize the impacts of these chemicals on receiving waters. In the case of de-icing practices, several studies have indicated that advances in technology or shifts in the types of materials applied (salt versus sand) can result in significant material savings (see Alwan and Casey, 2014 and Barbaro, 2006). If the type of materials applied are changed, the list of monitoring parameters would need to be altered accordingly to reflect the difference in pollutant type and concentration.

Adjustments in vegetation management can also be considered a water quality BMP. For example, reducing the frequency of mowing would allow grass and vegetation to grow along the road shoulder. Alternatively, reseeding a vegetated area would affect the density and height of growth. A study by Caltrans found a substantial reduction in pollutants in road-side areas where sufficient vegetation was allowed to grow in (Caltrans, 2003b). Monitoring of mowing practices and/or reseeding could therefore provide a DOT with a simple, cost-effective means of managing stormwater in a localized area. Less frequent mowing would also save on labor costs.

---

[4] While there many examples of DOT maintenance and source control activities, for the purposes of this report, we focus on stormwater monitoring of street sweeping, catch basin clean-outs, de-icing/salt applications during adverse weather conditions, vegetation maintenance, and herbicide applications.

[5] Note that monitoring the volume and quality of swept material is not addressed in this report since it does not directly involve wet weather monitoring.

Evaluating the efficiency (meaning the ability to remove pollutants) of routine DOT maintenance activities requires a unique monitoring strategy. Maintenance activities include street sweeping, catch basin and ditch clean-outs, vegetation control, and use of de-icing materials[6]. One approach is to conduct before/after stormwater sampling to measure the effect of a particular maintenance activity on pollutant concentrations or loads. A good example of this is a study by Law et al. (2008). In this study, street sweeping was evaluated in a residential area of Baltimore. Using a before/after strategy, these authors found that increased sweeping frequency and coverage resulted in a modest reduction in total solids in stormwater runoff.

Another way to evaluate the efficiency of maintenance activities is to conduct paired control-treatment monitoring. "Control" means monitoring a reference area where no maintenance occurs, while "treatment" means monitoring another nearby area where maintenance does occur. To provide a level of experimental control, the treatment and control sites must have similar characteristics (save for the maintenance activity being evaluated) to exclude unwanted variability. With paired control-treatment monitoring, samples should ideally be collected from both sites concurrently (i.e. during the same storm) over multiple storm events using the same exact methods. A good example of this is a study by the United States Geological Survey (USGS) that examined the water quality impacts of herbicide use along a highway shoulder in Oregon (USGS, 2001). The purpose of this study was to characterize the use of herbicides at the treatment site and compare the data to a control site where there is no use of herbicides to assess whether the use of herbicides could be a contributor to the load of herbicides. Samples were collected from an unsprayed control site on one side of the highway and from a nearby sprayed treatment site on the other side of the highway. Through a comparison of the control and treatment data, the study authors were able to demonstrate that herbicide concentrations in runoff gradually subsided over a 2-week period following application.

One final DOT activity is the maintenance of structural BMPs. Regular and frequent structural BMP maintenance can potentially improve their functioning, although a monitoring study would be needed to demonstrate this. BMP monitoring strategies are covered under a separate goal (see Goal 2 above).

Additional Goals to Monitor

There are additional goals, less typical, for state DOTs to monitor. These include, for example, measuring the effects of wildfires, spills, natural disasters, climate change, and emergent pollutants on long-term runoff quality and quantity. Based on the literature review, these additional monitoring goals are not typical for DOTs. Therefore, monitoring protocols for measuring the water quality impacts of wildfires, spills, natural disasters, climate change, and emergent pollutants are not presented in this report.

[6] Note trash pick-up is not covered in this report.

CHAPTER 2 – IDENTIFY KEY MONITORING ELEMENTS ESSENTIAL FOR THE DEVELOPMENT OF AN EFFECTIVE DOT MONITORING PROGRAM

2. Introduction

DOT storm runoff consists of complex dynamics and therefore requires a clear understanding of the governing factors in order to develop effective monitoring protocols. Stormwater monitoring is inherently challenging and many monitoring studies fail to produce useful data because of poor study design and quality control issues (Law et al., 2008). Therefore, it is essential that a high level of scientific rigor be applied to every facet of a monitoring program to ensure that the monitoring is consistent with the DOT's goals, provides meaningful and valid results, and helps guide DOT management actions and priorities.

Based on the literature review, the research team identified the key monitoring elements that need to be considered to develop an effective DOT storm monitoring program. These key elements include: what type of data to collect, how to set up the monitoring station network, how to choose appropriate monitoring equipment and sampling methods, how to identify a qualifying storm event, how to implement QC measures, and how to select an appropriate data analysis technique to draw meaningful conclusions from the data. This chapter presents a detailed discussion of each of the key monitoring elements, including the recommended protocols for each element.

Before developing the monitoring program, it can be helpful to review and evaluate any existing or historical data (if available) for the local study area. This may include data collected during a pilot study; data from other water quality monitoring programs (e.g. by states, consultants, universities, or government agencies); data from TMDL reports and implementation plans; data from the International BMP database or the National Highway Runoff Database; and/or data collected from similar sites (USEPA, 2011). Often, the data may be of limited value or too sparse (or dated) to be meaningful. However, they can still provide some insight into what to expect, as well as help identify what type of study design would be most suitable by revealing water quality trends or important covariates (USEPA, 2011). Trends suggest there is some detectable change in water quality over time in one or more variables, which further monitoring may either validate or refute. Covariates are variables that vary together. For example, water quality constituent[7] concentrations can vary with season, location, traffic, or precipitation characteristics; or one pollutant may vary with another. Existing data also helps establish baseline conditions which serve as a point of comparison for new monitoring data. This is especially useful for before/after studies which have many applications (discussed in later sections).

2.1. Key Element - Health and Safety Plan

The health and safety of field personnel is a primary concern. It is common for monitoring to be performed during heavy storm conditions, at night, on slippery or uneven surfaces, near traffic, and sometimes in relatively remote areas. As such, it is imperative that project managers and field technicians conduct a

[7] For the purposes of this report, a "water quality constituent" is considered to be a compound in water that can be quantified, either as a concentration or other measure of magnitude. A water quality constituent is not necessarily a pollutant source, unless it is present in excess or is inherently toxic to aquatic biota. See also Footnote 1.

thorough investigation of each prospective monitoring station to identify any possible hazards before monitoring begins to prevent injuries and accidents from occurring. As a minimum, every field crew member should have personal protective equipment and a cellular telephone on them at all times to keep in touch with each other during storm events. All project staff should also know the protocols for what to do in the event of an accident or other emergency. Before any field monitoring occurs, a Health and Safety Plan (HSP) should be developed and approved by all parties involved, and there should be some oversight to make sure it is followed. The HSP identifies any physical and health hazards that can potentially harm field personnel, procedures to prevent accidents and maintain safe working conditions, and protocols to follow in case of an accident.

Some other safety considerations applicable to stormwater monitoring are listed below. This is a general list of some of the more common safety hazards and may not be applicable everywhere. Any additional site-specific safety concerns should be identified in the HSP. The list is as follows (FHWA, 2001; Caltrans 2015):

- confined space entry
- use of power tools
- proximity to high-speed traffic
- biological hazards (wildlife, plants)
- disease or infection vectors (e.g. ticks, mosquitoes, rodents, pathogens)
- insect stings (e.g. bees, wasps, hornets)
- poor visibility at night or during adverse weather
- heat and cold
- fall hazards and poor footing on slippery surfaces
- potential presence of explosive/toxic gases or hazardous materials
- fast moving stormwater
- opening/closing manholes
- lifting of heavy equipment such as full samplers or batteries
- strangers at the monitoring stations

For more information on health and safety issues, see the FHWA *Guidance Manual for Monitoring Highway Runoff Water Quality* (2001) or the *Caltrans Stormwater Monitoring Guidance Manual* (Caltrans, 2015).

2.2. Key Element – Monitoring Station Network

2.2.1 Location of Monitoring Stations

Monitoring locations are typically identified based on a systematic application of specific site selection criteria. The criteria are intended to filter out undesirable locations and ensure that the selected sites are suitable for the collection of samples and flow measurements. It is important to select monitoring locations that will provide representative data that address the monitoring goals. A non-representative location, on the other hand, would provide data of limited value due to bias, systematic error, and/or lack of relevance to the project objectives.

Before monitoring begins, it is helpful to review geographic information system (GIS) data or satellite imagery of the study area to generate a preliminary list of possible monitoring sites. Some examples of criteria to use at this stage include:

- Site is within DOT property boundaries (and the source of runoff is within DOT property)
- Site has an existing drainage system where monitoring can be conducted (unless nonpoint runoff will be targeted)
- Site has a contributing drainage area with the desired characteristics (e.g. AADT)
- Site is in a geographically convenient location to minimize driving time
- Site has the desired surrounding land use
- Site is not affected by erosion from extremely steep slopes or cut and fill areas
- Site discharges to TMDL receiving waters and DOT contribution is significant (for TMDL compliance monitoring only)

Much of this information can be obtained from computer mapping tools; publicly-available software; data clearing houses; or from DOT in-house resources. It is especially helpful to review GIS data of the storm drainage system (if available) to look for possible monitoring locations such as channels, ditches, swales, storm drain pipes, and outfalls.

Once a preliminary list of candidate sites has been developed, the next step is to conduct a feasibility survey in the field to verify site conditions and determine whether the location is suitable for monitoring. At this point, a more in-depth assessment is needed. A list of items to consider during the feasibility survey include:

- Can field personnel safely and easily access the site?
- Is the site suitable for measuring flow, collecting samples, and installing a precipitation gauge?
- Can the equipment be secured from theft and vandalism?
- Is there a source of electrical power or would reliance on solar power be required?
- Is there flooding potential at the site?
- Does the site have any debris, trash, obstructions, or contamination sources that could interfere with sampling and flow measurements?
- If in a coastal area, is there any tidal influence at the site?
- Are there any external influences at the site from outside the DOT right-of-way, such as runoff or other contributing factors from non-DOT controlled areas?
- Are there any insects or wildlife in the area that could infest or damage the monitoring equipment?
- What is the slope of the flow pathway? (Steep slopes should be avoided as they cause high velocity turbulent flows which are difficult to measure accurately.)
- Is the discharge at the site expected to be uniform (laminar), stable, and well-mixed in response to precipitation?
- Are there any bends or turns in the flow pathway which could cause the discharge to change direction suddenly and impede measurements and sampling? (For example, a 90-degree bend in a storm drain pipe could potentially cause runoff to flow up the sidewalls rather than along the invert).
- What method of flow conversion will be used to convert the measured level to discharge? Is the method appropriate for the site?

- What, if any, infrastructure will need to be constructed to conduct the monitoring? (For example, flow channels, test plots, conduits, concrete slabs, poles for mounting precipitation gauges or a solar array, earth anchors to secure equipment, and other constructed devices)

When conducting the feasibility survey, field personnel should complete a standardized form to record all site information in a consistent manner. In some cases, a ranking system or prioritizing scheme may be useful as some criteria will be more important than others. In general, it is important to conduct site feasibility surveys carefully and systematically. Taking extra time at the outset is better than hastily selecting a poor monitoring site which would not provide the data needed. Selecting suitable sites helps ensure that the monitoring proceeds smoothly and the data collected are representative and pertinent to the monitoring goals.

2.2.2 Number of Monitoring Stations

The number of sites to monitor is also an important consideration which depends primarily on the heterogeneity and scale of the study area. For small-scale targeted monitoring, the level of heterogeneity (e.g. in terms of road design, traffic, facility type, surrounding land use, climate) is likely to be relatively small. For example, if the goal is to characterize a single hotspot or a small stretch of a single highway, only a few monitoring stations are needed. At regional or state-wide scales, however, additional locations are needed to capture the geographic variability in precipitation and runoff characteristics across the study area. For example, Caltrans currently monitors over 100 locations throughout California to characterize runoff from their highway network. California is particularly heterogeneous in terms of climate, land use, precipitation patterns, traffic volumes, elevation, and terrain. As such, a more extensive monitoring network is needed. In relatively homogeneous states, however, highway runoff can be adequately characterized with fewer stations. For example, the Massachusetts DOT (MassDOT) monitored only 12 locations in a 2009 state-wide characterization study (USGS, 2009) while the North Carolina DOT monitored only 10 locations in a similar study in 2001 (NCDOT, 2001). Of course, there are other factors involved that dictate the number of monitoring stations needed such as regulatory requirements. Nonetheless, these examples provide a general idea of how many stations are needed to characterize runoff at different spatial scales. For BMP effectiveness monitoring, the minimum number of stations is 2, however this depends on the purpose and configuration of the BMP. As a rough guideline, 1 to 3 stations are recommended for small-scale monitoring (small highway sections, transportation facilities, and hotspots), 3 to 5 stations are recommended for medium-scale monitoring[8], and 10 to 50 stations are recommended for large-scale monitoring (state-wide characterization), again depending on the size and heterogeneity of the state.

2.3. Key Element – Water Quality Pollutants

Selecting appropriate water quality pollutants to monitor requires careful planning. There is no universal list of pollutants to monitor since it depends on the DOT's monitoring goals, data needs, regulatory requirements (NPDES permits, TMDLs), and available resources. If existing data are available, there may be some value in continuing to monitor the same parameters for trend analysis purposes, depending on the quality and size of the existing dataset. Some DOTs may also wish to monitor pollutants that are known to be treatable by stormwater BMPs for general stormwater management planning purposes. In

[8] The number of stations will vary depending on the diversity of highways to be monitored, as well as other factors such as ecoregions, land use, TMDLs, and traffic counts.

specialized cases where a DOT is involved in a superfund clean-up, toxic organic compounds may need to be monitored, depending on what is found at the clean-up site.

One basic recommendation is to always monitor for total suspended solids (TSS). Nationwide, TSS is the most common stormwater pollutant both by volume and by weight (King County Washington, 2016). In the highway environment, possible sources of sediment include rust and wear of vehicles, eroded soils, particles from pavement wear, routine maintenance activities, atmospheric deposition, and vehicles that carry solids from off-site (WSDOT, 2007). Importantly, TSS is often a carrier for other contaminants such as metals, bacteria, phosphorus, PCBs, and PAHs, particularly in the smaller size fractions which tend to be dominant in highway runoff (Geosyntec and Wright Water Engineers, 2009; Opher and Friedler, 2010). Some studies have found that most of the organic contaminants, metals, and PAHs found in urban runoff are associated with particulate matter such as TSS (Zawlocki et al. 1980, Bannerman et al. 1996, Corsi et al. 1999).

Some DOTs are required to analyze stormwater runoff for potential toxicity due to the presence of toxic compounds such as PCBs, pesticides, and/or metals. Toxicity itself is not considered a chemical pollutant but rather it is a measure of the effect of certain harmful chemicals on aquatic biota. To measure toxicity, a whole effluent toxicity (WET) test is performed in which live aquatic organisms (e.g. *Ceriodaphnia dubia*, a small invertebrate) are exposed to the sample water under controlled laboratory conditions to evaluate the effect of the sample on their survival. If rapid die-off occurs, it indicates that highly toxic contaminants are present in the source water, which may warrant further testing to identify which contaminant(s) in the sample are causing the toxicity (Caltrans, 2015). In certain cases, species-specific toxicity testing may be needed if stormwater runoff is known, or suspected to be, toxic to certain species. For example, in Washington State, previous studies have demonstrated that stormwater runoff is unusually lethal to adult Coho Salmon, which would warrant a specific WET test geared toward salmon species (Feist et al., 2017).

A biological water quality pollutant that is commonly monitored is bacteria. Fecal indicator bacteria are important to monitor in cases where the DOT's highways, facilities, and/or BMPs discharge to an impaired waterway with a bacteria TMDL, or is on the 303(d) list of impaired waters under the Clean Water Act (in which case a TMDL is pending), or is otherwise an area of special biological concern. There are two primary types of bacteria for monitoring and measurement purposes: pathogenic bacteria, which can adversely impact human health; and "indicator" bacteria (e.g. fecal coliforms, total coliforms, *E. coli*, and *Enterococcus* species), which provide an indirect measure of the likelihood that pathogens are present. Pathogens are rarely tested for directly because they are difficult to isolate and analyze in a laboratory. In contrast, indicator bacteria are relatively easy to measure and hence are included in many monitoring programs. If indicator bacteria are present at very high concentrations, it is more likely that pathogens are present in the sample, as well. In general, monitoring of indicator bacteria can be challenging and it is important to follow a carefully established protocol to achieve quality results. Careful handling of the sample is crucial and absolutely sterile conditions are needed to avoid contamination. Another disadvantage is that bacteria samples have a very short hold time (typically 6 to 8 hours), and therefore must be collected as grabs, not by automated sampler. Appendix C provides more information on the laboratory methods, including holding times.

The following sections provide recommendations for pollutants to monitor for each of the four DOT monitoring goals. This information was compiled from a review and synthesis of several literature sources, including:

- The Federal Highway Administration's *Guidance Manual for Monitoring Highway Runoff Water Quality* (FHWA, 2001)
- A DOT guidance manual entitled, *Caltrans Stormwater Monitoring Guidance Manual* (Caltrans, 2015)
- The American Association of State Highway and Transportation Officials' (AASHTO's) *State-of-the-Practice Report: Water Quality Monitoring* (AASHTO, 2013)
- A white paper by the East-West Gateway Coordinating Council entitled, *Highway Runoff and Water Quality Impacts*

The monitoring pollutants were selected conservatively, meaning that, if in doubt, a pollutant was excluded from the list in the interest of saving on laboratory costs, reducing the volume of sample needed, and streamlining the data analysis process. Note that the list of monitoring pollutants can potentially be modified. For example, if prior monitoring has consistently found that a given pollutant is not detected, or a pollutant is not yielding any useful information, it should be dropped from the list. In general, the DOT will need to weigh the costs and labor involved in the sampling against the expected value of the results.

2.3.1 Goal 1: Characterize Stormwater Runoff Quality

Table 1 lists the recommended water quality pollutants to characterize stormwater runoff quality. Again, these pollutants were identified and compiled from several literature sources as identified above.

Table 1. Recommended Pollutants to Monitor for Goal 1

Pollutant Category	Pollutants
Particulates	TSS
Nutrients – Nitrogen	TKN, NO3-2, NH3-N, Organic Nitrogen[1], TN[1]
Nutrients - Phosphorus	TP, PO4
Metals	Cadmium, Lead, Zinc, Copper, and Iron (total and dissolved)
Bacteria[H]	E. Coli, Fecal Coliform, and Enterococcus[5]
Organic Contaminants[H]	TPH, Oil and Grease
	PCBs (Total)
	PAHs (Total)
General/Conventional	Water Temperature[4], DO, Specific Conductivity, pH, Hardness, and TDS
Herbicides[H]	Varies based on chemical control utilized[2]
De-Icing[H]	Chlorides, TDS, Conductivity[3]

[1]note that these chemicals are calculated, not measured. Organic nitrogen = TKN – NH3, while TN = TKN + NO3-2.
[2]varies depending on which herbicides are being used by the DOT for vegetation control. Examples of common herbicides include glyphosate, diuron, and others.
[3]these three indicators are only applicable to cold weather states where de-icing materials are applied to roadways during winter storms.
[4]in cold weather states, air temperature is also sometimes measured (in addition to water temperature) to determine whether the measured precipitation was rain or snow
[5]for discharges to marine waters
[H] = suitable pollutant for hotspot monitoring (though not exlusively for hotspot monitoring). For example, TPH is often monitored at hotspots such as vehicle fueling stations.

2.3.2 Goal 2: Assess BMP Performance

Table 2 lists the recommended water quality pollutants to monitor for assessing BMP performance. These pollutants were identified and compiled from several literature sources. In many cases, it is also important to monitor volume reductions, especially for LID practices, filter strips, or other BMPs with a permeable bed. However, the focus of this section is on pollutant monitoring.

Table 2. Recommended Pollutants to Monitor for Goal 2

Pollutant Category	Pollutants
Particulates	TSS
Nutrients – Nitrogen	TKN, NO3-2, NH3-N, Organic Nitrogen[1], TN[1]
Nutrients - Phosphorus	TP, PO4
Metals	Cadmium, Copper, Lead, Zinc (total and dissolved)
General/Conventional	DO, temperature, pH, hardness, and TDS
Bacteria	Fecal Coliform and Enterococcus[2]
Organic Contaminants[H]	PAHs, Oil and Grease

[1]note that these pollutants are calculated, not measured. Organic nitrogen = TKN – NH3-N, while TN = TKN + NO3-2.
[H] = applicable mainly to BMP monitoring in hotspot areas, for example oil-grit separators at vehicle fueling stations at a DOT maintenance facility.
[2]for discharges to marine waters

2.3.3 Goal 3: TMDL Effectiveness Monitoring

The types of pollutants to monitor is dependent on the nature of the impairments in the local waterways. The most common TMDL pollutants include sediment, pathogens, nutrients, and metals (USEPA, 2017). Certain waterbodies, specifically tie to less typical pollutants such as biological integrity, turbidity, and fecal coliform. DOTs should monitor all pollutant(s) identified in EPA-approved TMDLs in which they are a stakeholder. It is recommended, however, to also include any known or suspected covariates associated with the TMDL pollutant. For example, the TMDL pollutant of concern may potentially vary with precipitation, other pollutants, flow rates, runoff volumes, and discharge velocities. The purpose of including covariates is to provide a better overall understanding of the contributing factors that are causing the TMDL pollutant to be discharged into the impaired receiving waters. This is particularly important if the monitoring shows that TMDL implementation has not been effective, in which case some knowledge of the covariates and their interactions with the TMDL pollutants may provide insight into why this might be the case. It can also help identify suitable management actions to correct or mitigate the problem. DOTs may also wish to monitor certain conventional parameters such as water temperature, pH, hardness, and/or DO.

2.3.4 Goal 4: Evaluate Efficiency of Routine Maintenance Activities

Table 3 lists the recommended water quality pollutants for evaluating the efficiency of routine maintenance activities. This information was compiled from several literature sources.

Table 3. Recommended Pollutants to Monitor for Goal 4

Maintenance Activity	Pollutants
Street Sweeping	TSS, TDS, Volatile Suspended Solids (VSS), particle size distribution, Metals (Cadmium, Copper, Lead, Zinc, and Iron)
Catch Basin Clean-outs	TSS, TDS, VSS, particle size distribution
Vegetation Control	Varies depending on what herbicide was used for vegetation control, hardness, water temperature, pH
Application of De-icing Materials	Chloride, conductivity, water temperature

2.4. Key Element - Duration of Monitoring

As a general guideline, monitoring for about two years in humid temperate regions of the country and two to three years in arid and semi-arid regions[9] is recommended (Law et al., 2008). This allows enough time to build a sufficiently large and representative dataset that captures the variations in precipitation and runoff characteristics from storms of different types and sizes, as well as any seasonal effects and inter-annual variations in precipitation. In arid regions, persistent drought conditions may warrant an even longer monitoring period. On the other hand, if the study period is unusually wet, or data collection is particularly efficient, a shorter monitoring period may be sufficient.

Note that this recommendation of 2 to 3 years is somewhat arbitrary and is subject to change as the project evolves. The actual length of the monitoring period will always be a function of how many qualifying storm events occur (see following section for discussion of how many storms to sample) and the objectives of the monitoring. However, for the purposes of this report, 2 to 3 years is considered a reasonable goal and may be used for general planning purposes.

To put this into context, Table 4 below provides a summary of 10 different DOT studies from throughout the U.S. which successfully generated meaningful storm data for the benefit of the DOT. Most of these studies were 2 to 3 years long, regardless of the monitoring goal and climatic region where the study took place. The table also shows how many storms were sampled during this period and how many stations were monitored, which gives a general idea of how much data can be collected over a 2 to 3-year time frame. The number of storms sampled varied widely but in most cases, was in the 10 to 30 range (see following section for number of storms to target). This is probably due to variations in climate, the frequency of sampling opportunities, and other factors. The number of monitoring locations also varied widely among the studies, anywhere from 2 locations up to 14 depending on the monitoring objectives. Overall, these studies support the recommendation that a 2 to 3-year monitoring period is a reasonable goal.

[9] For discussion purposes, "humid" states are considered to be those in the eastern half of the U.S. (roughly east of the Mississippi River) which on average receive at least 30 inches of precipitation per year based on NOAA's 1981 to 2010 climate normals. "Arid" states are considered to be those west of the Mississippi River which typically receive less than 30 inches of precipitation annually, except for portions of northern California and the Pacific Northwest. See a NOAA report by Arguez et al. (2012) for more information.

Table 4. Summary of Data Collected for Ten Different DOT Studies.

DOT Monitoring Goal	State DOT	Type of Monitoring	Number of Monitoring Stations	Total Length of Monitoring Period*	Number of Storms Sampled	Reference
1	Massachusetts DOT	General state-wide characterization of highway runoff	12**	2 years (Sept. 2005 to Sept. 2007)	16 to 18 at primary monitoring stations	(1)
1	North Carolina DOT	General state-wide characterization of highway runoff	10	1.6 years (May 1999 to December 2000)	20 to 27 depending on the monitoring station	(2)
1	Texas DOT	General highway characterization monitoring in Austin, Texas Area	3	1.7 years (Sept 1993 to May 1995)	49, 36, and 20 storms at the three monitoring stations	(3)
1	Michigan DOT	General highway characterization monitoring in three NPDES-regulated municipalities in Michigan (this study required by MDOT's permit)	3	2.3 years (June 1995 to Oct 1997)	3 storms at each of 3 stations	(4)
1, 2	California DOT (Caltrans)	General characterization monitoring of highways and transportation facilities in the high-elevation Lake Tahoe region. Some BMP performance monitoring of sand traps also conducted	6	2.7 years (August 2000 to April 2003)	30 to 38 depending on the monitoring station	(5)
2	Texas DOT	BMP monitoring to evaluate the effects of an existing Permeable Friction Course (PFC) overlay on highway stormwater quality in Texas	2 paired stations (one for a PFC highway and one for a non-PFC highway)	3.8 years (Feb 2004 to Dec 2007)	13 storms at two paired stations	(6)
2	Washington State DOT (WSDOT)	BMP monitoring of existing highway vegetated filter strips to satisfy WSDOT's NPDES permit requirements	14 paired influent-effluent stations at five vegetated filter strip BMPs	3 years*****	5 to 11 storms at the five vegetated filter strips (varies by station)	(7)
2	Caltrans	BMP performance monitoring of existing vegetated slopes (biofilters) adjacent to freeways in northern and southern California.	8***	1 year total, equivalent to 2 wet seasons (October to April of	6 to 22 storms at 8 stations***	(8)

DOT Monitoring Goal	State DOT	Type of Monitoring	Number of Monitoring Stations	Total Length of Monitoring Period*	Number of Storms Sampled	Reference
				2001-2002 and 2002-2003)		
4	Oregon DOT	Monitoring of water quality in streams near areas where herbicides were used by Oregon DOT to control roadside vegetation	4 stations total (1 station with triplicate test plots, 1 station with triplicate control plots, and 2 stations in the receiving water (upstream and downstream)	About 0.5 year, including periodic monitoring of 6 different herbicide applications from May 1999 to January 2000	8 total (3 simulated storms and 5 natural storms)	(9)
4	N/A****	Determining efficiency of street sweeping and its impact on reducing pollutant loads to surface waters	2	2.3 years (0.9 years pre-sweeping, 1.4 years post-sweeping	17 and 15 storms at the two stations before sweeping; 11 and 7 storms at two stations after sweeping	(10)

(1) Quality of Stormwater Runoff Discharged from Massachusetts Highways, 2005-07.
(2) Sampling and Testing of Stormwater Runoff from North Carolina Highways (2001).
(3) Characterization of Highway Runoff in the Austin, Texas Area
(4) Highway Stormwater Runoff Study. Prepared for the Michigan Department of Transportation. April 1998
(5) Caltrans Tahoe Highway Runoff Characterization and Sand Trap Effectiveness Studies – 2000 to 2003 Monitoring Report
(6) Effects of the Permeable Friction Course (PFC) on Highway Runoff.
(7) WSDOT NPDES Municipal Stormwater Permit BMP Effectiveness Monitoring Status Report (S7.C and S7.D) Water Years 2012 – 2015.
(8) Roadside Vegetated Treatment Sites (RVTS) Study. Caltrans. November 2003.
(9) Herbicide Use in the Management of Roadside Vegetation, Western Oregon, 1999-2000: Effects on the Water Quality of Nearby Streams
(10) Deriving Reliable Pollutant Removal Rates for Municipal Street Sweeping and Storm Drain Cleanout Programs in the Chesapeake Bay Basin. CWP, 2008)).

*Monitoring period is approximate. For some of the studies listed, different stations had different monitoring periods.

**The 12 monitoring stations included 4 primaries, 4 secondary, and 4 test stations. Most of the data were collected at the primary and secondary stations.

***Each monitoring "site" consisted of multiple biofilter strips in close proximity along the same highway. For clarity, each site was considered one monitoring station (for 8 total), and the average number of storms sampled at each biofilter strip (which ranged from 6 to 22) is listed.
****Study was not conducted by a DOT but rather the non-profit Center for Watershed Protection located in Ellicott City, Maryland.
*****While this study included only 3 years of BMP effectiveness monitoring, WSDOT is required by their permit to monitor for 5 years total. Their 2014 permit requires a maximum sampling effort of 35 sampling events in 5 years, or 7 storms per year on average.

Finally, it should be noted that it is possible to conduct too much monitoring. At a certain point, the collected dataset will be large enough that any further sampling would not yield any additional information and would be a drain on DOT resources. For this reason, monitoring beyond two to three years is not recommended.

2.5. Key Element – Number of Storms to Sample

Sampling of 10 storms per year is a fairly reasonable goal (Law et al., 2008). This would result in 20 to 30 storms over the suggested 2 to 3-year monitoring period, which is consistent with most of the DOT studies listed in Table 4. Of course, this assumes the DOT's permit does not require otherwise and a sufficient number of sampling opportunities actually occur. Again, 10 storms a year is only a broad guideline. Inevitably, sampling will include some false starts where the samples are discarded due either to limited sample volume, equipment malfunction, and/or errors in weather forecasting. Monitoring tends to proceed haphazardly due to adaptations to unforeseen circumstances and variations in precipitation patterns, especially in the early stages of a monitoring program. Thus, only a few storms may be sampled in one year and many more in the next, even if weather conditions are similar. The exact number of storms is less important than ensuring that the monitoring data are clean, representative, and scientifically valid. The sampled storms should ideally cover a broad range of storm types and sizes occurring at different times of the year (or at least different times within a given season), and under varying antecedent conditions.

Alternatively, there are statistical methods for estimating the number of storms to sample. One method, developed by Pitt and Parmer (1995), is called a power analysis. It calculates the number of samples needed to achieve a 95% confidence level with a power of 80%. A drawback is that this method requires a prior knowledge of the data variability, specifically the coefficient of variation[10] (COV). The COV can be estimated from existing data (if available), from prior experience, or from other local studies. Alternatively, the COV can be calculated directly from the monitoring data once a few storms have been sampled. This direct approach will likely yield a more representative COV so that sample size can be more accurately computed. For more information on the power analysis, see Pitt and Parmer (1995).

2.6. Key Element - Sampling Methods

In this report, the term "sampling" refers to the physical collection and chemical analysis of stormwater samples where the sample is collected in the field at a monitoring station and the analysis is performed at a certified laboratory. There are many ways to collect a sample depending on the type of data needed and the goals of the monitoring. The particular method to use should be considered carefully to make

[10] The coefficient of variation is the standard deviation divided by the mean. It indicates the variability of the data.

sure it is appropriate for the application. Some of the more common sampling methods are discussed below. Note that in-situ water quality monitoring, which consists of permanently mounting an electrical probe in the flow path to make continuous water quality measurements, is not discussed in this report. This is because in-situ monitoring is not considered "sampling" since there is no physical collection of the water for laboratory testing. Similarly, continuous measurements of precipitation using an automated precipitation gauge is also not considered sampling since, again, there is no physical collection of water by field personnel (the exception would be if the precipitation water was sampled and sent to a laboratory to evaluate precipitation chemistry). A discussion of various sampling methods follows.

2.6.1 Grab Sampling

Grab sampling means the sample is collected directly from the runoff flow path (either by hand or with an extension pole) at a single point in time during the storm event. This type of sampling is relatively easy to do and requires little in the way of equipment or training. A major disadvantage, however, is that it has a high labor cost since field personnel need to be at the monitoring station at the time of the storm. In addition, it is often difficult to sync grab sampling with the desired flow rate due to logistical constraints such as driving time, staff availability, and the timing and duration of the storm. It is particularly easy to miss the "first flush" of pollutants and/or the peak of the hydrograph when using grab sampling. Moreover, it can be potentially dangerous to collect a grab sample in inclement weather. Note, however, that grab sampling is mandatory for certain contaminants such as oil and grease, TPH, volatile organic compounds, low-level mercury, and bacteria, since these pollutants degrade or volatilize if left in an automated sampler (Caltrans, 2015).

The procedure for collecting a grab sample is as follows:

1) Wearing powder-free nitrile gloves, wade into the flow path with a clean unpreserved sample bottle in hand (from the lab) of an appropriate volume and material.

2) Fill the sample bottle at the centroid of the water column in the area of main flow (where the stormwater is well-mixed and most turbulent) upstream of where you are standing. Do not touch the bottom to avoid stirring up particles. Do not sample stagnant pools or eddies if present. If sampling for TPH and/or oil and grease, fill the bottle at the surface, not at the centroid, since these contaminants float.

3) Dump the first collected sample to rinse the container (unless sampling for TPH or oil and grease) and then fill the bottle again. Repeat until the bottle has been rinsed with site water three times.

4) After rinsing, collect the final sample and pour it into the sample bottle provided by the laboratory. If the sample bottle contains acid or another preservative, do not over-fill the bottle. Continue collecting samples with the original unpreserved bottle until all the laboratory bottles have been filled.

5) Carefully cap the sample bottles. Do not touch the inside of the bottle or the underside of the cap. Label the bottles, complete the laboratory's chain of custody (making sure the chain of custody form matches the sample labels exactly), and immediately pack the samples in a cooler with ice. Transport the samples to the laboratory.

Grab samples can also be collected with an extension pole if it is unsafe to wade into the flow path. It is important not to sample runoff that has touched the pole as this could introduce contamination. In cases where the flow is too shallow to submerge a bottle (e.g. sheet flow), a sandbag, dustpan, or other barrier device can be used to pond the water. For more information on techniques to sample sheet flow, see WSDOE (2010) or Caltrans (2015).

2.6.2 Automated Sampling

Automated sampling is conducted using a portable automated sampler to collect samples from the flow pathway rather than collecting them by hand. Automated samplers are installed at the monitoring station adjacent to the flow path at the beginning of the monitoring period. They have a built-in peristaltic pump which pumps a storm water sample from the flow path via a plastic suction tube. The samples are stored inside the sampler until field personnel arrive to collect them. Automated samplers may be programmed before-hand to collect samples in many different ways. Some samplers have refrigeration systems to keep the samples cool, typically at 4 °C.

There are several major advantages to automated sampling. The primary one is that field staff are not required to be on-site during the storm since the sampler can collect samples on its own based on pre-programmed instructions. Thus, storms occurring after-hours can potentially be sampled. This saves considerably on labor costs and expands the number of sampling opportunities. In addition, automated sampling is the easiest and most cost-effective way to develop an Event Mean Concentration (EMC) (see Section 2.9.4) which is the average flow-weighted concentration of a given pollutant during a storm. In most circumstances, EMCs provide the most useful means of characterizing stormwater quality (Geosyntec and Wright Water Engineers, 2009). Finally, as mentioned, automated samplers can be programmed in many different ways, which provides a great deal of flexibility in how and when the samples are collected. Disadvantages associated with automated sampling are 1) the samplers are expensive and time-consuming to set up; 2) the samplers require a power source; 3) the operation of the samplers requires a higher level of training; 4) the samplers must be maintained regularly; and 5) the samplers may potentially be a source of contamination if appropriate QC protocols are not followed. In some areas, vandalism or theft may also be a risk. Nonetheless, automated samplers are a very common means of collecting storm samples since the advantages usually outweigh the disadvantages. Again, note that automated sampling cannot be done for oil and grease, TPH, volatile organic compounds, low-level mercury, and bacteria, as noted above.

2.6.3 Composite Sampling

Composite samples consist of a series of small sample aliquots (sub-samples) that are mixed together into a single large sample. The aliquots are typically between 50 and 500 milliliters depending on laboratory volume requirements. The aliquots are collected at specified intervals during the storm runoff period, typically based on either flow rate, flow volume, or time (more on this below). The easiest way to collect a composite sample is by using an automated sampler with a single large container. While composite samples can also be collected by combining multiple grab samples (collected either by the automated sampler into discrete bottles, or by hand), this is not recommended as it is very labor-intensive and usually not feasible if monitoring more than a few sites.

A major advantage of composite sampling is that it reduces the number of samples to be analyzed for each storm event, which saves on laboratory costs. In addition, composite sampling with an automated sampler is the easiest way to develop EMCs. A disadvantage is that composite sampling does not provide information on the variability in runoff quality over the aliquot collection period. For example, if 12 aliquots are collected at various points in the storm and are combined into 3 different composites (e.g.

rising limb, peak, and falling limb), 75% of the variation in pollutant concentrations is lost. This is only a concern, however, if the DOT wishes to measure the within-storm variability in pollutant concentrations, which is not always necessary or feasible.

The USEPA defines four basic ways to collect composite samples depending on how and when the sample aliquots are collected. These are described as follows (USEPA, 1992):

1) Sample aliquots of equal volume are collected at fixed time intervals (also known as time-weighted sampling).
2) Sample aliquots of variable volume are collected at fixed time intervals. The sample aliquot volume varies in proportion to the flow volume.
3) Sample aliquots of variable volume are collected at fixed time intervals. The sample aliquot volume varies in proportion to the flow rate.
4) Sample aliquots of equal volume are collected at unequal time intervals. The time interval varies in proportion to flow volume.

The first method is referred to as “time-weighted” composite sampling because the sample aliquots are collected at specific times during the runoff period irrespective of flow conditions. This type of sampling is appropriate when the flow rate is relatively constant over time (i.e. does not vary more than ± 10% of the average flow rate), or when flow monitoring equipment is not available. The other composite sampling methods (2 through 4) are types of “flow-weighted” sampling and thus require having an in-situ flow sensor in place to measure the amount of flow present. Flow-weighted sampling is appropriate in cases where stormwater quality varies dramatically during a storm (which is usually the case) because a greater quantity of pollutants will be discharged during the high-flow periods (USEPA, 2010). For NPDES permit applicants, EPA requires that flow-weighted composites be taken for either the entire duration of storm discharge (if the storm is less than three hours long) or for the first three hours of storm discharge (if the storm is greater than three hours long) (40 CFR 122.21(g)(7)(ii)). As mentioned, flow-weighted composite sampling is one of the most common means of developing EMCs. Many DOT studies employ automated flow-weighted composite sampling.

The four types of composite samples are summarized in Table 5 below. For more information on composite sampling, see USEPA (1992), DeLeon and Lowe (2009), and USEPA (2010).

Table 5. Types of Composite Samples from USEPA (1992).

No.	Composite Category	Time Increment Between Sample Aliquots	Sample Aliquot Volume
1	Time-Weighted	Fixed	Fixed
2	Flow-weighted	Fixed	Varies in Proportion to Flow Volume
3	Flow-weighted	Fixed	Varies in Proportion to Flow Rate
4	Flow-weighted	Varies in Proportion to Flow Volume	Fixed

2.7. Key Element - Monitoring Equipment

There are many options on the market for automated monitoring equipment, including flow sensors, precipitation gauges, automated samplers, data loggers, and other devices. Before a monitoring program begins, it is important to carefully consider the options and make sure that the equipment selected are 1) suitable for addressing the monitoring goals; 2) have a high level of measurement accuracy; 3) can accommodate the range of flows expected; and 4) can be protected from vandalism, flooding, and exposure to the elements. Cost is an important factor, as well. It is crucial to consider not only the initial investment but also the labor costs associated with regular maintenance, testing, calibration, and trouble-shooting of the equipment over the life of the study.

The following is a brief discussion of the different types of equipment available. For more information on equipment considerations, see Geosyntec and Wright Water Engineers (2009).

2.7.1. Precipitation Measurement Equipment

There are various types of precipitation gauges to measure precipitation. The simplest and least expensive is a manual gauge which consists of a graduated cylinder with parallel sides and an open top. The precipitation falls into the cylinder and is stored inside. Field personnel then measure the depth of water in the cylinder using the graduated markings on the side and then empty the cylinder for the next measurement. Measurements can be made daily, weekly, monthly, or following each storm event. If long periods elapse between measurements, oil can be placed in the cylinder to retard evaporation. Although inexpensive and easy to use, manual gauges are not recommended. Measuring precipitation at such infrequent intervals is not useful for detailed analyses of stormwater quality.

In contrast, automated precipitation gauges provide a continuous record of precipitation without the need for field personnel to be on-site. Since these gauges are automated, they require a data logger to record the measurements. Measurements should be taken every 15 minutes or less. One of the most common types is the tipping bucket gauge, which consists of a cylinder, usually 8 to 12 inches in diameter, and a mechanical pivot device or "bucket" similar to a see-saw. When the bucket fills with a pre-calibrated precipitation amount (typically 0.01 inches of precipitation), the pivot tips to one side due to the weight of the water, which closes an electrical contact. A built-in sensor records each tip and sends the information to the data logger. In general, tipping gauges are useful for a wide range of applications and are reasonably accurate provided they are installed in an open area away from obstructions. In colder states where snowfall occurs, antifreeze (ethylene glycol or an ethylene glycol/methanol mixture) can be added to the cylinder to convert the snow to liquid, or the gauge may be heated. In some cases, air temperature sensors can be used to determine if the precipitation is rain or snow (Smith and Granato, 2010). A drawback of tipping buckets, however, is that they tend to under-estimate heavy precipitation because the tipping mechanism cannot keep up with the high precipitation rate. Tipping gauges also require annual calibration and need to be cleaned regularly. Other types of automated precipitation gauges include 1) weighing gauges, in which the weight of the precipitation in the gauge is recorded with a scale; 2) float gauges, in which the precipitation is stored in a chamber with a float that rises in proportion to the depth of precipitation; and 3) optical gauges, which use optical scattering to measure the amount of precipitation (the latter are rarely used).

Regardless of the type of precipitation gauge, or the type of precipitation being measured (rain, sleet, snow), there will inherently be some level of error in the measurements. For example, measurements made during very windy conditions where precipitation falls at an angle may underestimate precipitation because the precipitation gauge acts as an obstacle to the air flow and affects the wind field above the gauge opening (Kuligowski, 1997). If this is a concern, baffles may be set up around the gauge to minimize the impact of wind. Another source of error is poor siting of the precipitation gauge. Gauges should be in an open area away from trees, buildings, power lines, bridges, or other objects. A general rule of thumb is to make sure that the nearest object is at least twice the distance away as the object is tall. For example, if a nearby tree is 50' tall, the precipitation gauge should be located at least 100' away from the tree. Another potential source of error is wildlife. Precipitation gauges naturally provide attractive perches and feeding spots for raptors or other birds, and bird droppings, feathers, or seeds can clog the cylinder. Finally, if the precipitation gauge is not perfectly level, it can bias the readings. Due to these various possible sources of error, the precipitation gauge must be sited and installed carefully. In some instances, a second back-up precipitation gauge may be needed. For more information on precipitation monitoring, see Kuligowski (1997), EPA (2000), and WMO (2012).

2.7.2. Flow Measurement Equipment

Measuring the rate of flow provides basic information on runoff variability during the storm and allows for collecting flow-weighted samples and calculating total storm event volumes. There are many types of equipment for measuring flow. It is common to use some type of electronic sensor that is permanently installed in the flow pathway as it provides a continuous record of flow rather than a single measurement point in time. Typically, the flow sensor measures water level, not flow rate. The recorded water level is converted to a flow rate using a flow conversion method (e.g. weir equation, stage-discharge rating curve, Manning's Equation). It is common to use the level measurement and/or precipitation to trigger storm sampling. Some common types of level sensors are: 1) pressure transducers, which measure depth as a function of the hydrostatic water pressure exerted on the sensor; 2) bubblers, which measure depth in proportion to the pressure required to push a compressed air bubble through a small vinyl tube; 3) weirs or flumes, which control the flow of water such that the depth of water conveyed over the weir or flume has a known mathematical relationship with flow (see example of a flume set-up in Figure 2); 4) area-velocity sensors, which measure both stage and velocity which are then converted to flow based on a user-specified level to cross-sectional area relationship (i.e. flow = cross-sectional area x flow velocity); and 5) ultrasonic devices, which are non-invasive acoustical instruments mounted above the flow path (e.g. on the ceiling of a storm drain pipe) that measure water depth by emitting an ultrasonic pulse and then measuring the time it takes for the signal to bounce off the water and return to the sensor.

For more information on this topic, see Geosyntec and Wright Water Engineers (2009) and Caltrans (2015).

Figure 2. Example of a flume set-up to capture storm runoff discharged from a small HDPE outfall pipe. From King County (2013).

Note that the type of sensors described above focus on measuring channelized flows where the runoff is contained within a geometric shape such as a round pipe, ditch, or weir. With additional measures in place, however, water quality monitoring may also be conducted on more diffuse runoff or sheet flow. The basic approach is to build a structure or conveyance of some kind to channelize the flow, making it more conducive to measure and collect samples. Several studies have successfully monitored sheet flow in this fashion. For example, the Oregon DOT evaluated the performance of vegetated filter strips by using a metal sheet wall to confine the runoff within the experimental plot (Figure 3). The bottom wall of the plot served as the collection channel for the runoff. Poly(vinyl chloride) (PVC) pipes were buried just below the ground surface to convey runoff to the flow sensor for measurement (ODOT, 2016).

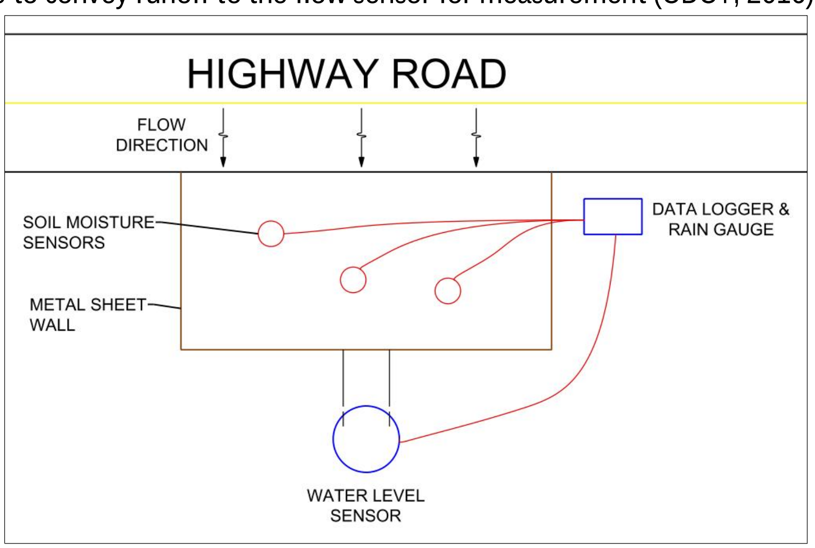

Figure 3. Example of Sheet Flow Monitoring Set-up.

In another study by the Washington State DOT, a high-density polyethylene (HDPE) pipe was notched along a portion of highway to allow sheet flow to enter the pipe (Figure 4). The sheet flow collector sloped downhill where a weir inside the pipe was used to concentrate and convey a sufficient volume of runoff for measurement and sampling purposes.

Figure 4. Sheet Flow Collector. From WSDOT (2014), page 37

If only discrete measurements of flow are needed (rather than continuous monitoring), a very simple way to quantify the flow rate is by measuring the time it takes to fill a bucket of a known volume (e.g. a 5-gallon bucket) using a stopwatch. It is recommended to make multiple measurements of the time and use the average value. The rate of flow can then be calculated as the volume of the bucket divided by the average time it takes to fill the bucket. Although this method can be quite accurate, it has several drawbacks, such as 1) field personnel need to be on-site when runoff is occurring to make the measurement; 2) it may not always be possible to direct the runoff into a bucket (e.g. if flows are very low or very high); and 3) it only provides a single discrete measurement of flow as mentioned. The bucket method is most useful for conducting limited research and for calibrating equipment (Geosyntec and Wright Water Engineers, 2009).

In some cases, it may be useful to have a second flow sensor nearby to provide back-up data in case the first sensor fails. If this is desired, both sensors should be of the same exact type to ensure consistency in measurement methods. Back-up sensors, however, require additional labor for routine maintenance and data processing.

2.7.3. Sample Collection Equipment

There are many types of equipment available to collect stormwater samples. In this report, we focus on two of the more common types: automated samplers and equipment for grab sampling. These are discussed in Sections 2.6.1 and 2.6.2, respectively.

2.7.4. Equipment to Set up Telemetry

Telemetry is the ability to remotely access the monitoring data, and in some cases programmable functions of automated equipment, through telephone modem communication. Cellular service is therefore required, as well as a PC vendor software package or cloud server to interface with the monitoring stations. There are two main types of telemetry transmission: one-way and two-way. One-way transmission means the data (e.g. water level, flow rate, precipitation, sample collection times) can only be transmitted from the monitoring station to a central location such as a cloud server or software package. In two-way transmission, however, the data can be relayed back and forth and the user has the ability to remotely program the sampler and change any settings as needed (USEPA, 2016). Although one-way telemetry is usually less expensive than two-way, it has limited applications since it can only be used to view and download data.

Two-way telemetry, on the other hand, is more versatile since it can be used to program the monitoring equipment for approaching storm events, view and retrieve data, and check the station's operational status (e.g. battery level). While two-way telemetry is more expensive to set up, it can greatly reduce the expertise and training necessary for field crews because many of the technical aspects of equipment set-up and shut-down can be conducted remotely by a system supervisor (Geosyntec and Wright Water Engineers, 2009). Some telemetry models can transmit data via line-of-sight, ultra high frequency/very high frequency (UHF/VHF), iridium modem receiver, or satellite radio. Baud rates can sometimes be selected by the user (Geosyntec and Wright Water Engineers, 2009). In climates where precipitation is extremely flashy and highly variable, telemetry can be very useful as, without it, it is almost impossible to predict a storm well enough that the field crew can visit the site the day before to prepare the samplers. In addition, for a large monitoring station network, telemetry eliminates the need to travel from station to station to access the monitoring data and programmable functions.

In addition, some specialized transmitters can send data through Geostationary Operational Environmental Satellites (GOES). GOES satellites have orbits that correspond with the earth's rotation, permitting each satellite to remain above a precise region. GOES satellites supply meteorologists and hydrologists with weather measurements such as rainfall and snowfall estimates, and track atmospheric patterns for severe weather conditions like tornadoes and hurricanes. GOES transmitters are normally owned and operated by government agencies and research organizations.

There are some disadvantages of telemetry, as well. Cellular coverage can be sporadic or unavailable at the monitoring stations, and any equipment malfunctions such as a clogged bubbler or a missed sample, are usually not conveyed by the telemetry unit. In general, while telemetry can be useful in many ways and increases productivity, it is still recommended to conduct routine maintenance site visits to verify that the equipment is working and are in optimal condition.

2.7.5. Equipment Power Needs

Automated samplers can be powered using an external 12V direct current (DC) battery, a solar array, or alternating current (AC) power. If using an external battery, a high-capacity deep-cycle battery is recommended. Some data loggers and certain flow measurement devices have their own internal batteries. It is often easier, however, to have a single external battery that powers all the equipment components to simplify battery maintenance and replacement. For non-refrigerated automated samplers, most of the power drawdown is due to the pumping action when the samples are being collected. Based on the manufacturer specified amperage draw of the pump, the expected power requirements for running a complete sampling program during a storm event can be calculated, although actual power use may vary somewhat depending on site conditions. For refrigerated samplers, power needs can be especially high as the sampler housing needs to be cooled and maintained at 4°C. This can be the most power-intensive component of the set-up.

Solar arrays are sometimes used to top off external batteries during dry sunny periods to help extend battery life. They tend to be more common in cases where the monitoring station(s) are remote, AC power is not available, and/or field personnel have limited availability to replace and maintain fully charged batteries. While solar arrays can be useful, they also have downsides. For example, they are relatively expensive and time-consuming to set up, and converters are needed to regulate the solar energy to a suitable voltage that is compatible with the automated sampler power requirements. They also make the monitoring station more visible which in some areas could promote vandalism. In addition, during storm events when it is cloudy and the equipment's power needs are most intensive, the solar array has no recharge capacity. It is important to weigh the additional cost and set-up time of using a solar array against its expected value to the project. In cases where refrigeration is needed, the power needed to maintain the samples at a cold temperature can potentially outstrip the ability of the solar array to recharge the batteries. Thus, as long as the refrigerator is active, the solar arrays are of little value.

2.7.6. Grab Sampling Equipment

Collecting a grab sample requires very little equipment. All that is needed is: a clean set of sample containers, powder-free nitrile gloves, an extension pole (if wading into the flow stream is not feasible), a cooler with ice to hold the samples, and a chain of custody and pen for transferring the samples to the laboratory. Clean sample containers can be obtained from the laboratory. Depending on which pollutants are being monitored, some of the sample containers may require acid preservatives. For example, if measuring nutrients, the sample container must contain dilute sulfuric acid. Similarly, for metals, nitric acid is required. If these containers are needed, the container cannot be immersed into the flow path since the acids will be washed out. In these cases, it is recommended to obtain a preservative-free container from the lab which can be immersed, and then fill the other acid-preserved containers with that one. Alternatively, field personnel can use an intermediate sample container but it must be washed thoroughly before-hand. In addition, certain contaminants such as TPH, oil and grease, and bacteria cannot be sampled with an intermediate container, as previously described (Caltrans, 2015).

2.8. Key Element – Storm Event Criteria

The following section provides some information that is useful to know when preparing to sample a forecasted storm. This includes the use and development of specific storm event criteria to identify what constitutes a "qualifying" storm event, as well as different types of weather forecasting services which aid in tracking storms.

2.8.1. Historic Weather Records

Before monitoring begins, it can be useful to compile and review local, long-term weather records. There are several benefits to doing this, including 1) long-term weather data provide some understanding of what to expect during the monitoring period; 2) the data help identify what constitutes a "small", "medium", and "large" storm for the study area; 3) it puts the monitoring data into a broader context; and 4) it can help the DOT track progress during the monitoring. As an example of the latter, if five storms have been sampled but all of them are relatively small for the local area, additional sampling may be warranted to capture a wider range of storm sizes. Of course, precipitation patterns can change over time, particularly due to the impacts of climate change. Nonetheless, reviewing weather records can still be a useful exercise and is not particularly labor-intensive.

A good source of weather data is NOAA's National Centers for Environmental Information (NCEI, formerly the National Climatic Data Center). The NCEI has historical precipitation data going back many decades for thousands of locations throughout the U.S. NOAA has compiled the data into climate "normals" which represent long-term average conditions for different regions of the country. Climate normals are updated every 10 years. The latest normals are for the time period 1981 to 2010. NCEI data can be obtained from the mapping tool at the following link:

https://www.ncdc.noaa.gov/cdo-web/datatools/findstation

Some examples of long-term weather records are provided in Figure 5 and Figure 6 based on the 1981 to 2010 climate normals. These graphs show precipitation characteristics for two locations in the U.S. with very different climates: Phoenix Sky Harbor International Airport in Arizona and Alabaster Shelby County Airport in Alabama. Figure 5 compares the average monthly precipitation at the two airports. The Phoenix airport is clearly drier overall and exhibits a stronger seasonal pattern, with a distinct dry season from April to June. In contrast, Alabama has a wetter climate overall but has less seasonal variability.

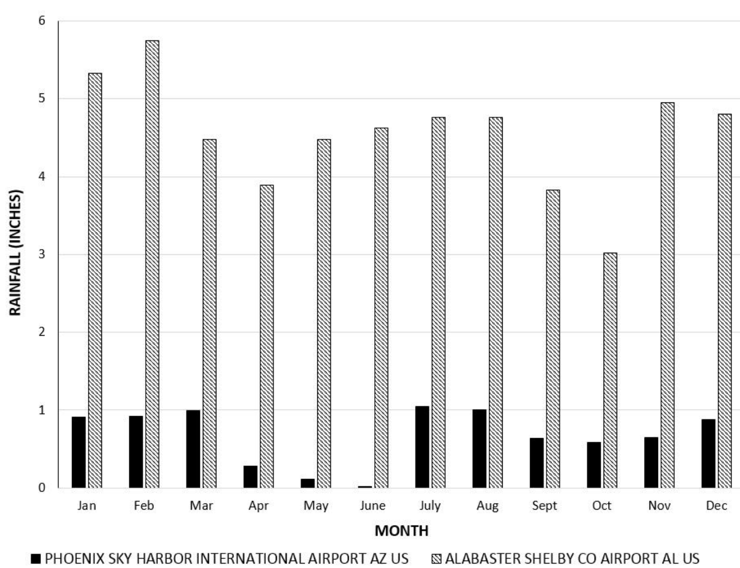

Figure 5. Average Monthly Precipitation for Phoenix Sky Harbor International Airport in Arizona (dark bars) versus Alabaster Shelby County Airport in Alabama (light bars). Data Source: NOAA NCEI, 1981 to 2010 Climate Normals.

Figure 6 below illustrates the differences in precipitation frequency at the two airports. Shown are the number of days per month with ≥ 0.01-inch precipitation, ≥ 0.1-inch precipitation, ≥ 0.5-inch precipitation, and ≥ 1 inch precipitation for both locations. At the Phoenix airport (Figure 6, Panel A), large storms were exceedingly rare. For example, in February, only 0.4 day had precipitation ≥ 0.5 inch and 0.1 day had precipitation ≥ 1.0 inch. In contrast, small to medium storms were relatively common, especially in the wet season (July through March). At the Alabama airport (Figure 6), Panel B; note the difference in the y-axis scale), large storms were more frequent. For example, in the same month (February), 4.2 days had precipitation ≥ 0.5 inch and 1.6 days had precipitation ≥ 1.0 inch, although most events were smaller.

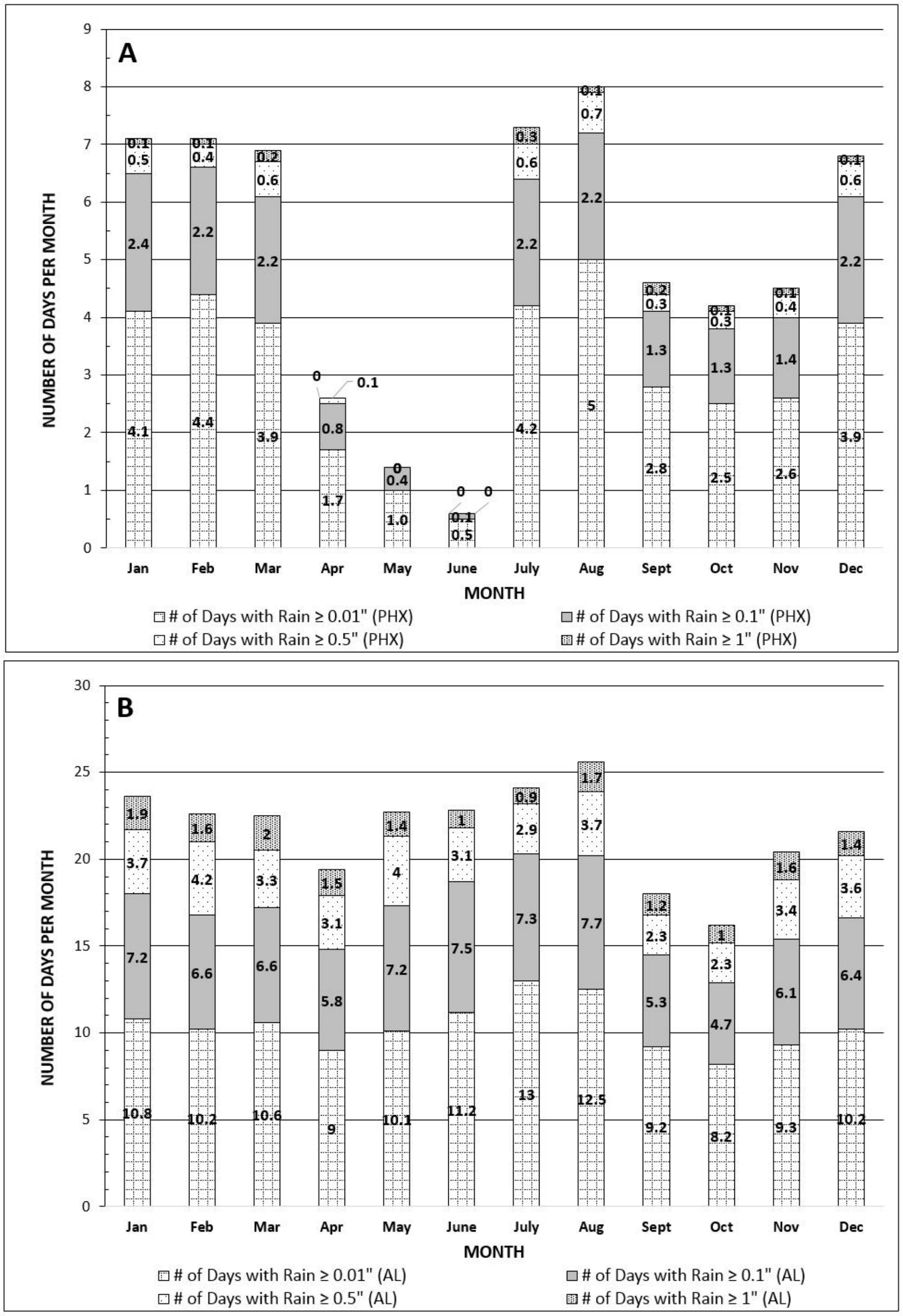

Figure 6. Monthly Precipitation Frequencies at Phoenix Sky Harbor International Airport in Arizona (Panel A – Top) versus Alabaster Shelby County Airport in Alabama (Panel B – Bottom). Data Source: NOAA NCEI, 1981 to 2010 Climate Normals.

2.8.2. Developing Storm Event Criteria

Storm event criteria are certain characteristics of a forecasted storm that determine whether or not it qualifies as a sampling opportunity. For NPDES permitting purposes, the USEPA has established criteria based on the expected size of the storm (total precipitation depth), the antecedent dry period (duration of dry weather leading up to the storm), and a variance criterion indicating whether or not the storm is expected to be representative of a “typical” runoff event. These criteria are as follows (from 40 CFR 122.21(g)(7)(ii)):

- The storm must have greater than 0.1-inch precipitation
- The storm must be preceded by at least 72 hours from the previous measurable (greater than 0.1-inch precipitation) storm event
- Where feasible, the variation in the duration of the event and the total precipitation of the event should not exceed 50 percent from the average or median precipitation event in that area

The purpose of these criteria is to ensure that enough flow is present to collect a sample, to allow some build-up of pollutants during the dry weather interval, and to ensure that the storm is representative of the most common conditions for the site (USEPA, 1992).

In some cases, however, these criteria may not be appropriate, and therefore site-specific storm event criteria can be developed (FHWA, 2001). This can be important for arid regions where, due to the relative infrequency of storms, the criteria may need to be more lax to minimize the risk of missing a rare sampling opportunity. The opposite is true in very wet areas which have such a high frequency of precipitation that meeting the antecedent dry period requirement can be challenging, resulting in many fewer sampling opportunities. In portions of the western U.S. with distinct wet and dry seasons, more than one set of criteria may be needed to reflect seasonal changes in precipitation patterns. For example, in King County, Washington, the criterion for the antecedent dry period is 72 hours during the dry season but only 24 hours during the wet season. This creates additional sampling opportunities during the wet season when storms are more frequent.

Additional storm event criteria include the duration of the storm and the inter-event period. The WSDOT and Caltrans require that a qualifying storm be at least one hour long with no fixed maximum. The inter-event period refers to the amount of time that must elapse before a new storm can begin. This criterion is useful for regions where precipitation tends to be intermittent, making it difficult to determine what constitutes an individual “event” for sampling and analysis purposes.

The FHWA recommends storm event criteria that are somewhat different from those of the USEPA described above. FHWA does not have a variance criterion and their antecedent dry period is shorter (24 hours, not 72). They also specify an inter-event period of 6 hours.

A summary of storm selection criteria used by various agencies, including four state DOTs, is presented in Table 6 below. The USEPA and FHWA criteria are also included for reference.

Table 6. Storm Event Criteria Used by Various Agencies

Criterion	USEPA	Washington State DOT (WSDOT)	King County, WA	California DOT (Caltrans)	Arizona DOT	North Carolina DOT	FHWA
Precipitation Depth	Min. of 0.1" ± 50% of avg. or median event	0.15" minimum	0.20" minimum, no fixed max.	Min. of 0.15" over 24-hour period	Min. of 0.1"	Min. of 0.1"	Min. of 0.1"
Precipitation Duration	Not specified	1-hr minimum	No fixed min or max	1-hr minimum	Not specified	Not specified	No fixed min or max
Antecedent Dry Period	At least 72 hours	< 0.04" in previous 6 hours	≤0.02" precipitation in previous 24-hrs (wet season), ≤0.02" precipitation in previous 72-hours (dry season)	< 0.04" in previous 6 hours	Not specified	At least 72 hours	At least 24 hours
Inter-Event Dry Period	Not specified	Not specified	6 hours	Not specified	72 hours	10 hours	6 hours

2.8.3. Tracking the Weather

Tracking the weather forecast regularly is a basic need of all stormwater monitoring programs. The reliability of the forecast is of key importance, but in many places, forecast accuracies are limited and subject to change even up to an hour before the impending storm. When tracking the weather, the predicted storm characteristics should align with the established storm event criteria discussed above. Even if the storm qualifies, however, it may not result in any actual viable samples if the forecast is poor. This is often the cause of "false start" samples. This happens when the storm was much smaller than predicted, in which case there was not enough runoff to meet the sample volume requirements, or the storm was much larger than expected, in which case the sample container filled too quickly. When programming an automated sampler, it is important to incorporate some flexibility to allow for uncertainty in the weather forecast. For example, the sample volume can be set a little higher than needed to ensure there is enough volume to run the laboratory tests.

Weather forecasts are available from multiple sources. This includes the National Oceanic and Atmospheric Administration's National Weather Service (NOAA-NWS) located at www.weather.gov, the Weather Channel (www.weather.com), Accuweather, or Weather Underground. These sources estimate the probability of precipitation (0 – 100%) as well as the amount of precipitation expected in inches. As a general rule of thumb, a 70% chance and at least 0.10" to 0.25" of precipitation predicted is considered adequate for sampling mobilization. In extremely arid climates, a lower threshold may be needed to avoid the risk of missing rarely occurring storms.

The amount of precipitation expected is especially important, yet this is often the part of the forecast that has the most uncertainty. This information is usually not available until a day or two before the predicted storm when the forecasting is relatively more accurate. NOAA-NWS reports a "quantitative precipitation forecast" or QPF which is an estimate of how much precipitation is expected for the local region, usually accompanied by some estimate of the storm's expected duration. However, QPF estimates are usually developed for relatively large geographic regions and may be too coarse to be useful for a specific monitoring location, or in cases where storms are small or isolated. Alternatively, NOAA-NWS publishes a real-time hourly weather forecasting graph (see weather.gov) that shows the amount and percent (%) chance of precipitation in the next 48-hours.

Regardless of what forecasting service is used, there will be some judgment and interpretation involved. This is where storm event criteria, flexibility in sample collection methods, availability of staff, and personal experience and knowledge of local weather patterns is useful.

2.9. Key Element – Data Analysis Technique

The ability to work efficiently with data and draw appropriate conclusions is crucial to any monitoring program. Ultimately, the outcome and usefulness of a study will only be as good as the ability of a data analyst to make sense of the collected data and draw meaningful conclusions using appropriate statistical techniques. When working with data, there are typically many steps involved. This includes validating the raw data, developing and populating a database, visualizing the data in graphs and tables, conducting statistical testing to check for trends and patterns, and finally drawing valid conclusions that address the goals of the monitoring. Only trained data analysts should attempt to analyze the data as it is very easy to draw erroneous or misleading conclusions if the data are not handled in the proper way.

2.9.1. Data Validation

Data validation is the process of reviewing the laboratory reports for any typographical errors, suspect results, data qualifiers, blank contamination, or any other problems encountered by the laboratory from the time the samples were delivered to the time when the laboratory issued the final report. It is important to evaluate the lab results as they are received to correct any potential mistakes in a timely manner. Some errors may simply be typographical errors, which are easily fixed, while others may be due to contamination, sample dilution issues, or other errors that are more difficult to correct. In these cases, it may be useful to have the laboratory re-run the sample in question to see if the result differs. However, this is only appropriate if the hold time for the sample has not expired (see Appendix C for hold times). Lab results for samples that are out of hold time are not valid.

Laboratory reports are usually standardized for a particular laboratory, but not necessarily across multiple laboratories. As such, lab reports have varying levels of detail. Some consist of hundreds of pages of detailed information, including instrument calibrations, narratives of sample processing and handling, a glossary of laboratory terms, definitions of data qualifiers, a copy of the submitted chain of custody (COC), QC data produced and evaluated internally by the laboratory, and any corrective actions taken. In contrast, other laboratories may only produce a short report with relatively few details, making data validation more difficult. Therefore, before monitoring begins, it is important to select a suitable laboratory that will provide the detailed information needed to properly validate the data.

As a basic review, it is recommended to check the laboratory reports for the following:

- Were the correct tests run as indicated on the COC?
- Were the correct detection limits achieved for each of the requested tests?
- Do the sample IDs and sample collection dates/times match with what was on the COC?
- Are the units correct for each result?
- Are the results within the expected range of values? If not, were there any unusual conditions at the time of sampling, or if not, could it be a laboratory error?

For more information on data validation and reviewing laboratory reports, see the *Caltrans Stormwater Monitoring Guidance Manual*, Chapter 13 (Caltrans, 2015).

2.9.2. Database Development

Data management refers to compiling all the data into a useable, organized master database so that the user can easily access records for data analysis or manipulation purposes. Database structures and content vary. It is generally useful to organize the data in relation to the monitoring goals. As an example, for a study of BMP performance, the data should be grouped as paired influent-effluent sample concentrations and loads. Similarly, if evaluating the effect of an independent variable (e.g. traffic volume) on runoff quality, the data should be grouped by traffic condition, for example high, medium, and low ADT data groups. The following list provides some examples of attributes that can potentially be included in a database, depending on the monitoring goals and the level of detail needed.

- Overall storm summary
    - # of storms successfully sampled
    - # of failed attempts to sample (false starts)

- Sampling activities
    - Date, location, number, and type (grab or composite) of samples collected
    - Season(s) in which sampling occurred
    - Portion of total storm volume captured by sampling (%) and/or portion of total storm duration captured by sampling (hours)*

- Precipitation characteristics
    - Total event precipitation (inches)
    - Peak precipitation intensity (inches/hour)
    - Duration of precipitation (hours)
    - Antecedent precipitation (# of hours or days leading up to the sampling event)

- Runoff characteristics
    - Total storm volume (ft^3)
    - Peak flow rate (ft^3/sec)
    - Flow and precipitation hydrograph data

- Water quality results
    - Sample concentrations from laboratory reports
    - QC sample results

Generally, the portion of the total storm volume captured is more important for flow-weighted samples, and some DOTs use percent capture to evaluate the validity of a composite sample. The portion of the total storm duration captured by the sampling is more important for time-weighted samples.

One example of an existing stormwater database is the National Highway Runoff (NHR) Database, developed by the U.S. Geological Survey in cooperation with the Federal Highway Administration. The structure of this database is a useful reference. The NHR Database is very large and currently (as of v 1.0.0a) includes 38 tables, fields for almost 200 water quality parameters, and data from many highway monitoring studies conducted throughout the coterminous U.S. An example of a table from the NHR Database is provided in Table 7 below summarizing some basic information about some monitoring stations. Shown from left to right are the location of each station (state, city, latitude/longitude), the average daily traffic (ADT) at each location, and the size (acres) and impervious fraction (0 to 1) of the contributing drainage area for each station.

Table 7. Example of a database table from the National Highway Runoff Database, v.1.0.0a.

Site_ID	State	tCountyCity	dLatitude	dLongitude	sADT	dDrainageArea	sImperviousFraction
91	WI	Milwaukee	42.9923	-88.0374	133900	4.56	0.95
92	WI	Milwaukee	42.9957	-88.0377	133900	5.51	0.628
93	TX	Travis, Austin	30.31	-97.75	58150	1.32	1
94	TX	Travis, Austin	30.217	-97.711	8780	0.1389	1
95	TX	Travis, Austin	30.41	-97.71	47240	25.847	0.37
96	TX	Travis, Austin	30.41	-97.71	47000	0.2619	1
97	TX	Travis, Austin	30.3794	-97.738	111000		1
98	TX	Travis, Austin	30.3794	-97.738	111000	3.212	0.52
99	WA	Near Auburn	47.3271	-122.2443	96000	0.051	1
100	WA	King County, Bo	47.77	-122.19	95000	28	
101	WA	Near Olympia, \	47.0453	-122.9913	42000	3.8	0.68
102	WA	Olympia, Washi	47.0339	-122.8901	130000	6	1
103	WA	Near Olympia, \	47.0325	-123.11	14000	3.4	0.78
104	WA				90000	8	
105	WA		47.0339	-122.8901	130000		
106	WA	Everett			160000		
107	WA	King County, Bo			90000	1	
108	WA	King County, Bo			90000		
109	WA	Mukilteo			20000	1.75	
110	WA	Mukilteo			20000	1.8	
111	CO	Grand Rapids	42.9562	-85.6715	120000	0.46373	1
112	MI	Ann Arbor	42.2816	-83.7836	41000	0.41781	1
113	MI	Flint	43.044	-83.676	51000	1.22	1
114	OH	Cincinnati	39.126	-84.534	150000	0.0675	1
115	MA	Ashburham	42.7025	-71.91433	3000	0.59667	1
116	MA	Littleton	42.50733	-71.487	39693	0.26349	1
117	MA	Boxborough	42.4725	-71.5555	81900	0.38056	1
118	MA	Lexington	42.43883	-71.25917	154500	0.72324	1
119	MA	Ashburham	42.69862	-71.90878	3000	0.31685	1

Another example of a large existing database is the International Stormwater BMP (ISB) Database, originally developed by the American Society of Civil Engineers and USEPA in the 1990s. The ISB Database now includes over 500 BMP studies (as of version 2016 11 17) from around the U.S. Some studies are related to highways, but not all. An example of a table from the ISB Database is provided in Table 8 below, which includes the results of various sampled storm events from 2008 to 2010. Shown from left to right are the monitoring station name, storm number, sample date and time, the water quality parameter (zinc in this case), the sample concentration and units, and the sample type (grab or flow-weighted composite EMC).

Table 8. Example of a database table. Source: International Stormwater BMP Database version 20161117 accessed at www.bmpdatabase.org

SITEID	SITENAME	Storm #	SAMPLEDATE	SAMPLETIME	WQX Parameter	WQ Analysis Valu	WQ Units	STCODEDescp
829589938	I-95 Plaza Bioretention Cell	15	9/21/2009	11:15 AM	Zinc	71	µg/L	Grab
829589938	I-95 Plaza Bioretention Cell	15	9/21/2009	1:00 PM	Zinc	44	µg/L	Grab
829589938	I-95 Plaza Bioretention Cell	15	9/21/2009	1:15 PM	Zinc	50	µg/L	Grab
829589938	I-95 Plaza Bioretention Cell	15	9/21/2009	12:45 PM	Zinc	30	µg/L	Grab
829589938	I-95 Plaza Bioretention Cell	16	10/29/2009	5:54 AM	Zinc	40	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	17	7/5/2010	3:32 PM	Zinc	77	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	18	7/11/2010	8:36 AM	Zinc	85	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	19	7/20/2010	5:24 PM	Zinc	59	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	20	7/24/2010	4:39 PM	Zinc	166	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	21	7/29/2010	12:26 PM	Zinc	49	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	22	8/20/2010	9:39 PM	Zinc	93	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	23	9/1/2010	7:11 AM	Zinc	77	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	24	9/13/2010	10:23 PM	Zinc	56	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	25	9/15/2010	9:31 AM	Zinc	32	µg/L	Flow Weighted Composite EMCs
829589938	I-95 Plaza Bioretention Cell	26	9/15/2010	5:31 PM	Zinc	46	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	1	9/12/2008	10:32 AM	Zinc	23	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	2	9/24/2008	7:17 AM	Zinc	28	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	4	10/15/2008	3:56 AM	Zinc	46	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	5	10/21/2008	6:31 PM	Zinc	43	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	6	11/5/2008	11:40 PM	Zinc	16	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	8	3/10/2009	7:55 AM	Zinc	29	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	9	3/24/2009	4:00 AM	Zinc	31	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	10	4/26/2009	11:00 PM	Zinc	37	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	11	5/8/2009	8:07 AM	Zinc	9	µg/L	Flow Weighted Composite EMCs
269023093	Highland View	12	5/15/2009	6:27 PM	Zinc	33	µg/L	Flow Weighted Composite EMCs

Links to the current versions of the NHR Database and ISB Database are provided below.

National Highway Runoff (NHR) Database:
https://pubs.usgs.gov/sir/2009/5269/disc_content_100a_web/ReadMe.htm

International Stormwater BMP (ISB) Database:
http://www.bmpdatabase.org/download-master.html

In general, when developing a database, it is important to follow standardized protocols so there is consistency among the data records. For example, all date fields should be in the same format (e.g. mm/dd/yyyy) and sample concentrations should retain the same number of significant digits as reported by the laboratory. For continuous variables, such as flow, the number of decimal places depends on the precision of the flow sensor. In addition, it can be useful to incorporate drop-down menus with look-up tables. Once the database is developed, an independent analyst should check for typing errors by comparing the data entries with the original source documents. For large, complex databases, it may not be feasible to check all the data, in which case a "spot check" may be performed for only a portion of the entries. Any errors should be flagged and corrected (GeoSyntec Consultants and Urban Water Resources Research Council of ASCE, 2002).

For more complex databases, records are often linked among multiple tables to reduce clutter and to group the data in some logical fashion. For example, the water quality results for a given storm may be in one table while the precipitation measurements and total runoff volumes for the same storm may be in another linked table. These links should be checked for errors, for example by ensuring that the start and end dates/times of the storm matches for all the measured parameters. This can be done using database queries to identify any dates or times that do not pair up, or visually by the analyst if there are a limited number of records. This process ensures the database is internally consistent. If any errors are encountered, the original source documents should be consulted (GeoSyntec Consultants and Urban Water Resources Research Council of ASCE, 2002).

2.9.3. Data Visualization

Data visualization means summarizing the data into graphs and tables to provide information on the characteristics of the data, help identify any trends or patterns in the data, and identify suitable data analysis methods. Graphs are a very efficient means of conveying information. According to Helsel and Hirsch (2002), computing statistical measures without looking at a plot is an invitation to misunderstanding the data. For these reasons, graphing is the focus of this section, although a brief discussion of summary tables is also presented. In general, there are many different types of graphs. However, only some of the more common ones are included here. For more information on graphing, see *The Visual Display of Quantitative Information* by Edward Tufte (Tufte, 1983) or *Visualizing Data* by William Cleveland (Cleveland, 1993).

Summary Tables

Summary tables present the monitoring results in a concise fashion. Unlike a graph, tables have the benefit of providing exact numerical values. It can be difficult to discern patterns in the data, however, without an accompanying graph. Summary tables can either provide a general overview of the data ranges and variability, or can be more detailed (e.g. storm-by-storm EMCs). In the case of overview summary tables, it is common to include some descriptive statistics to characterize the dataset. These include measures of location, measures of spread, and measures of skew, which are described as follows:

- Measures of Location – conveys the central tendency of the data, usually expressed as the mean, mode, median, or geometric mean (the latter two are preferable to reduce the influence of extreme outliers (Horwatich and Bannerman 2009)

- Measures of Spread – conveys the variability of the data, often expressed as the standard deviation, minimum/maximum, or coefficient of variation

- Measures of Skew – conveys the degree of asymmetry in the data distribution, expressed as the kurtosis or the coefficient of skewness.

Descriptive statistics should only be calculated if at least 30% of the pollutant concentrations are detected, i.e. above the laboratory's reporting limit. This can be done using a spreadsheet program or statistical software package. If less than 30% are detected (i.e. reported as "less than" values or "ND" for non-detects, an alternative procedure developed by Helsel and Cohn (1988) and Helsel (2005) is recommended. The procedure is based on using the detected concentrations to extrapolate the non-detected values. Briefly, the process is as follows:

1. Rank the detected values in descending order so that the highest value is 1, the second highest is 2, and so on.
2. Log-transform the detected values.

3. Plot the Z-statistic, which is a type of statistical distribution that is the inverse of a normal distribution.
4. Plot the log-transformed values on top of the Z curve.
5. Fit a regression line through the values and use the regression equation to extrapolate values for the non-detects.
6. Calculate any descriptive statistics desired for the complete data set, including the extrapolated values.

For more information on this method, see the studies noted above.

Some examples of summary tables are presented in Table 9, Table 10, and Table 11 below. Table 9 shows an example of TSS monitoring results for highway runoff. This is a general overview table that presents the TSS results by water year (October 1 to September 30) to show the annual variability in the data. The descriptive statistics include the minimum, median, maximum, mean, standard deviation, and number of samples (n). Data for multiple monitoring stations are presented to allow a comparison among sites. As it is an overview table, there is no information on individual storms.

Table 9. Example of an Overview Summary Table Showing TSS Results for Multiple Monitoring Stations.

Highway Monitoring Site	Minimum	Median	Maximum	Mean	Standard Deviation	n
I-95	20.0	50.0	75.2	45.3	20.3	10
RT 1	28.0	51.5	130.0	65.0	46.6	5
RT 27	40.5	75.2	100.5	70.8	19.9	8
RT 75	50.0	78.0	117.0	81.0	65.2	6

Table 10 is an example of a summary table representing a highway runoff study. This table shows individual storm EMCs for many different water quality pollutants. As only one monitoring station is included, no comparison can be made with other sites. This table provides a high level of detail and allows one to quickly discern variations in storm EMCs from one storm to the next. No descriptive statistics are included because the table is intended to show more detailed information on individual storms, not overall data characteristics.

Table 10. Example of a Detailed Summary Table of Individual Storm EMCs.

Date	Precipitation (mm)	Flow (L)	TSS (mg/L)	VSS (mg/L)	BOD (mg/L)	COD (mg/L)	N (mg/L)	TP (mg/L)	Zn (mg/L)	F. col. (mg/L)	F. Strep (mg/L)
10/17/15	5.3	760	220	50	5	45	0.3	0.2	0.1	N/A	N/A
11/5/15	5.1	2270	40	12	10	70	N/A	0.2	0.1	1360	4800
5/8/16	58	4400	30	10	5	20	1.6	0.05	0.2	1200	4400

Table 11 presents a summary of findings for different monitoring studies. This is a very broad overview table showing the general prevalence and magnitude of various pollutants detected in runoff samples collected throughout the state. Many different descriptive statistics are shown, including the mean, median, 25th and 75th percentiles, interquartile range, and others.

Table 11. Example of an Overview Summary Table Showing the Prevalence of Various Water Quality Pollutants

Parameter	Number of Site with Data	Average Percent Detected	Mean	Median	Min	Max	25th Percentile	75th Percentile	Std. Dev	Inter-quartile Range
Solids										
TSS (mg/L)	27	99.5%	118.9	93.0	2.7	294.6	60.4	190.7	82.5	130.3
VSS (mg/L)	5	100%	196.2	81.0	19.0	460.0	65.8	355.0	200.0	290.0
Metals										
Copper, total (µg/L)	29	98.3%	28.0	24.4	4.6	72.0	18.0	38.0	17.0	20.0
Mercury, total (µg/L)	1	100%	0.02	0.02	0.02	0.02	0.02	0.02	-	0
Zinc, total (µg/L)	29	99.2%	162	116	26.0	395	95	230	111	140
Nutrients										
Nitrate + Nitrite Nitrogen (mg/L)	6	100%	1.5	1.6	0.5	3.0	0.7	1.9	0.9	1.2

When preparing summary tables, it is important to avoid clutter. This can be done, for example, by dividing the data into subsets and developing multiple smaller tables rather than a single large table (for example, one table of results for urban monitoring stations and one table for rural stations). In addition, it is generally easier to view a table if the number of rows is larger than the number of columns. For more information on summarizing data in tables, see Helsel and Hirsch (2002), Chapter 1.

Bar Graphs

One common type of graph is a bar graph, which conveys the magnitude of one or more variables as a function of the height of a series of rectangles of equal width. Bar graphs can be used, for example, to show variations in storm EMCs or storm total precipitation. An example is provided in Figure 7 below. This figure shows the variation in storm total precipitation (y-axis) for a series of sampled storm events over a 3-year period (x-axis) at two monitoring stations (dark bars versus light bars). The bar graph quickly conveys which storms were small and which were large. In some cases, it may be helpful to plot multiple variables on the same bar graph, for example concentrations of copper, lead, and zinc, all on the same y-axis. If reporting concentrations, however, the y-axis scale should be appropriate for the variables being plotted. For example, metals, usually measured in μg/L, should not be on the same graph as nutrients, usually measured in mg/L. While a secondary y-axis can be added to avoid this problem, it generally makes the bar graph more difficult to read.

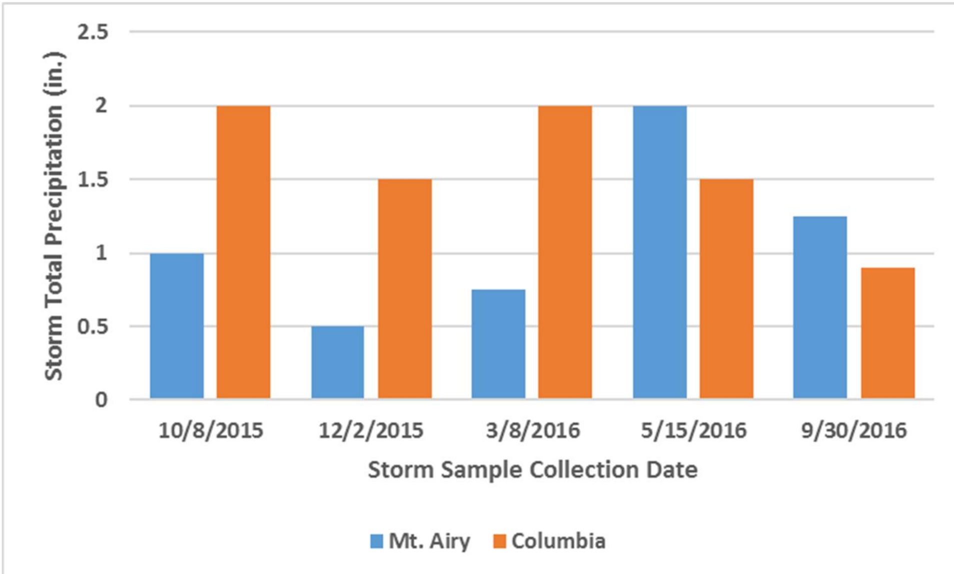

Figure 7: Example of a Bar Graph.

In some applications, a stacked bar graph may be more appropriate. Stacked bar graphs have rectangles that are plotted on top of each other rather than next to each other. A common use of such graphs is to show the relative magnitude or frequency of a variable that all the monitoring locations have in common. In Figure 8, a stacked bar graph illustrates different water quality pollutants that are related, for example the various chemical forms of nitrogen (NO_3-NO_2, TKN, NH_3, and organic N). In this example, each chemical component is assigned an individual bar, and the total sum of all the bars represents the total nitrogen concentration of the samples. This enables the viewer to quickly discern which forms of nitrogen were dominant in the samples.

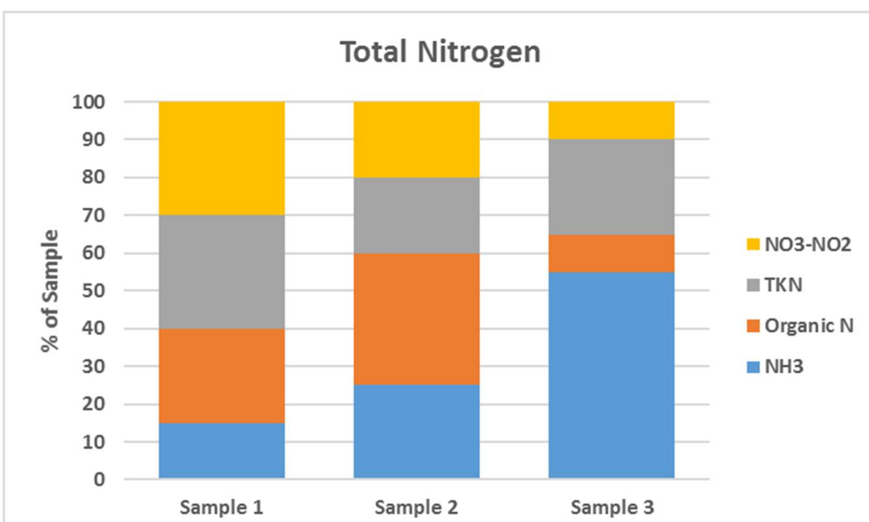

Figure 8. Example of a Stacked Bar Graph

Box Plots

A box plot (sometimes called a box and whisker plot) provides crucial information on the central tendency and spread of the monitoring data. Box plots are more complicated than some other graphs as they have many visual components. A standard box plot includes the following elements: the box itself, which represents the data values between the 25th and 75th percentiles (called Q1 and Q3, respectively); 2) a horizontal line through the center of the box representing the median; 3) a point symbol somewhere inside the box representing the average; 4) two vertical lines or "whiskers" extending from either end of the box which define the inter-quartile range (defined below); and 5) additional point symbols (if needed) outside the box representing outliers.

As mentioned, the height of the whiskers above and below the box represent the interquartile range or IQR, which can be calculated as follows:

- Bottom of whisker = Q1 – (1.5 * IQR)
- Top of whisker = Q3 + (1.5 * IQR])

If any of the data points extend beyond the top or bottom of the whiskers, they are considered outliers. Outliers are plotted as point symbols such as circles or asterisks. If there are many outliers, it suggests a skewed (non-normal) distribution. Sometimes, outliers are further classified into "outside values" and "far-out values" depending on how extreme the values are. See Helsel and Hirsch (2002) for more information.

An example of a box plot is shown in Figure 9. This figure shows total zinc concentrations in storm runoff at two monitoring stations.

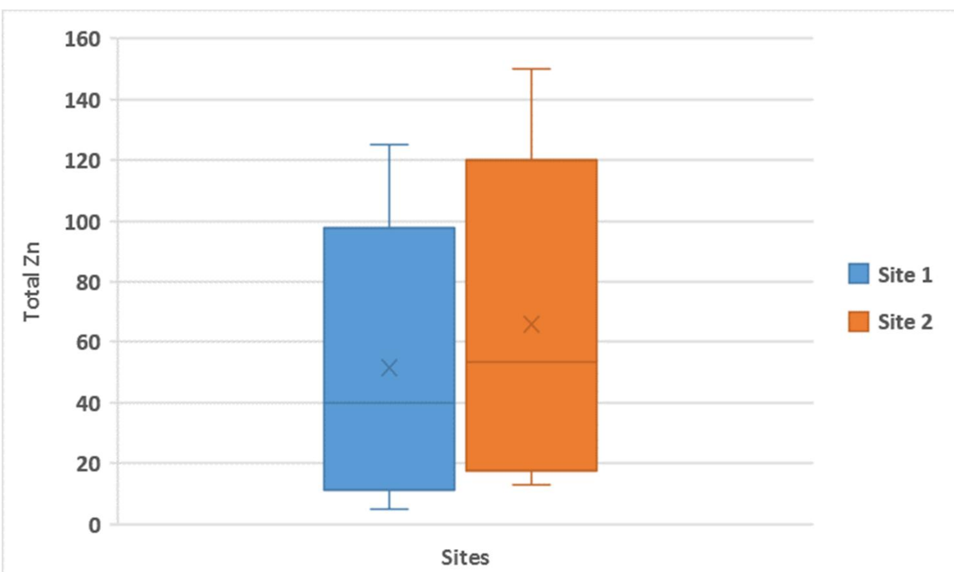

Figure 9: Example of a Box Plot.

Scatter Plots

Scatter plots are a basic tool of graphical investigation. They are a plot of a dependent variable on the y-axis versus an independent variable on the x-axis where each individual data point is represented as a point symbol (e.g. circle, square). Scatter plots quickly convey if there are any patterns to the data, or if they are randomly distributed. For example, a scatter plot might show that one variable increases while the other decreases (suggesting an inverse correlation; more on this below), or they might show that the data are completely unrelated (no correlation). If a pattern is known or suspected, a curve may be fit to the data points to illustrate the pattern. This is typically done as part of a correlation or regression analysis (discussed further below). An example of a scatterplot is shown in Figure 10. In this case, the fraction of the total mass of TSS in storm runoff (y-axis) is plotted against the amount of runoff (x-axis). The fraction appears to generally increase with the amount of runoff. To test if there is a statistically significant association between these two variables, a regression analysis would be needed. It appears unlikely, however, given the number of outliers.

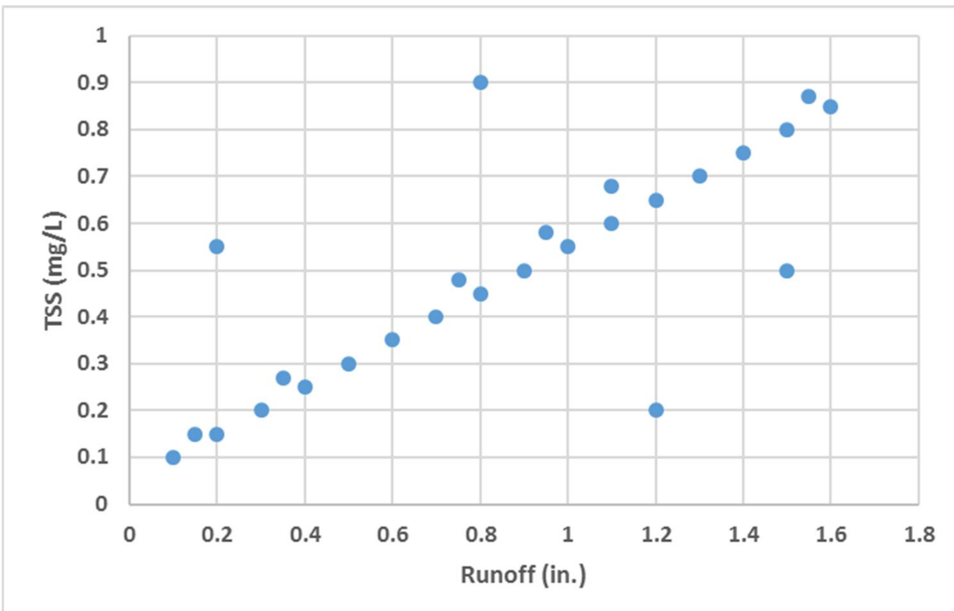

Figure 10: Example of a Scatterplot

Trend Plots (Time Series Plots)

Trend or time series plots show a dependent variable (e.g. water quality indicator concentrations) as a function of time. Trend plots are useful for showing if there are any significant changes in the data over a specified period of time, for example whether runoff quality is improving or declining. In some cases, it may be possible to extrapolate a trend into the future provided the data set is sufficiently large and the trend is strong enough to provide a meaningful prediction. An example of a trend plot is presented in Figure 11 showing an increasing trend in the concentrations of chloride and sodium in Pike Lake, Wisconsin from 1960 to 2000 (from Rose et al. 2004).

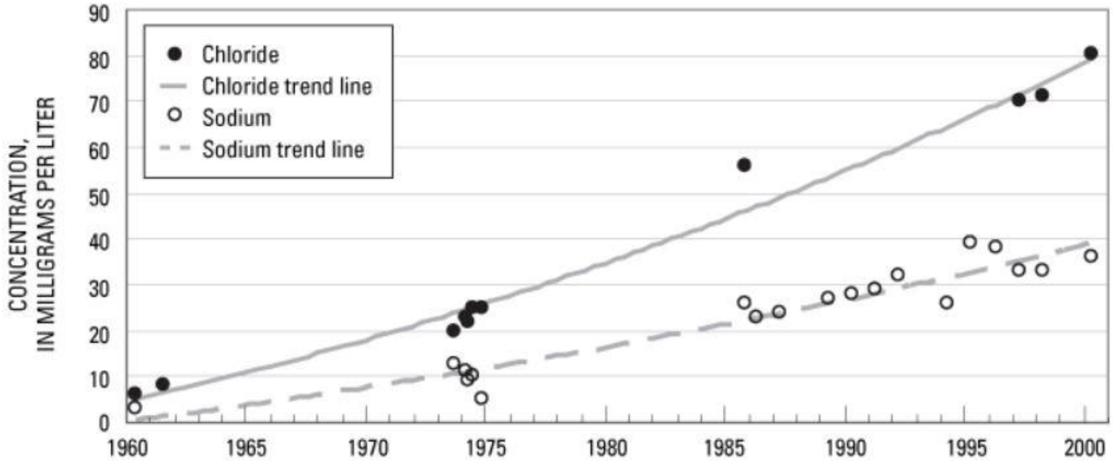

Figure 11: Example of Trend or Time Series Plot. Trends in Chloride and Sodium Concentrations in Pike Lake near Hartford, Wisconsin. From Rose et al. (2004)

Hydrographs

Hydrographs are a type of trend or time series plot. They are discussed separately because they are very common and are considered a fundamental tool for understanding the physical and chemical nature of storm runoff. Hydrographs show the variation in flow rate over a specified period of time, usually over a single storm event but sometimes over longer periods such as a season or year, depending on the goals of the monitoring. Hydrographs are often combined with other types of time series data. These may include, for example, 1) the hyetograph, a plot of precipitation over time; 2) a pollutograph, a plot of pollutant concentration or load over time; and 3) the sample collection times (assuming multiple discrete samples were collected). An example of a hydrograph is provided in Figure 12. Note the hyetograph is plotted on the secondary y-axis with the values in reverse order (low values on top, high values on bottom) for clarity.

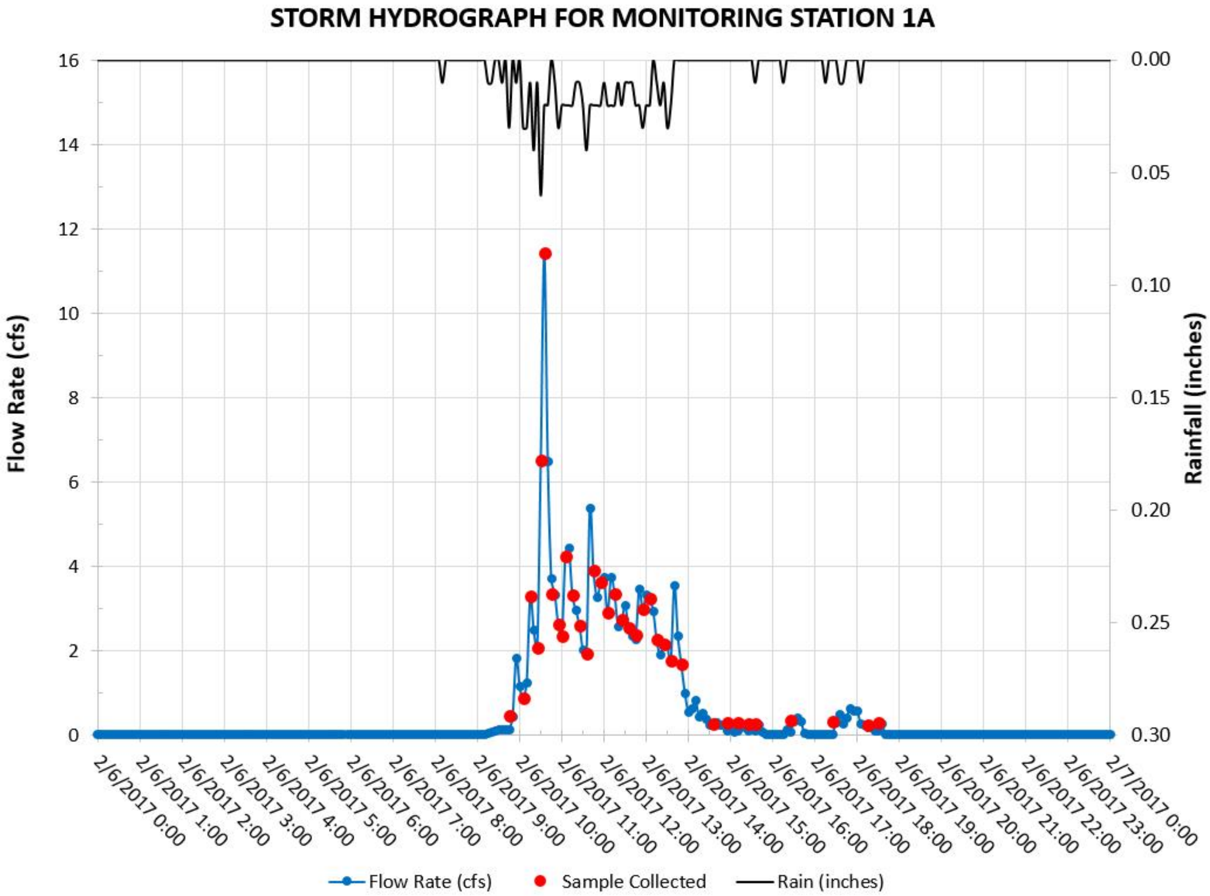

Figure 12: Example of a Storm Hydrograph.

2.9.4. Data Analysis Technique Per Goal

The basic goal of data analysis is to identify any patterns, trends, or effects in the data that provide insight into the physical and chemical processes of stormwater runoff while satisfying the goals of the monitoring. Data analysis is a large field and there are many different ways to analyze data. The various methods should be evaluated at the outset to determine the most suitable approach given the specific monitoring goals. In general, it is important to analyze the data in a scientifically valid manner as a misguided analysis can lead to erroneous or misleading conclusions and limit the value of the study to the DOT.

This section presents a general overview of some typically used analysis procedures for stormwater studies. Some of these procedures can be applied regardless of the DOT's monitoring goals, while others are more typically used for one particular goal or another based on general tendencies that have been observed in the literature. This does not mean that a given analysis technique should never be applied to other goals, only that the literature suggests that it is relatively more common for certain goals than others. Throughout this section, example studies are presented to illustrate how the data analysis methods have been applied by others. Note that the techniques presented here are not the only ones available and some have been omitted for the sake of clarity and brevity. There are many resources on data analysis methods where more detailed information may be obtained, including the following:

- The Federal Highway Administration's *Guidance Manual for Monitoring Highway Runoff Water Quality* – see especially Appendix A (FHWA, 2001)
- *The Stormwater Effects Handbook* – see especially Chapter 7 (Burton and Pitt, 2002)
- The USGS' *Statistical Methods in Water Resources* – see especially Chapter A3 (Helsel and Hirsch, 2002)
- A research report for the USEPA entitled, *Identification and Treatment of Emerging Contaminants in Wet Weather Flows* (Pitt et al., 2013)

Basic Statistical Concepts

The basic objective of statistical analysis is to infer characteristics of a large group of data (the "population") by analyzing the characteristics of a small sampling of that group (Zar, 1999). In stormwater studies, the population is the complete range of all possible precipitation and runoff conditions for the local study area, while the "sampling" represents all the monitoring data and measurements being gathered to represent those conditions. A central difficulty is that stormwater is highly variable over space and time. As a result, it is often unclear if the observed variation in measurements is random coincidence or if it is representative of the sampled population as a whole. The usefulness of statistics is that it provides a means of describing this variation objectively to separate the signal (a real trend or pattern in the data) from the noise (the "unexplained" component) (Reckhow et al., 1993).

The two major types of statistics are parametric and non-parametric. Parametric statistics assume the data are normally distributed (symmetric), have constant variance, and consist of independent observations. Overall, parametric tests tend to be more robust and diverse than non-parametric ones (Law et al., 2008) but have the downside of being very sensitive to outliers and are difficult to use for censored data (i.e. pollutant concentrations reported as "less than" or non-detected values). In contrast, non-parametric statistics make no assumption about the data distribution and generally work better for data gaps, censored data, and outliers (Mozejko, 2012). Some examples of parametric statistical tests include the t-test (both paired and two-sample), analysis of variance (ANOVA), and linear regression. Examples of non-parametric tests include the Mann-Whitney, Signed Rank, Kruskal-Wallis, and Friedman's tests. Statistical testing will be discussed later in this section. For general statistical background, see any statistics textbook, for example *Biostatistical Analysis* by Zar (1999).

Data Analysis Methods Common to All Monitoring Goals

The following data analysis methods have many different applications and, as such, are not considered unique to any DOT monitoring goal. These techniques are a part of basic exploratory analysis and serve as a starting point for the more in-depth analyses to follow.

Testing Data for Normality

A good starting point for any type of data analysis is to determine whether or not the sampled population follows a normal probability distribution. This is important because the applicability of most statistical tests is dependent on the type of distribution. Generally, it is preferable to have normally distributed data as this allows a broad range of parametric statistical tests to be used. If this is not the case, the data can be transformed to approximate a normal curve so that the assumptions of parametric testing can be met. The most common type of transformation is a log10 transformation (Pitt, 2013).

There are various statistical tests that can be used to test for data normality. These include the Kolmogorov-Smirnov test, the D'Agostino-Pearson K^2 test, the Shapiro-Wilk test, or the chi-square (χ^2) goodness-of-fit test. Here, the latter is recommended for its high level of power and relative ease of calculation (most spreadsheet programs and statistical software packages can compute the chi-square statistic). This test assumes that the observations are independent, that the dependent variable is numerical (e.g. pollutant concentration or load), and that the independent variable is qualitative or categorical (e.g. land use) and has two or more values (e.g. urban land use and rural land use).

To run the chi-square goodness-of-fit test, the chi-square statistic (χ^2) must first be calculated, as follows:

$$\chi^2 = \sum_{1}^{i} \frac{(O_i - E_i)^2}{E_i}$$

Where
i = total number of observations
O_i = observed value for i^{th} observation
E_i = expected value for i^{th} observation

Chi-square values range from 0 to 1. Relatively low values indicate that the observed values are close to the expected values, which supports the null hypothesis, while relatively high values indicate the opposite. To determine statistical significance, compare the calculated value of chi-square (from equation above) to the critical value of chi-square from a statistical textbook. Critical values vary with the degrees of freedom (v) and the acceptable level of error (traditionally 95% confidence level or $p < 0.05$). The degrees of freedom equal the number of category values minus one. In the above land use example, there is one category (land use) and two values (urban and rural), hence v = 1. If the critical value is greater than the calculated value, the null hypothesis cannot be rejected, meaning the sampled population is normally distributed. If the critical value is less than the calculated value, the null hypothesis is rejected, in which case the alternative hypothesis holds true (with 95% confidence) and the sampled population can be assumed to be not normally distributed.

Calculating Event Mean Concentrations (EMCs)

Event Mean Concentrations or EMCs are a basic measure of storm runoff quality and often provide the most useful means of quantifying the magnitude of pollution resulting from a runoff event (Geosyntec and Wright Water Engineers, 2009). Many types of monitoring programs include development of EMCs, for example highway runoff characterization studies (Goal 1), BMP effectiveness studies (Goal 2), and others. EMCs are also a common input into water quality models for TMDL planning purposes (Goal 3) and are a primary focus of existing databases such as the ISB database. When combined with flow data, EMCs can be used to derive pollutant mass loads for one or more storm events.

EMCs are defined as the flow-weighted average concentration of a given pollutant during a storm event. They are often determined from analysis of multiple grab samples collected at different times during a storm event. In this case, the calculation is as follows:

$$EMC = \frac{M}{V} = \frac{\int_0^t C_t Q_t dt}{\int_0^t Q_t dt} \cong \frac{\sum C_t Q_t \Delta t}{\sum Q_t \Delta t}$$

Where:

M = pollutant mass
V = volume of runoff
C_t = concentration of sample collected at time t
Q_t = discharge rate or volume at time t
Δt = time period between samples

If flow-weighted composite samples are collected instead, the analytical result from the laboratory is equivalent to the EMC and no computations are necessary. This is often a more cost-effective and representative way of estimating EMCs relative to grab sampling.

EMCs can also be developed as flow-weighted average EMCs based on event runoff volume for a series of events over a specified time period (Geosyntec and Wright Water Engineers, 2009). Flow-weighted average EMCs are computed as follows (from WSDOE, 2009):

$$EMC_{Avg} = \frac{\sum_1^n (EMC_i \, xF_{Avg})}{\sum_1^n F_{Avg}}$$

Where:

EMC_{Avg} = Flow-weighted average EMC
EMC_i = EMC for event i
F_{Avg} = mean storm flow rate for event i, obtained as total storm volume divided by duration of storm event
n = number of events
i = 1 to n events

Calculating Pollutant Loads

Pollutant loads are a basic means of characterizing storm runoff and represent a mass quantity (typically lb or kg) of a water quality constituent. Loading rates have various uses. For example, they are a common means of evaluating the overall water quality impact of storm runoff on receiving waters. In some cases, they can be as important to formulating stormwater management strategies as EMCs (Pitt et al., 2013). Loads may also be used to assess BMP performance (see Section 2.9.4.2) or to calibrate a water quality model for pollutant load analyses. Calculating pollutant loads requires that the EMC and total runoff volume are known. Loads can be calculated over different time periods, for example for an individual storm event or annually. The following equation may be used, in this case for an individual storm event:

$$L_s = EMC \ x \ V$$

Where:

L_s = load per storm event (lb.)
EMC = event mean concentration (lb./L)
V = volume of runoff for storm event (L)

A unit pollutant load may also be calculated by scaling the load to the contributing drainage area of the monitoring station. The calculation is as follows:

$$L_u = \frac{(EMC \ x \ V_r)}{DA}$$

Where:

L_u = unit load per storm event (lb./acre)
EMC = event mean concentration (lb./L)
V_r = volume of runoff for storm event (L)
DA = contributing drainage area (acres)

Water Quality Modeling

One final analysis method that is widely applicable is the use of water quality models to characterize runoff and receiving water impacts. Models can provide planning-level estimates of pollutant concentrations (EMCs), loads, and flow peaks and volumes. Some examples include the Stochastic Empirical Loading and Dilution Model (SELDM), Hydrologic Simulation Program-Fortran (HSPF), and the Driscoll Model which have been used and tested by many DOTs. As it is not possible to monitor everywhere, the use of a model greatly extends the ability of a DOT to understand the spatial variation in runoff from their highways and transportation infrastructure. It is important to always populate/calibrate a water quality model with empirically collected monitoring data to ensure that the model predictions are realistic. If water quality conditions change, or the model results appear unrealistic, additional targeted monitoring may be done in certain areas for re-calibration purposes. Ultimately, models are a useful and cost-effective analytical tool, and for certain applications are the only tool available. As this report focuses on monitoring protocols, not water quality models, a detailed discussion of modeling is not presented.

2.9.4.1 Goal 1: Characterize Stormwater Runoff Quality

This section describes some of the more common data analysis methods for Goal 1. Again, these methods are not exclusive to this goal but were commonly observed in the literature on characterizing stormwater runoff quality.

Comparative Analysis

One common type of data analysis in stormwater runoff characterization studies is to compare data between two or more data groups. The data groups should be established in some logical way, for example the urban monitoring locations can be one group while the rural monitoring locations can be the other group. Data grouped by season is also common, which allows, for example, wet season data to be compared to dry season data. If only two data groups are being compared, the parametric Student's t-test (unpaired) is suitable. This test evaluates if there are any significant differences in the sample means between the two data groups. Again, as a parametric test, normality, equal variance, and independent observations are assumed. If the data are skewed, the non-parametric Mann-Whitney test is more appropriate.

If comparing more than two data groups, the parametric analysis of variance (ANOVA) test is frequently used. There are different types of ANOVA tests depending on how many independent or qualitative variables are being tested. A one-way ANOVA tests a single independent variable (e.g. traffic volume), a two-way ANOVA tests two independent variables (e.g. traffic volume and highway design), and so on. Non-parametric equivalents include the Kruskal-Wallis and Friedman's tests. Table 12 below summarizes the different types of statistical tests to use for comparative analyses.

Table 12. Examples of Statistical Tests to Use for Comparative Analysis. Observations Must be Independent.

Data Distribution	No. of Data Groups Being Compared	Variance of Data Groups	Category of Statistical Test	Statistical Test to Use
Normal	2	equal	parametric	Student's t-test (unpaired)
Normal	>2	equal	parametric	ANOVA (one-way, two-way, or three-way)
Non-normal or no distribution	2	unequal	non-parametric	Mann-Whitney
Non-normal or no distribution	>2	unequal	non-parametric	Kruskal-Wallis, Friedman's test

Regression or Correlation Analysis

A regression or correlation analysis is a way of determining if there is a relationship between two quantitative variables: an independent variable (x) and a dependent variable (y). This analysis is included under Goal 1 because it is frequently used as part of an exploratory data analysis to identify covariates and determine which factors correlate with the observed pollutant concentrations or loads.

A common way to test for a correlation is through graphical analysis of a scatter plot. In this method, the analyst develops a line-of-best-fit or regression curve for the data points and evaluates the strength of the regression using certain metrics. One very common metric is the coefficient of determination or R^2 value (ranging from 0 to 1) which is a statistical measure of how close the observations are to the fitted curve. Relatively high values of R^2 indicate a strong correlation between the two variables while low values indicate a weak correlation or none at all. An inverse correlation means that one variable increases while the other decreases, whereas a positive correlation means that both variables either increase or decrease together. The numerical relationship between the two variables is defined by the regression equation, typically shown on the graph accompanied by the R^2 value or a similar metric. Sometimes, the p-value is also shown indicating the statistical significance of the regression model. Using the regression equation, the value of y can be estimated for any value of x.

An alternative (though less common) way of doing a regression analysis is to develop a correlation matrix or table. This can be done using various types of statistical tests, for example the parametric Pearson correlation test or the non-parametric Spearman correlation test. These tests can be run using statistical software packages. The outcome of the test is a series of correlation coefficients (ranging from 0 to 1) which indicate the strength of the relationship between the independent and dependent variables. Higher values indicate a stronger correlation. As a rough guideline, values at or above 0.75 are considered high (Pitt, 2013).

2.9.4.2 Goal 2: Assess BMP Performance

Data analysis techniques for assessing BMP performance are somewhat different than those for runoff characterization because the monitoring data are typically in the form of influent-effluent pairs. As a result, the observations are not independent since the runoff discharging from the BMP is a function of the runoff that entered the BMP, as well as processes occurring within the BMP itself. Parametric tests for paired influent-effluent data include ANOVA or the paired t-test. Non-parametric tests include the Signed Rank test, Friedman's test, Wilcoxon test, or Kruskal-Wallis test. As an example, the Signed Rank test is described below in detail. For information on the other tests, refer to a statistical textbook or the sources cited in Section 2.9.4 above.

The Signed Rank test is based only on whether there is a positive or negative difference between the data pairs irrespective of the magnitude of the differences. This makes the test very resistant to outliers[11]. An example is provided below showing a paired comparison of influent-effluent storm EMCs of NO_3-NO_2 for 10 sampled events. The purpose of this comparison is to determine if the effluent EMCs are significantly lower than the influent EMCs, which in this case is being used to assess the BMP's effectiveness.

1. Make a list of all ten (10) paired influent-effluent EMCs for NO_3-NO_2.

2. If needed, remove all cases where there is no difference between the paired NO_3-NO_2 values. The Sign test is based only on non-zero differences.

3. Assign a plus ("+") if the effluent EMC is higher than the influent EMC and a minus ("- ") if it is lower.

4. Count the number of "+" values. In this example, assume there are nine of them, meaning in most cases the effluent NO_3-NO_2 EMCs were higher. The total number of "+" values is called S^+. In this example, $S^+ = 9$.

5. Look up the p-value for $S^+ = 9$ from a standard statistical table[12]. In this example, the p-value is 0.011. Since this is less than the p-value of 0.05 (i.e. 95% confidence level), the conclusion is that the effluent EMCs for NO_3-NO_2 are significantly greater than the influent EMCs for NO_3-NO_2. This suggests that the BMP is not effective at removing NO_3-NO_2.

Table 13 below summarizes the statistical tests to use for BMP performance studies. Again, these tests require that the observations are paired, not independent.

Table 13. Examples of Statistical Tests to Use for Paired Observations of BMP Influent and Effluent

Data Distribution	No. of Data Groups Being Compared	Variance for Different Data Groups	Type of Statistical Test	Statistical Test to Use
Normal	2	equal	parametric	paired t-test
Normal	>2	equal	parametric	ANOVA (one-way, two-way, or three-way)
Non-normal or no distribution	2	unequal	non-parametric	Signed Rank test, Friedman's test, Wilcoxon
Non-normal or no distribution	>2	unequal	non-parametric	Kruskal-Wallis

BMP Performance Calculations

BMP performance may also be evaluated through relatively simple calculations without the need for statistical testing. Here, two primary techniques are discussed. These include pollutant load/ concentration reductions and runoff volume reductions.

BMP Pollutant Load or Concentration Reductions

Assessing BMP performance by comparing influent and effluent pollutant loads or concentrations is a straightforward way of gauging the effectiveness of a BMP in treating stormwater runoff. If the effluent load/concentrations are found to be less than the influent load/concentration, the reduction may be attributed to the ability of the BMP to remove contaminants from runoff. Higher reductions indicate a relatively high-performing BMP while lower reductions suggest the opposite. If no reduction is found (i.e. effluent > influent), it can be concluded that the BMP is not functioning as a water quality treatment device and is exporting pollutants to downstream receiving waters.

The following equations may be used to evaluate differences in influent-effluent pollutant loads or concentrations. The differences in loads may be expressed as an absolute mass quantity (i.e. lbs. or kg, see first equation) or as a percent load reduction (second equation). If using concentrations rather than

[12] A useful website for obtaining the S+ value is: http://vassarstats.net/binomial01.html.

loads (e.g. EMCs), the difference is computed as a percent only (third equation) since concentrations are generally not subtracted from one another.

$$BMP\ pollutant\ load\ reduction\ (lbs\ or\ kg) = influent\ load - effluent\ load$$

$$BMP\ \%\ pollutant\ load\ reduction = \frac{(influent\ load - effluent\ load)}{influent\ load}\ x\ 100$$

$$BMP\ \%\ EMC\ reduction = \frac{(influent\ EMC - effluent\ EMC)}{influent\ EMC}\ x\ 100$$

An example of this approach to BMP performance evaluation is provided in Table 14. The table shows the annual influent pollutant loads of total nitrogen and total phosphorus, the percent removal in the wet detention pond, and the annual load reduction due to the removal of nutrients by the BMP. The results suggest that the BMP is more effective at removing total phosphorus (90% removal) than total nitrogen (45% removal).

Table 14. Annual Load Reductions for Total Nitrogen and Total Phosphorus

Parameter	Annual Loading (kg/yr.)	Removal in Pond (%)	Annual Load Reduction (kg/yr.)
Total Nitrogen	600	45	270
Total Phosphorus	80	90	72

Note that this type of analysis also pertains to TMDL pollutant load analyses (Goal 3). The load reductions observed for a single BMP can be used to calibrate a water quality model for evaluating the overall water quality impact of TMDL implementation over large areas. This enables DOTs to quantify both the current baseline loads given the current level of BMP implementation, as well as predict the potential load reductions that could be achieved by implementing additional BMPs to meet TMDL requirements.

BMP Runoff Volume Reductions

Another metric to quantify the performance of a BMP is runoff volume reduction. This is particularly useful for BMPs designed to store, infiltrate, and evapotranspire captured stormwater. Evaluating volume reductions requires accurately measuring paired influent and effluent volumes, typically based on continuous flow measurements over a specified duration. This can be difficult for BMPs with poorly defined or non-existent inlets and/or outlets where runoff primarily occurs as sheet flow. In some cases, model simulations may be needed to estimate the volumes. The calculation of runoff volume reduction is similar to those for load or EMC reduction described above:

$$BMP\ flow\ volume\ reduction = \frac{(influent\ volume - effluent\ volume)}{influent\ volume}\ x\ 100$$

Note that it is possible for the effluent volume to be greater than the influent volume. For example, if two relatively large storms occur back-to-back, the runoff from the earlier storm may still be stored within the BMP when the second storm begins. In this case, the runoff from the second storm can displace the stored water from the first storm, causing additional discharge at the outlet. This indicates that the BMP is exporting water downstream rather than reducing the influent volume.

For additional data analysis methods to evaluate BMP effectiveness, see Geosyntec and Wright Water Engineers (2009).

2.9.4.3 Goal 3: TMDL Effectiveness Monitoring

Data analysis methods for TMDL effectiveness monitoring differ depending on whether a before/after study design or a trend monitoring study design was used. In the former case, the before and after datasets are considered independent because the monitoring occurred at different times, even if all the data were collected at the same monitoring station. Hence, a parametric test (which assumes independent observations) is appropriate such as the unpaired t-test. The non-parametric equivalent is the Rank Sum test. These tests will determine whether there is a statistically significant difference in the sample means of the two data groups.

For trend studies, which require long-term data records, there is usually some data preparation needed before any analysis can be done. For example, if the measurement interval varied over the monitoring period (e.g. 15-min flow measurements in year 1, hourly flow measurements in year 2), the data should be reduced or aggregated as needed to provide consistent, fixed measurement intervals for the entire record (Meals et al., 2011). Any gaps in the data should also be identified, as well as any known or suspected measurement errors. In addition, the data may need to be transformed to a normal distribution to satisfy the assumptions of parametric testing.

Once this has been completed, a parametric linear regression analysis is recommended to test for trends. Linear regression determines if a dependent or response variable (e.g. pollutant concentration or load) varies linearly over time (independent variable). Linear regression is commonly done graphically by fitting a linear curve to a scatter plot of individual observations (see example of scatter plot in Figure 11). If the curve has a high R^2 value and a p-value of < 0.05 (indicating statistical significance at a 95% confidence level), it can be concluded that the response variable is either increasing or decreasing in a linear fashion over time. Trends in the data often help inform management actions. For example, if water quality is found to be declining over time despite TMDL implementation, additional pollution control strategies may be warranted. The non-parametric equivalent of linear regression is the Mann-Kendall test. This test is recommended either for highly skewed data which cannot be transformed or if there are many gaps or outliers in the data record. For more information on trend analysis, see Meals et al. (2011).

Table 15 below summarizes some statistical tests that can be used for evaluating TMDL effectiveness monitoring data.

Table 15. Examples of Statistical Tests to Use for TMDL Effectiveness Monitoring. From USEPA (2011).

Type of Data Analysis	Study Design	Type of Statistical Test	
		Parametric	Non-parametric
Compare two independent groups	Before/After	Student's t-test (unpaired)	Rank Sum test
Evaluate the relationship between one data group and time	Trend Monitoring	Linear Regression Analysis	Mann-Kendall Test
Evaluate the relationship between one data group, time, and other variables	Trend Monitoring	Multiple Linear Regression Analysis	Mann-Kendall Trend Test with LOWESS

2.9.4.4 Goal 4: Evaluate Efficiency of Routine Maintenance Activities

There are various data analysis procedures for evaluating DOT routine maintenance activities such as street sweeping or catch basin clean-outs.

For street sweeping studies with a paired drainage area approach (i.e. one control drainage and one treatment drainage), the EMC medians for drainage areas with sweeping and no sweeping can be compared to test for statistically significant differences (use paired t-test if equal variance or Signed Rank test if unequal). If the median of the swept EMCs is significantly lower (at $p < 0.05$), then sweeping can be considered effective. A similar comparison can be done for pollutant loads. The efficiency of the sweeping can be reported as the percent reduction in EMCs or loads. For an example of this, see Terstriep (1983).

2.10. Key Element - QC Procedures

Quality Control (QC) procedures ensure the integrity of the monitoring data by minimizing the chances of contamination and preventing any bias or systematic error in the collected data. This section presents QC protocols for sample collection, handling, storage, and transport.

- Sample Container Type: The type of sample container (e.g. plastic, glass) required is based on the water quality indicator being measured. The Code of Federal Regulations (40 CFR 136.3) identifies the required sample container type. If a composite sample is collected and sub-sampled into other containers, the material of the composite bottle must be compatible with *all* the pollutants being measured.

- Sample Preservatives: 40 CFR 136.3 identifies the required sample preservative based on the analytical method. For example, samples for nutrients and metals require the addition of a sample preservative to ensure that the sample remains stable from the time of collection to the time of delivery to the laboratory. In addition, all samples, must be placed in a cooler packed with ice to maintain a cool temperature (approximately 4° C) for transport to the laboratory.

- Sample Holding Time: The sample holding time refers to the length of time the sample can be held in storage before it must be processed by the laboratory. Holding times vary widely depending on the type of water quality indicator being measured. The water quality indicators with the shortest holding times dictate the logistics of delivering the samples to the laboratory. 40 CFR 136.3 identifies USEPA approved holding times for various types of pollutants.

- Sample Volume: Collection of a sufficient volume of sample is required to meet the laboratory's minimum requirements for analysis. Sample volumes vary depending on the pollutant being measured. Before monitoring begins, the required sample volume for each pollutant should be confirmed with the laboratory. Sample bottles must be of the correct size to ensure that the appropriate volume is collected. If collecting QC samples (discussed below), extra sample volume is needed.

- Collection of QC Samples: QC samples are required to evaluate the quality of the reported data. There are many different types of QC samples that verify various aspects of field sampling and laboratory data quality. QC samples should be tested for the same water quality methods as the stormwater samples and should be collected at the same time as the stormwater samples. The following types of QC samples are recommended:

    - Equipment Blanks: Equipment blanks test for possible contamination due to the equipment being used to physically collect the samples. Also, equipment blank samples are typically prepared when samples are being collected for metals, nitrates, and organic contaminants such as pesticides, herbicides, PAHs, organic carbon, and phthalate compounds (Caltrans, 2012). Equipment blanks are especially important when using an automated sampler which have a number of moving parts and components (e.g. suction tubing, pumps and pump tubing, sample bottles), all of which can potentially introduce contamination. For example, the inside walls of the suction tubing tend to accumulate sediment over time which can potentially contaminate the samples if the sediment dislodges due to the pumping action of the sampler. Equipment blanks should be collected in the exact same manner as the stormwater samples. The only difference is they are collected using clean, analyte-free water (which can be obtained from the lab), not from stormwater runoff.

    - Field Blanks: Field blanks test for possible contamination during the handling, storage, and transport of the stormwater samples to the laboratory. Field blanks are typically used only when samples are being collected for laboratory analysis for bacteria, trace metals, nitrates, and trace organic contaminants such as pesticides, herbicides, PAHs, organic carbon, and phthalate compounds (Caltrans, 2012). A field blank is collected by pouring clean, analyte-free water directly into the laboratory bottle in the field and then process the sample like the other stormwater sample. In contrast to an equipment blank, the physical process of collecting the sample is omitted. Field blanks should be prepared at least once by each field sampling team at a rate of 5% of all storm samples (Ohio EPA, 2012). Additional field blanks are needed if field personnel, equipment, or sampling procedures change.

    - Trip Blanks: Trip blanks are recommended but only if monitoring volatile organic compounds (VOCs). They are very similar to field blanks in that they test for potential contamination from sample handling, storage, and transport. Trip blanks are provided by the laboratory and must remain unopened until they reach the laboratory. If multiple coolers are used to store the samples, a trip blank is needed in each cooler (Ohio EPA, 2012).

- o Field Duplicates: Field duplicates (sometimes called split samples or replicates) provide a measure of the reproducibility and precision of the sampling and laboratory analysis procedures. Field duplicates are collected by pouring or splitting a stormwater sample into two identical sets of laboratory sample bottles. Thus, twice the normal sample volume is needed. Field duplicates should be collected at each monitoring station either once per season or once every 10 monitoring events, whichever is greater (Caltrans, 2013).

- Chain of Custody: A chain of custody (COC) form is provided by the laboratory and acknowledges the transfer of the stormwater samples from field personnel to the testing laboratory. The COC includes basic information about the samples such as the sample ID, date and time collected, and the pollutants to be tested by the laboratory. The COC must be completed by field sampling personnel (in pen) and signed by both field and laboratory personnel upon delivery of the samples to the laboratory. It is important that COCs are completed correctly to avoid any confusion with the laboratory and a possible mix-up of the samples. The information on the COC should always be double-checked against the sample bottle labels to ensure consistency.

A more formal means of documenting and describing QC procedures is a Quality Assurance Project Plan or QAPP. Some DOTs may want to develop (or have already developed) a QAPP for their monitoring programs. The procedures for developing QAPPs are not discussed here. For more information, see the following:

- Caltrans' *Stormwater Monitoring Guidance Manual* (Caltrans, 2015)
- Geosyntec and Wright Water Engineers' *Urban Stormwater BMP Performance Monitoring* (Geosyntec and Wright Water Engineers, 2009)
- The USEPA's *Volunteer Monitor's Guide to Quality Assurance Project Plans* (USEPA, 1996)
- The USGS' *National Field Manual for the Collection of Water-Quality Data* (USGS, 2015)

CHAPTER 3 – EFFECTIVE MONITORING PROTOCOLS FOR A STATE DOT MONITORING PLAN

The term "protocols" used in Chapter 3 refers to a set of comprehensive key elements essential to establishing an efficient way to monitor stormwater quality suited for specific monitoring goals and objectives. Water quality protocol is important in order to produce quality data to meet monitoring goals. This chapter presents a set of stormwater quality monitoring protocols summarizing each of the key monitoring elements needed to create an effective state DOT stormwater monitoring program. These recommendations are based on the findings of the literature review and synthesis which, as mentioned previously, included a wide range of monitoring studies by state DOTs; universities and research organizations; local, state, and federal government agencies; and many others. Since the nature of monitoring varies depending on a DOT's goals, recommendations are provided separately for each of the four DOT monitoring goals identified previously (see Chapter 1).

This chapter is intended as a condensed version of what was previously presented in more detail in Chapter 2. In this chapter, the information is presented as a series of simple, user-friendly tables and figures where the DOT can quickly obtain information on a specific monitoring element for a specific goal. Throughout this chapter, if more detailed information is needed, the reader may refer to Chapter 2. Note that the stormwater monitoring protocols presented in this chapter are not intended to be prescriptive so much as provide a broad overview of the types of considerations to make when designing an effective monitoring program.

The key monitoring elements are summarized as follows and form the structure of this chapter.

- Monitoring Station Network
    - Location of Monitoring Station
    - Number of Monitoring Stations
- Water Quality Pollutants
- Duration of Monitoring
- Number of Storms to Sample
- Sampling Method
- Monitoring Equipment
- Storm Event Criteria
- Data Analysis Technique

3.1. Goal 1: Characterize Stormwater Runoff Quality Monitoring Plan

Characterizing stormwater runoff quality from DOT highways and transportation infrastructure is associated with NPDES permit requirements or as part of a broader data collection effort to support State DOT stormwater management programs. The following protocol described in the subsequent sections apply to Goal 1.

3.1.1 Key Element – Monitoring Station Network

3.1.1.1 Location of Monitoring Stations

Monitoring stations should be located at DOT discharge points along the DOT storm drainage network (ditches, curb inlets, catch basins, pipe outfalls) that receive storm runoff from the roadway or facility of interest. Pipe outfalls are generally recommended if possible. For monitoring sheet flow, the stations may be located at the edge of pavement or other areas of the road shoulder.

3.1.1.2 Number of Monitoring Stations

Table 16 below summarizes the recommended number of monitoring stations for Goal 1. This information is based on a review of 10 different DOT studies which were previously summarized in Table 6 of Chapter 2.

Table 16. Recommended Number of Monitoring Stations for Goal 1

Type of Monitoring Study	Recommended Number of Monitoring Stations
State-Wide	10 to 50 stations, depending on the area of the state and geographic variation within a state.
Transportation Facility/Hotspot	1 to 3 stations
Highway Specific	3 to 5 stations

3.1.2 Key Element - Water Quality Pollutants

Table 17 identifies the recommended pollutants for characterizing stormwater runoff quality.

Table 17. Typical Pollutants to Monitor for Goal 1

Pollutant Category	Pollutants
Particulates	TSS
Nutrients – Nitrogen	TKN, NO3-2, NH3-N, Organic Nitrogen[1], TN[1]
Nutrients - Phosphorus	TP, PO4
Metals	Cadmium, Lead, Zinc, Copper, and Iron (total and dissolved)
Bacteria	E. Coli, Fecal Coliform, and Enterococcus[6]
Organic Contaminants[5]	TPH, Oil and Grease
	PCBs (Total)
	PAHs (Total)
General	Water Temperature[3], DO, Specific Conductivity, pH, Hardness, and TDS
Herbicides[5]	Varies based on chemical control utilized[2]
De-Icing[4, 5]	Chlorides, TDS, Conductivity

[1] Note that these chemicals are calculated, not measured. Organic nitrogen = TKN – NH3, while TN = TKN + NO3-2.
[2]Varies depending on which herbicides are being used by the DOT for vegetation control. Examples of common herbicides include glyphosate, diuron, and others.
[3]In cold weather states, air temperature is also sometimes measured to determine whether the measured precipitation was rain or snow.
[4]These pollutants are only applicable to cold weather states where de-icing materials are applied to roadways during winter storms.
[5]Suitable pollutant for hotspot monitoring
[6]For discharges to marine waters

3.1.3 Key Element – Duration of Monitoring

In general, water quality monitoring is recommended for a minimum of 2 years for states with at least 30” of annual precipitation (mainly Eastern U.S., Northern California, and Pacific Northwest) or 3 years for states with less than 30” annual precipitation (mainly western U.S.). This recommendation is subject to change due to the variability of how many sampling oppurtunities present themselves and DOT permit requirements. See Section 2.4 for more information on estimating the duration of monitoring period.

3.1.4 Key Element - Number of Storms to Sample

Ten (10) storms per year are recommended to be sampled over the monitoring period. Again, this is based on a number of DOT and other studies, and is considered a broad guideline only (see Section 2.4 for more information). This would result in 20 to 30 storms over the course of 2 to 3 years. The power analysis method (see Section 2.5) is recommended once a few storms have been sampled to obtain minimum number of sampled storm events for statistical significance .

When collecting grab samples to characterize stromwater quality runoff from DOT facilities or hotspots, a minimum of three samples should be collected during the first flush of a storm event that meets the project's storm event criteria (See Section 2.8). The first sample would represent the uncontamnated sample that is collected upgradient from the hotspot. The remaining two samples allows for the ability to compare the water quality results among all of the samples. In some cases, the laboratory results of one sample may differ from the other. This may be due to collecting the samples at different points during the storm or the concentration of the pollutant associated with the pollutant source.

3.1.5 Key Element - Sampling Method

A flow-weighted composite sample is the recommended sampling method for characterizing stormwater runoff quality. Other sampling methods such as grab sampling, however, may be applicable. For example, characterizing stormwater quality runoff associated with hotspots in the DOT right of way can be approached by using either grab sampling or composite sampling. The collection of grab samples is recommended when a DOT facility needs to quickly characterize the runoff and determine the primary pollutant associated with the hotspot. The grab sample should be collected at the most downstream point of the hotspot or at a nearby inlet to the assumed pollutant source. A second grab sample, upgradient from the hotspot should also be collected. This provides information on background conditions and allows a comparison among the contaminated and uncontaminated sample.

Automated sampling, is recommended when the DOT facility needs to identify the specific pollutant source and conduct a detail study over a period of time. The automated sampling approach utilizes several monitoring locations to narrow down the pollutant source. Monitoring stations should be located within each pipe network at the hotspot to determine the pollutant source. Once the pollutant source is verified, stormwater quality monitoring should be continued until the pollutant is minimized. Chapter 2, Section 2.6 provides information on sampling methods. Table 18 presents the recommended sampling method for Goal 1.

Table 18. Recommended Sampling Method for Goal 1

Goal	Discrete (by Automated Sampler) *	Flow-Weighted Composite
1	X	X

*The term "discrete" is used here to differentiate from "grab" samples. Discrete in this context refers to samples that are collected by an automated sampler, as opposed to grab samples which are collected by hand. Discrete sampling by an automated sampler means collecting a series of individual samples for laboratory analysis to provide data on the within-storm variability of the water quality pollutants.

3.1.6 Key Element - Monitoring Equipment

Table 19 identifies precipitation, flow and sampling equipment along with the necessary equipment required for monitoring station power and telemetry. Maintenance information for these materials and equipment is provided in Appendix A.

Table 19. Recommended Monitoring Equipment for Goal 1

Monitoring Parameter or Equipment Type	Recommended Equipment
Precipitation	tipping bucket precipitation gauge[1]
Flow	depends on nature of flow at monitoring station[2]
Sample Collection – Automated[3]	automated sampler (suggest refrigerated model in warm climates)
Sample Collection – Grab[4]	laboratory bottles[5], powder-free nitrile gloves, cooler with ice, extension pole[6], intermediate container[6], chain of custody, pen
Telemetry	cellular modem, antenna, vendor software, cellular phone account, data logger
Power	Power is only needed if using automated samplers. AC power is ideal. If not available, use a 12 V deep-cycle marine battery and possibly a solar array.

[1]Nearby weather station data should be used in absence of onsite precipitation gauge.
[2]See Section 2.7.2 for equipment options to measure flow
[3]Automated samples are recommended for most pollutants except temperature, pH, dissolved oxygen, oil and grease, TPH, and bacteria. See Chapter 2 for details.
[4]Grab samples are mandatory if monitoring the pollutants listed in Footnote 3.
[5]With acid preservatives if needed
[6]Only needed if it is not safe or feasible to wade into the flow pathway
[7]Only needed if not able to collect the sample directly into the laboratory containers

3.1.7 Key Element - Storm Event Criteria

Table 20 presents the recommended storm event criteria to use based on the type of climate.

Table 20. Recommended Storm Event Criteria for Goal 1

Storm Event Criteria	Arid or Semi-Arid Climates[1]	Humid Climates[2]
Precipitation Depth	Minimum 0.1”	Minimum 0.1”
Precipitation Duration	At least 1 hour	At least 1 hour
Antecedent Dry Period	At least 72 hours	At least 24 hours
Inter-Event Dry Period	10 hours	10 hours

[1] Arid or semi-arid climates refers to regions with relatively low precipitation amounts and distinct wet and dry seasons. Typically, these climates occur in the western U.S.
[2] Humid climates refers to regions with relatively high precipitationprecipitation amounts and fairly consistent precipitationprecipitation from month to month. Typically, these climates occur in the eastern U.S.

3.1.8 Key Element - Data Analysis Technique

The specific type of data analysis to perform varies by monitoring goal. Refer to Chapter 2, Section 2.9 for more information on data analysis methods. Table 21 and Table 22 present the recommended data analysis methods for Goal 1.

Table 21. Recommended Data Visualization for Goal 1

Goal	Summary Table	Bar Graph	Hyetograph	Hydrograph	Box Plot	Scatter Plot	Trend Plot
1	X	X	X	X	X	X	X

Table 22. Recommended Data Analysis for Goal 1

Goal	Calculation of EMCs	Calculation of Pollutant Loads	Calculation of Treatment Efficiency	Conducting Statistical Analysis	Conducting a Regression Analysis	Conduct a Trend Analysis
1	X	X		X	X	X

3.2. Goal 2: Assess BMP Performance Monitoring Plan

Assessing the performance of existing stormwater treatment BMPs is a key component of a DOT's stormwater management program and is often required by NPDES permits or other regulatory requirements. Assessing BMP performance by monitoring is beneficial in that it identifies which structural BMP types and designs (including retrofits) are most cost-effective at addressing water quality impairments. The following protocol described in the subsequent sections apply to Goal 2.

3.2.1 Key Element – Monitoring Station Network

3.2.1.1 Location of Monitoring Stations

Monitoring locations should be located at the influent and effluent point(s) of the BMP being monitored. The stations must be located as close as possible to the BMP, yet far enough away that the data collected at the influent and effluent will not be influenced by the BMP itself. The effluent station(s) should be located before the discharge exiting the BMP enters the receiving water or flows off the DOT's right of way. If monitoring for within-BMP processes (or if there is no feasible way to sample the effluent), an interior monitoring station is needed. For monitoring individual treatment train components, stations are needed at the inflow and outflow points of the component(s) of interest.

3.2.1.2 Number of Monitoring Stations

The number of monitoring stations needed to assess a BMP's performance depends on the drainage configuration of the BMP (e.g. whether there are multiple influent points or a single influent point), whether interior within-BMP monitoring is desired, and whether monitoring a treatment train or a single self-contained BMP. The following are suggested:

- BMPs with a single influent and effluent point: 2 monitoring stations, one at each
- BMPs with a single influent point and no effluent point: 2 monitoring stations, one at influent and one interior
- BMPs with multiple influent points and multiple effluent points: Monitor at each influent and effluent point to capture all inflows and outflows, up to a maximum of 4 monitoring stations. BMPs with more than 4 influent and effluent locations are not recommended for monitoring.
- Treatment train BMPs: two paired inflow-outflow stations for each individual treatment train component.

3.2.2 Key Element - Water Quality Pollutants

BMP performance monitoring programs typically monitor TSS, nutrients, metals, and certain "general" parameters (e.g. temperature, DO) (Struck, 2006) which are considered to be properties of the water, not contaminants per se. Sometimes, chemical oxygen demand (COD) and/or biological oxygen demand (BOD) are also measured, but are not always necessary. More comprehensive BMP studies might include total PAHs, a large array of metals, or others. Table 23 presents the recommended pollutants to monitor as part of BMP performance.

Table 23. Recommended Pollutants to Monitor for Goal 2

Pollutant Category	Pollutants
Particulates	TSS
Nutrients – Nitrogen	TKN, NO3-2, NH3-N, Organic Nitrogen[1], TN[1]
Nutrients - Phosphorus	TP, PO4
Bacteria	E. Coli and Fecal Coliform
Metals	Total and Dissolved Cadmium, Zinc, Copper, and Lead
Organic Contaminants[2]	TPH, PAHs, Oil and Grease
General	DO, temperature, pH, hardness, and TDS

[1]Note that these pollutants are calculated, not measured. Organic nitrogen = TKN – NH3-N, while TN = TKN + NO3-2.
[2]Applicable mainly to BMP monitoring in hotspot areas, for example oil-grit separators at vehicle fueling facility.

3.2.3 Key Element – Duration of Monitoring

Paired influent-effluent monitoring is recommended to evaluate BMP efficiency. Water quality monitoring is recommended for a minimum of 2 years for states with at least 30" of annual precipitation (mainly Eastern U.S., Northern California, and Pacific Northwest) or 3 years for states with less than 30" annual precipitation (mainly western U.S.).This recommendation is a general guidleine and subject to change due to the variability of how many sampling oppurtunities present themselves and DOT permit requirements. See section 2.4 and Table 6 for more information on estimating the duration of the monitoring period.

3.2.4 Key Element - Number of Storms to Sample

Ten (10) storms per year are recommended to be sampled over the monitoring period. This would result in 20 to 30 storms over the course of 2 to 3 years. The power analysis method (see Section 2.5) is recommended once a few storms have been sampled to obtain minimum number of sampled storm events for statistical significance .

3.2.5 Key Element - Sampling Method

A flow-weighted composite sample is the recommended sampling method for assessing BMP performance. Flow-weighted composite samples should be collected at each influent and effluent station. Chapter 2, Section 2.6.3 provides detailed information on the collection of a flow-weighted composite sample.

3.2.6 Key Element - Monitoring Equipment

Table 24 provides the recommended monitoring equipment for measuring precipitation and flow, for collecting the samples (either with automated samplers or by hand); for accessing the monitoring data remotely (telemetry); and finally, for powering the equipment. Automated sampling is recommended for Goal 2. Grab sampling may be used in instances where recommended equipment is not available, however the data may be less meaningful. Site maintenance is required for the listed recommended equipment. See Appendix A for more information.

Table 24. Recommended Monitoring Equipment for Goal 2

Monitoring Parameter or Equipment Type	Recommended Equipment
Precipitation	tipping bucket precipitation gauge[1]
Flow	depends on nature of flow at monitoring station[2]
Sample Collection – Automated[3]	automated sampler (suggest refrigerated model in warm climates)
Sample Collection – Grab[4]	laboratory bottles[5], powder-free nitrile gloves, cooler with ice, extension pole[6], intermediate container[6], chain of custody, pen
Telemetry	cellular modem, antenna, vendor software, cellular phone account, data logger
Power	Power is only needed if using automated samplers. AC power is ideal. If not available, use a 12 V deep-cycle marine battery and possibly a solar array.

[1]Nearby weather station data should be used in absence of onsite precipitation gage.
[2]See Section 2.7.2 for equipment options to measure flow
[3]Automated samples are recommended for most pollutants except pH, temperature, dissolved oxygen, oil and grease, TPH, and bacteria. See Chapter 2 for details.
[4]Grabs are mandatory if monitoring the pollutants listed in footnote 3.
[5]With acid preservatives if needed
[6]Only needed if it is not safe or feasible to wade into the flow pathway
[7]Only needed if not able to collect the sample directly into the laboratory bottles

3.2.7 Key Element - Storm Event Criteria

Table 25 presents the recommended storm event criteria to use based on the type of climate.

Table 25. Recommended Storm Event Criteria for Goal 2

Storm Event Criteria	Arid or Semi-Arid Climates[1]	Humid Climates[2]
Precipitation Depth	Minimum 0.1”	Minimum 0.1”
Precipitation Duration	At least 1 hour	At least 1 hour
Antecedent Dry Period	At least 72 hours	At least 24 hours
Inter-Event Dry Period	10 hours	10 hours

[1] Arid or semi-arid climates refers to regions with relatively low precipitation amounts and distinct wet and dry seasons. Typically, these climates occur in the western U.S.
[2] Humid climates refers to regions with relatively high precipitation amounts and fairly consistent precipitation from month to month. Typically, these climates occur in the eastern U.S.

3.2.8 Key Element - Data Analysis Technique

The specific type of data analysis to perform varies by the monitoring goal. Refer to Chapter 2, Section 2.9 for more information on data analysis methods. Table 26 and Table 27 presents the recommended data analysis methods for Goal 2.

Table 26. Recommended Data Visualization for Goal 2

Goal	Summary Table	Bar Graph	Hyetograph	Hydrograph	Box Plot	Scatter Plot	Trend Plot
2	X	X	X	X	X		X

Table 27. Recommended Data Analysis for Goal 2

Goal	Calculation of EMCs	Calculation of Pollutant Loads	Calculation of Treatment Efficiency	Conducting Statistical Analysis	Calculation of Runoff Volume Reductions	Conducting a Regression Analysis	Conduct a Trend Analysis
2	X	X	X	X	X		X

3.3. Goal 3: TMDL Effectiveness Monitoring Plan

The purpose of TMDL effectiveness monitoring is to evaluate if the DOT's TMDL Implementation Plan has resulted in any measurable water quality benefits in the impaired TMDL receiving waters[13]. The following protocol described in the subsequent sections apply to Goal 3.

3.3.1 Key Element – Monitoring Station Network

3.3.1.1 Location of Monitoring Stations

The monitoring aspect of this goal represents a combination of Goals 1 and 2. Therefore, see sections 3.1.1.1 and 3.2.1.1 for information pertaining to monitoring station locations.

3.3.1.2 Number of Monitoring Stations

TMDL compliance monitoring applies to DOTs that are subject to TMDL requirements. For these DOTs, it is difficult to recommend a specific number or even range of stations since it differs case by case. Some factors to consider include: 1) the relative contribution of pollutant loads of the DOT and size of the area to the impaired TMDL waterway; 2) the size of the area where stormwater treatment is needed to comply with the TMDL; 3) the heterogeneity of these areas; and 4) how much monitoring is needed to populate and calibrate a water quality model for pollutant load analysis. In general, homogeneous highways require fewer stations than heterogeneous highways. Unless there is a presumed contaminant loading rate from highways (that has been accepted by both the regulator and the DOT), then achieving compliance with TMDLs requires both characterization monitoring (Goal 1) and BMP performance monitoring (Goal 2). See sections 3.1.1.2 and 3.2.1.2 for more information.

3.3.2 Key Element - Water Quality Pollutants

The types of pollutants to monitor for is dependent on the nature of the impairments in the local waterways. The most common TMDL pollutants include sediment, pathogens, nutrients, and metals (USEPA, 2017). Certain waterbodies, specifically tie to less typical pollutants such as biological integrity, turbidity, and fecal coliform. DOTs should monitor all pollutant(s) identified in EPA-approved TMDLs in which they are a stakeholder.

3.3.3 Key Element - Duration of Storms to Sample

Water quality monitoring is recommended for a minimum of 2 years for states with at least 30" of annual precipitation (mainly Eastern U.S., Northern California, and Pacific Northwest) or 3 years for states with less than 30" annual precipitation (mainly western U.S.). This recommendation is a general guidleine and subject to change due to the variability of how many sampling oppurtunities present themselves and DOT permit requirements. See section 2.4 for more information on estimating the duration of a monitoring period.

---

13 This assumes that the DOT is currently subject to TMDL requirements, which is not the case in every state. As new TMDLs are continually being issued nationwide, however, it is expected that more and more DOTs will need to conduct TMDL effectiveness monitoring in the future.

3.2.4 Key Element - Number of Storms to Sample

Ten (10) storms per year are recommended to be sampled over the monitoring period. This would result in 20 to 30 storms over the course of 2 to 3 years. The power analysis method (see Section 2.5) is recommended once a few storms have been sampled to obtain minimum number of sampled storm events for statistical significance. The sampling duration should be based off number of storm events calculated using the power analysis method in order to produce a robust data set. The minimum number of storms will vary between arid/semi-arid and humid climates. For example, an arid climate may only be able to sample 3 storms per year where as a humid climate may have 20-30 qualifying storm events in one-year. The sampled storms should cover a broad range of storm types and sizes occurring at different times of the season or year and under varying antecedent conditions. For more information, see section 2.5.

3.3.5 Key Element - Sampling Method

A flow-weighted composite sample is the recommended sampling method for Goal 3. Chapter 2, Section 2.6 provides detailed information on the collection of a flow-weighted composite sample.

3.3.6 Key Element - Monitoring Equipment

Table 28 provides the recommended monitoring equipment for measuring precipitation and flow, for collecting the samples (either with automated samplers or by hand), for accessing the monitoring data remotely (telemetry), and finally for powering the equipment. Automated sampling is recommended for Goal 3. Site maintenance is required for the listed equipment. See Appendix A for more information.

Table 28. Recommended Monitoring Equipment for Goal 3

Monitoring Parameter or Equipment Type	Recommended Equipment
Precipitation	tipping bucket precipitation gauge[1]
Flow	depends on nature of flow at monitoring station[2]
Sample Collection – Automated[3]	automated sampler (suggest refrigerated model in warm climates)
Sample Collection – Grab[4]	laboratory bottles[5], powder-free nitrile gloves, cooler with ice, extension pole[6], intermediate container[6], chain of custody, pen
Telemetry	cellular modem, antenna, vendor software, cellular phone account, data logger
Power	Power is only needed if using automated samplers. AC power is ideal. If not available, use a 12 V deep-cycle marine battery and possibly a solar array.

[1]Nearby weather station data should be used in absence of onsite precipitation gauge.

[2]See Section 2.7.2 for equipment options to measure flow

[3]Automated samples are recommended for most pollutants except temperature, pH and dissolved oxygen See Chapter 2 for details.

[4]Grabs are mandatory if monitoring the pollutants listed in Footnote 3.

[5]With preservative if needed
[6]Only needed if it is not safe or feasible to wade into the flow pathway
[7]Only needed if not able to collect the sample directly into the laboratory bottles
[6]Only needed if not able to collect the sample directly into the laboratory bottles

3.2.7 Key Element - Storm Event Criteria

Table 29 presents the recommended storm event criteria to use based on the type of climate.

Table 29. Recommended Storm Event Criteria for Goal 3

Storm Event Criteria	Arid or Semi-Arid Climates[1]	Humid Climates[2]
Precipitation Depth	Minimum 0.1"	Minimum 0.1"
Precipitation Duration	At least 1 hour	At least 1 hour
Antecedent Dry Period	At least 72 hours	At least 24 hours
Inter-Event Dry Period	10 hours	10 hours

[1] Arid or semi-arid climates refers to regions with relatively low precipitation amounts and distinct wet and dry seasons. Typically, these climates occur in the western U.S and receive less than 30 inches of precipitation annually.
[2] Humid climates refers to regions with relatively high precipitation amounts and fairly consistent precipitation from month to month. Typically, these climates occur in the eastern U.S and receive at least 30 inches of precipitation per year.

3.3.8 Key Element - Data Analysis Technique

The specific type of data analysis to perform varies by monitoring goal. Refer to Chapter 2, Section 2.9 for more information on data analysis methods. Table 30 and Table 31 presents the recommended data analysis methods for Goal 3.

Table 30. Recommended Data Visualization for Goal 3

Goal	Summary Table	Bar Graph	Hyetograph	Hydrograph	Box Plot	Scatter Plot	Trend Plot
3	X	X	X	X	X	X	X

Table 31. Recommended Data Analysis for Goal 3

Goal	Calculation of EMCs	Calculation of Pollutant Loads	Calculation of Treatment Efficiency	Conducting Statistical Analysis	Conducting a Regression Analysis	Conduct a Trend Analysis
3	X	X	X	X	X	X

3.4. Goal 4: Routine Maintenance Activities Monitoring Plan

Routine DOT maintenance activities, such as street sweeping; catch basin clean-outs; trash pick-up; and ditch cleaning, provide pollutant source control. These practices can potentially improve stormwater quality if they are performed on a regular basis. Other routine DOT maintenance includes road-side vegetation control using herbicides and the application of de-icing materials to the road surface during winter storms (in cold climates). The following protocol described in the subsequent sections apply to Goal 4.

3.4.1 Key Element – Monitoring Station Network

3.4.1.1 Location of Monitoring Stations

Evaluating the efficiency of routine DOT maintenance activities (street sweeping, catch basin and ditch clean-outs, vegetation control, and use of de-icing materials) requires specific monitoring locations. The following presents the recommended monitoring locations based on DOT maintenance activities:

- Herbicide Applications: Monitor at centroid of herbicide spray area and immediately downslope of the spray area. It is also recommended to monitor at a control site where herbicides were not sprayed but all other conditions are the same. This can either be at a separate location nearby, or along the road shoulder just before runoff enters the spray area.
- De-icing/Salt Applications: Monitor along the road shoulder or in a catch basin adjacent to roads where de-icing materials are applied. Stations may also be located at bridge decks where runoff is discharged into a downstream BMP.
- Street Sweeping/Catch Basin Clean-Outs: Storm drain pipe outfalls at the base of the catchment area where street sweeping occurs.
- Vegetation Maintenance (Mowing): Edge-of-pavement at the road shoulder, in the centroid of the mowing area, immediately downslope of the mowing area.

3.4.1.2 Number of Monitoring Stations

Table 32 provides the recommended number of monitoring stations for evaluating routine DOT maintenance activities. This information is based on a review of 10 different DOT studies which were previously summarized in Table 6 in Chapter 2.

Table 32. Recommended Number of Monitoring Stations for Goal 4

Type of Routine Maintenance Activity	Recommended Number of Monitoring Stations
Herbicides[14]	2 to 3 stations at control sites and 2 to 3 stations at treatment areas
De-icing/Salt Applications during Adverse Weather Conditions	2 per bridge crossing or 1 station per catchment
Street Sweeping/Catch Basin Clean-Out	2 stations per monitoring study (1 station per catchment)
Vegetation Maintenance (Mowing)	1 to 3 stations per mowing area

3.4.2 Key Element - Water Quality Pollutants

Table 33 presents the recommended pollutants to monitor when evaluating efficiency of routine maintenance activites.

Table 33. Recommended Pollutants to Monitor for Goal 4

Pollutant Category	Pollutants
Street Sweeping	TSS, TDS, VSS, particle size distribution
Catch Basin Clean-outs	TSS, TDS, VSS, particle size distribution
Herbicide Applications	Varies depending on what herbicide was used for vegetation control (i.e. Glyphosate, Diuron, and etc.), hardness, temperature, pH
Application of De-icing Materials	Chloride, conductivity
Vegetation Maintenance (Mowing)	TSS, TDS, VSS, particle size distribution

3.4.3 Key Element - Duration of Storms to Sample

Paired control-treatment monitoring is recommended to evaluate BMP efficiency. “Control” means monitoring a reference area where no maintenance occurs, while “treatment” means monitoring another nearby area where maintenance does occur. To provide a level of experimental control, the treatment and control sites should have similar characteristics (save for the maintenance activity being evaluated)

[14] Within each control and treatment area, paired influent-effluent stations are recommended, or triplicate stations (influent, within interior area, effluent).

to exclude unwanted variability. With the exception of herbicide applications , water quality monitoring is recommended for a minimum of 2 years. In general, water quality monitoring is recommended for a minimum of 2 years for states with at least 30” of annual precipitation (mainly Eastern U.S., Northern California, and Pacific Northwest) or 3 years for states with less than 30” annual precipitation (mainly western U.S.). A three-month monitoring period is sufficient for monitoring herbicide applications. This recommendation is a general guidleine and subject to change due to the variability of how many sampling oppurtunities present themselves and DOT permit requirements. See section 2.4 and Table 6 for more information on estimating the duration of the monitoring period.

3.4.4 Key Element - Number of Storms to Sample

Five (5) to ten (10) storms per year are recommended to be sampled over the monitoring period. This would result in 10 to 30 storms over the course of 2 to 3 years. The power analysis method (see Section 2.5) is recommended once a few storms have been sampled to obtain minimum number of sampled storm events for statistical significance. The sampling duration should be based off number of storm events calculated using the power analysis method in order to produce a robust data set. The minimum number of storms will vary between arid/semi-arid and humid climates. For example, an arid climate may only be able to sample 3 storms per year where as a humid climate may have 20-30 qualifying storm events in one-year. The sampled storms should cover a broad range of storm types and sizes occurring at different times of the season or year and under varying antecedent conditions. For more information, see section 2.5.

3.4.5 Key Element - Sampling Method

A flow-weighted composite sample is the recommended sampling method for evaluating efficiency of routine maintenance activities. Other sampling methods such as grab sampling, may also be applicable. Automated sampling, is recommended when the DOT needs to identify the specific pollutant source and conduct a detail study over a period of time. The automated sampling approach utilizes several monitoring locations to narrow down the pollutant source. Once the pollutant source is verified, stormwater quality monitoring should be continued until the pollutant is minimized. Changes to the frequency of the maintenance activity may be necessary to minimize the pollutant. Chapter 2, Section 2.6 provides information on sampling methods. Table 34 presents the recommended sampling method for Goal 4.

The collection of grab samples is only necessary and useful when a DOT needs to quickly characterize the runoff and determine the primary pollutant associated at the maintenance activity of interest. An initial grab sample should be collected prior to beginning the maintenance activity. A second grab sample should be collected upon completion of the maintenance activity. This provides information on background conditions and allows a comparison among the contaminated and uncontaminated sample.

Table 34. Recommended Sampling Method for Goal 4

Goal	Discrete (by Automated Sampler) *	Flow-Weighted Composite
4	X	X

*the term "discrete" is used here to differentiate from "grab" samples. Discrete in this context refers to samples that are collected by automated sampler, as opposed to grab samples which are collected by hand. Discrete sampling by automated sampler means collecting a series of individual samples for lab analysis to provide data on the within-storm variability of water quality pollutants.

3.4.6 Key Element - Monitoring Equipment

Table 35 provides the recommended monitoring equipment for measuring precipitation and flow, for collecting the samples (either with automated samplers or by hand), for accessing the monitoring data remotely (telemetry), and finally for powering the equipment. Automated sampling is recommended for Goal 4. Grab sampling may be used in instances where recommended equipment is not available. Site maintenance is required for the listed recommended equipment. See Appendix A for more information.

An effective way to monitor application of de-icing materials on highways is by using conductivity meters (i.e., water quality sondes). Dissolved salt is highly conservative when entirely dissolved in runoff and correlates very closely with conductivity. There are various motives for measuring salt loading in runoff for example comparing it with the Winter Severity Index, with management decision systems, or even in certain areas for TMDL purposes.

Table 35. Recommended Monitoring Equipment for Goal 4

Monitoring Parameter or Equipment Type	Recommended Equipment
Precipitation	tipping bucket precipitation gauge[1]
Flow	depends on nature of flow at monitoring station[2]
Sample Collection – Automated[3]	automated sampler (suggest refrigerated model in warm climates)
Sample Collection – Grab[4]	laboratory bottles[5], powder-free nitrile gloves, cooler with ice, extension pole[6], intermediate container[6], chain of custody, pen
Telemetry	cellular modem, antenna, vendor software, cellular phone account, data logger
Power	Power is only needed if using automated samplers. AC power is ideal. If not available, use a 12 V deep-cycle marine battery and possibly a solar array.
Water Quality Sonde	Conductivity Meter

[1]Nearby weather station data should be used in absence of onsite precipitation gage.
[2]See Section 2.7.2 for equipment options to measure flow
[3]Automated samples are recommended for most pollutants except pH and temperature. See Chapter 2 for details.

[4]Grab samples are mandatory if monitoring the pollutants listed in Footnote 3.
[5]With preservatives if needed
[6]Only needed if it is not safe or feasible to wade into the flow pathway
[7]Only needed if not able to collect the sample directly into the laboratory bottles

3.4.7 Key Element - Storm Event Criteria

Table 36 presents the recommended storm event criteria to use based on the type of climate.

Table 36. Recommended Storm Event Criteria for Goal 4

Storm Event Criteria	Arid or Semi-Arid Climates[1]	Humid Climates[2]
Precipitation Depth	Minimum 0.1”	Minimum 0.1”
Precipitation Duration	At least 1 hour	At least 1 hour
Antecedent Dry Period	At least 72 hours	At least 24 hours
Inter-Event Dry Period	10 hours	10 hours

[1] Arid or semi-arid climates refers to regions with relatively low precipitation amounts and distinct wet and dry seasons. Typically, these climates occur in the western U.S.
[2] Humid climates refers to regions with relatively high precipitation amounts and fairly consistent precipitation from month to month. Typically, these climates occur in the eastern U.S.

3.4.8 Key Element - Data Analysis Technique

The specific type of data analysis to perform varies by monitoring goal. Refer to Chapter 2, Section 2.9 for more information on data analysis methods. Table 37 and Table 38 presents the recommended data analysis methods for Goal 4.

Table 37. Recommended Data Visualization for Goal 4

Goal	Summary Table	Bar Graph	Hyetograph	Hydrograph	Box Plot	Scatter Plot	Trend Plot
4	X	X	X	X			

Table 38. Recommended Data Analysis for Goal 4

Goal	Calculation of EMCs	Calculation of Pollutant Loads	Calculation of Treatment Efficiency	Conducting Statistical Analysis	Conducting a Regression Analysis	Conduct a Trend Analysis
4	X		X			

REFERENCES

Arguez, A., I. Durre, S. Applequist, R. S. Vose, M. F. Squires, X. Yin, R. R. Heim, Jr., and T. W. Owen, 2012: *NOAA's 1981-2010 U.S. Climate Normals: An Overview*. Bulletin of the American Meteorological Society, 93, 1687-1697.

AASHTO. 2013. State-of-the-Practice Report – *Water Quality Monitoring. Center for Environmental Excellence by AASHTO Stormwater Management Community of Practice (CoP).*

Bannerman RB, Legg AD, and Greb SR. 1996. *Quality of Wisconsin Stormwater, 1989–94. U.S. Geological Survey Open File Report* 96-458, 26 p.

Barbaro, H. 2006. *Environmental Concerns of Sand*. Boston, MA: Mass DOT Highway Division.

Barrett ME, Malina JF, Charbeneau RJ. 1995. *Characterization of Highway Runoff in Austin, Texas Area. Technical Report CRWR 263*. Center for Research in Water Resources, University of Texas at Austin.

Burton AG, and Pitt R. 2002. *Stormwater Effects Handbook – A Toolbox for Watershed Managers, Scientists, and Engineers.* Lewis Publishers.

Alwan, Wes & Casey, Patrick. 2013*. NCHRP Synthesis 444 - Pollutant Load Reductions for Total Maximum Daily Loads for Highways*. National Cooperative Highway Research Program.

Caltrans (State of California Department of Transportation). 2003. *A Review of the Contaminants and Toxicity Associated with Particles in Stormwater Runoff*. Document No. CTSW-RT-03-059.73.15. August.

Caltrans (State of California Department of Transportation). 2003a. *Statewide Stormwater Runoff Characterization Study, Monitoring Season 2002-2003*. Document No. CTSW-RT-03-052. June.

Caltrans (State of California Department of Transportation). 2003b. *Roadside Vegetated Treatment Sites (RVTS) Study*. Document No. CTSW-RT-03-028. November.

Caltrans (State of California Department of Transportation). 2012. *Construction Site Monitoring Program Guidance Manual Final Report.* Document No. CTSW-RT-11-255.11.01-F.

Caltrans (State of California Department of Transportation). 2015. *Caltrans Stormwater Monitoring Guidance Manual*. Document No. CTSW-OT-15-999.43.01. November.

Cleveland, William S. 1993. V*isualizing Data*. Hobart Press, P.O. Box 1473, Summit, NJ 07902.

Corsi SR, Greb SR, Bannerman RB, and Pitt RE. 1999. *Evaluation of the Multi-Chambered Treatment Train, a retrofit water-quality management device: U.S. Geological Survey Open-File Report 99–270*, 24 p.

CWP (Center for Watershed Protection). 2008. *Deriving Reliable Pollutant Removal Rates for Municipal Street Sweeping and Storm Drain Cleanout Programs in the Chesapeake Bay Basin*. A report prepared by the Center for Watershed Protection as fulfillment of the USEPA Chesapeake Bay Program grant CB-973222-01. September

De Leon, D.B., Lowe J. 2009. *Standard Operating Procedure for Automated Sampling for Stormwater Monitoring*. Version 1.0. Washington State Department of Ecology, 47 p.

FHWA (Federal Highway Administration). 2001. *Guidance Manual for Monitoring Highway Runoff Water Quality*. Publication No. FHWA-EP-01-022. June

Feist B, Buhle E, Baldwin D, Spromberg J, Damm S, Davis J, and Scholz N. 2017. *Roads to Ruin: Conservation threats to a Sentinel Species across an Urban Gradient*. Ecological Applications.

FHWA (Federal Highway Administration). 2001. *Guidance Manual for Monitoring Highway Runoff Water Quality*. Publication No. FHWA-EP-01-022. June

GeoSyntec Consultants and Urban Water Resources Research Council of ASCE. 2002. *Urban Stormwater BMP Performance Monitoring – A Guidance Manual for Meeting the National Stormwater BMP Database Requirements*. EPA-821-B-02-001. US Environmental Protection Agency. Washington, D.C.

Geosyntec Consultants and Wright Water Engineers, Inc. 2009. *Urban Stormwater BMP Performance Monitoring.* October.

Helsel, D. R., and T. A. Cohn. 1988. *Estimation of descriptive statistics for multiply censored water quality data, Water Resour. Res.*, 24(12), 1997–2004, doi:10.1029/WR024i012p01997.

Helsel, D.R. and R. M. Hirsch. 2002. *Statistical Methods in Water Resources*. Book 4, Chapter A3. U.S. Geological Survey. 522 pages. Available online at: https://pubs.usgs.gov/twri/twri4a3/

Helsel, D.R., 2005. *Nondetects And Data Analysis: Statistics for censored environmental data*. John Wiley and Sons, New York. 250 p. http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471671738.html

HI DOT (State of Hawaii Department of Transportation). 2014. *Storm Water Hawaii Glossary of Terms - Storm Water Hawaii*. Highways Division – Oahu District, Environmental Management Section. Accessed September 6, 2016 at: http://www.stormwaterhawaii.com/about/glossary-of-terms/#glossary-W

Horwatich, Judy A. and Bannerman, Roger T. 2009. *Pollutant Loading to Stormwater Runoff from Highways: Impact of a Highway Sweeping Program*. Wisconsin Department of Transportation. Publication No. 0092-04-04.

King County, WA. 2016. *The Science of Stormwater*. Available on-line at: http://www.kingcounty.gov/services/environment/water-and-land/stormwater/introduction/science.aspx

Kuligowski, RJ. 1997. *An Overview of National Weather Service Quantitative Precipitation Estimates*. National Oceanic and Atmospheric Administration. TLD Office Note 97-4. November.

Law, N.L., K. DiBlasi and U. Ghosh. 2008. *Deriving Reliable Pollutant Removal Rates for Municipal Street Sweeping and Storm Drain Cleanout Programs in the Chesapeake Bay Basin*. September.

Meals DW, Spooner J, Dressing SA, and Harcum JB. November 2011. *Statistical analysis for monotonic trends, Tech Notes 6*. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 23 p. Available online at https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoring-technical-notes

Mozejko, J., 2012. In: Voudouris (Ed.), *Detecting and Estimating Trends of Water Quality Parameters, Water Quality Monitoring and Assessment*. http://dx.doi.org/10.5772/33052. ISBN: 978-953-51-0486-5, InTech, Available from: http://www.intechopen.com/books/water-quality-monitoring-andassessment/detecting-and-estimating-trends-of-water-quality-parameters

MDOT (Michigan Department of Transportation). 1998. *Highway Stormwater Runoff Study*. Final Report. Prepared for the Michigan Department of Transportation. CH2MHILL. April.

Opher T, Friedler E. 2010. *Factors Affecting Highway Runoff Quality*. Urban Water Journal, Pg. 155-172.

Ohio Environmental Protection Agency. 2012. *Manual of Ohio EPA Surveillance Methods and Quality Assurance Practices*. Version 2.1. April.

ODOT (Oregon Department of Transportation). 2016. *Design Guide for Roadside Infiltration Strips in Western Oregon*. Document No. SPR 758. July.

Pitt, R. and K. Parmer. 1995. *Quality Assurance Project Plan: Effects, Sources, and Treatability of Stormwater Toxicants*. Contract No. CR819573. U.S. Environmental Protection Agency, Storm and Combined Sewer Program, Risk Reduction Engineering Laboratory. Cincinnati, OH. February.

Pitt R, Goodson K, Ogburn O, Eppakayala V, Veeravalli SU, Bathi J, Wilson B, Subramaniam S, Clark, S. 2013. *Identification and Treatment of Emerging Contaminants in Wet Weather Flows*. EPA Contract: EP-C-07-014. July 30.

Reckhow, KH., K. Kepford, and W. Warren Hicks. 1993. *Methods for the Analysis of Lake Water Quality Trends*. EPA 841-R-93-003.

Rose WJ, Robertson DM, Mergener EQ. 2004. *Water Quality, Hydrology, and the Effects of Changes in Phosphorus Loading to Pike Lake, Washington County, Wisconsin, with Special Emphasis on Inlet-to-Outlet Short-Circuiting*. USGS Scientific Investigations Report 2004-5141.

Smith, K.P. 2002. *Effectiveness of Three Best Management Practices for Highway-Runoff Quality along the Southeast Expressway, Boston, Massachusetts*, Prepared for U.S. Geological Survey.

Smith, K.P.,and Granato, G.E., 2009. *Quality of stormwater runoff discharged of Massachusetts highways, 2005-2007*: U.S. Geological Survey Scientific Investigations Report 2009-5269, 198 p.

Struck, S.D. 2006. *Stormwater Best Management Practice Monitoring. New Jersey Flows Publication of the New Jersey Water Resources Research Institute*, 7(1) p. 4-8.

Terstriep M, Noel D, Bender M. 1983. *Nationwide Urban Runoff Project, Champaign, Illinois: Assessment of the Impact of Urban Storm Runoff on an Agricultural Receiving Stream*. Contract No. 1-5-39600.

Tufte, Edward R. 1983. *The Visual Display of Quantitative Information*. Graphics Press, Box 430, Cheshire, CT 06410.

USEPA (U.S. Environmental Protection Agency). 1982. *Handbook for Sampling and Sample Preservation of Water and Wastewater*. EPA-600/4-82-029. Environmental Monitoring and Support Laboratory, Cincinnati, Ohio. September.

USEPA (U.S. Environmental Protection Agency). 1996. *The Volunteer Monitor's Guide to Quality Assurance Project Plans*. EPA 841-B-96-003. Office of Wetlands, Oceans, and Watersheds. September.

USEPA (U.S. Environmental Protection Agency). 2000. *Meteorological Monitoring Guidance for Regulatory Modeling Applications*. EPA-454/R-99-005. Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711. February.

USEPA (U.S. Environmental Protection Agency). 2011. *Technical Guidance for Exploring TMDL Effectiveness Monitoring Data*. Contract EP-C-08-002. December.

USEPA (U.S. Environmental Protection Agency). 2016. *Guidance for Designing Communication Systems – For Water Quality Surveillance and Response Systems*. EPA 817-B-16-002. September.

USEPA (U.S. Environmental Protection Agency). 2017. Impaired Waters and TMDL. Accessed October 24, 2017 at: https://www.epa.gov/tmdl/impaired-waters-and-stormwater

USGS (U.S. Geological Survey). 2001. *Herbicide Use in the Management of Roadside Vegetation, Western Oregon, 1999-2000: Effects on the Water Quality of Nearby Streams*. Oregon Department of Transportation. Report 01-4065.

USGS (U.S. Geological Survey). 2015. *National Field Manual for the Collection of Water-Quality Data*. October.

Washington Stormwater Center. 2008. *Guidance for Evaluating Emerging Stormwater Treatment Technologies – Modification: Evaluating Stormwater Treatment Technologies with Long Detention Times*. November.

WMO (World Meteorological Organization). 2012. *Guide to Meteorological Instruments and Methods of Observation*. 2008 edition, updated in 2010.

WSDOT (Washington State Department of Transportation). 2007. *Untreated Highway Runoff in Western Washington*. Prepared by Herrera Environmental Consultants, Inc. May.

WSDOT (Washington State Department of Transportation). 2014. *Quality Assurance Project Plan for Baseline Monitoring of WSDOT Highway Runoff. Prepared by Stormwater and Watersheds Program*. March.

WSDOT (Washington State Department of Transportation). 2014a. *WSDOT NPDES Municipal Stormwater Permit Highway Runoff Characterization Report (S7.B) Water Year 2013*. Prepared by Stormwater and Watersheds Program. October.

Zar, Jerrold H. 1999. *Biostatistical Analysis (Fourth Edition).* Pearson Education.

Zawlocki, K.R., Feguson, J.F., and Mar, B.W., 1980. *Survey of Trace Organics in Highway Runoff in Seattle, Washington*. WA-RD-39.9. Report to Washington DOT by the Department of Civil Engineering, University of Washington, Seattle.

APPENDIX A – SITE MAINTENANCE

All storm sampling equipment needs to be periodically maintained to ensure that it stays in optimal working condition. This includes checking the equipment for proper operation and looking for signs of wear or damage. Regular testing and calibration is required. In general, maintenance needs are typically higher for more complex set-ups. As a minimum, all equipment should be inspected and cleaned prior to each sampling event, especially following a major storm when the likelihood of damage from large debris is much higher. Below is a description of the various maintenance needs for different types of monitoring equipment.

- Primary Measurement Devices

    - Weirs/Flumes - Any sediment or debris that has accumulated behind the weir or is trapped on the crest, or within the flume should be cleaned out, as this blockage can invalidate the head-discharge relationship (Caltrans, 2015).

- Secondary Measurement Devices (Flow Sensors)

    - Pressure transducers: can be affected by conditions between storms, contaminants in the water, debris, bio-fouling, and cold temperatures. Typical maintenance of pressure transducers includes regularly cleaning them with mild soap and water, checking to make sure the O-ring is in place and not damaged, putting grease on the inside of the screw cap to maintain a water-tight seal, and performing periodic tests of the transducer to make sure it is measuring water depth accurately (this can be done in a bucket of water). Pressure transducers should never be exposed to freezing conditions.

    - Bubblers: bubbler tubing should be inspected periodically for damage, particularly the tip of the tube which can become clogged with sediment or algae, leading to erroneous measurements. If clogged, the tip should be cut off. If clogging is a continuing problem, the bubbler tube may be replaced with a larger inside diameter tube. Some flow meters and automated samplers have an internal purging mechanism to periodically clear the bubbler tube with a blast of air from the pump (Caltrans, 2015).

    - Area-Velocity (AV) Sensors: the accuracy of an AV sensor is potentially affected by conditions between storms, contaminants in the water, debris, and freezing conditions. During higher flows, the sensor may be damaged by heavy debris bouncing along the conveyance. Frequent inspection may be required, especially during periods of dry weather, as well as following periods of high flow (Caltrans, 2015).

- Desiccant - Maintaining desiccant is important as system malfunctions can occur due to high moisture levels inside the equipment. Some flow sensors and automated samplers contain moisture indicators that change color when the desiccant becomes saturated. Dry desiccant is generally brighter in color and turns darker when saturated, at which point it should be replaced with new desiccant. While desiccant can be re-activated in an oven, this is not recommended as the desiccant loses its absorption capacity each time it is re-activated, and most organizations are not equipped with an oven.

- Automated Samplers - Automated samplers should be checked regularly for wear and tear and cleanliness, including the suction line, the peristaltic pump tubing, and the strainer. These components should also be checked to ensure they are securely fastened. If damaged or clogged, they should be replaced. Some level of sediment or debris inside the suction line is inevitable. For minor amounts, pumping phosphate-free detergent and distilled water through the line is sufficient to clean the line. If larger amounts are present, the suction line should be replaced. In addition, the sample container(s) inside the automated sampler should be kept clean.

- Precipitation Gauges - Precipitation gauges require regular cleaning to ensure they remain free of bird droppings, leaves, or other materials. They should also be calibrated prior to each monitoring season, or more often if anomalous readings are observed (follow manufacturer directions). Precipitation gauges should also be checked periodically to make sure they are level (Caltrans, 2015).

- Power Supply - Relatively little maintenance is needed if AC power is being used, aside from checking to ensure that the power cord remains plugged into the outlet and is free of kinks or damage. For external DC batteries, all connections and wiring should be inspected regularly to ensure that all powered equipment are receiving adequate power. In addition, batteries should be checked for remaining capacity with a voltmeter. Drained batteries should be replaced immediately with fully charged ones.

APPENDIX B – CALIBRATION AND TESTING OF EQUIPMENT

Monitoring stations should be calibrated prior to the start of monitoring. Equipment calibrations may also be required during maintenance visits. Calibration and testing of equipment should be completed to the following components of a monitoring station:

- ❖ Automated Sampler
- ❖ Flow Meter
- ❖ Precipitation Gauge

<u>Automated Sampler:</u>

The sampler should be calibrated after installation of the sampler in the field. The automated sampler should be calibrated according to manufacturer specifications to collect the desired sample aliquot. If the sampler volume varies significantly from programmed values, then the automated sampler should be checked for proper installation. Proper installation includes making sure the suction line slopes downhill continuously and drains completely after each sampling round. Additionally, the length of the suction line should match the programmed suction line length. Excessive wear to the sampler pump can also produce inaccurate sample volumes.

The automated sampler may be tested prior to field deployment to ensure the pump and pump arm are working correctly. To test prior to field deployment, insert the suction line into a bucket of water and activate the sampler to take a specific volume of water. At the end of the test, check the sample bottles for the desired sample volume.

<u>Flow Meter:</u>

It is recommended that the flow meter be tested prior to field deployment. Testing the flow meter can be completed by placing the sensor in a pre-determined depth of water and programming the sensor to the specified depth. If the flow meter is working properly, then the readings will match fairly close to the measured depth of water in the bucket.

Flow meter calibration should occur after field installation. The flow meter can be calibrated to measure level at any depth, as specified by the manufacturer. To calibrate after field installation, enter the water level measurement, measured by the field team, into the meter module.

<u>Precipitation Gauge:</u>

All gauges should be calibrated by the manufacturer or according to manufacturer specifications prior to field deployment. Precipitation gauges can be tested prior to field deployment by manually tipping the bucket a number of times, ensuring that each tip is being recorded and that the tilting mechanism is operating freely (Hydrologic Services, 2014). Improper installation of the precipitation gauge can lead to erroneous data. Precipitation gauges should be installed on level ground, protected from strong winds, away from large obstructions such as trees and buildings, and checked regularly for debris.

APPENDIX C – LABORATORY METHODS

Laboratory analytical methods and detection limits should always be reviewed at the outset of a stormwater monitoring study. This should be done in consultation with the certified laboratory. It is particularly important to check the detection limits which the laboratory can attain for each specific water quality indicator. Detection limits that are too high will result in very imprecise data and many non-detects, which limits the types of data analyses that can be performed. The recommended analytical methods for some common monitoring constituents are listed in the table below. For more information, see CFR Title 40, Part 136 - *Guidelines Establishing Test Procedures for the Analysis of Pollutants*.

Water Quality Indicator	Preservative	Sample Holding Time	Type of Container[1]	Recommended Analytical Method	Detection Limit
Bacterial Tests					
Coliform, total, fecal, and *E. coli*	Cool, <10 °C	8 hours	PA, G	SM 9221E/SM 9221B	2 MPN /100 mL
Enterococci	Cool, <10 °C, 0.0008% $Na_2S_2O_3$	8 hours	PA, G	SM 9230B	1 MPN /100 mL
Physical and Aggregate Tests					
Chloride	None Required	28 days	P, FP, G	300.0, Rev 2.1 (1993) and 300.1-1, Rev 1.0 (1997)	0.004 mg/L
Kjeldahl and organic N	Cool, ≤6 °C, H_2SO_4 to pH <2	28 days	P, FP, G	SM 4500-N or B-1997 or C-1997 and 4500-NH3B-1997	0.1 mg/L
pH	N/A	Field/Immediate	N/A	EPA 150.2; SM 4500-H+B-2000	±0.1 (pH units)
Residue, Nonfilterable (TSS)	Cool, ≤6 °C	7 days	P, FP, G	SM 2540D-1997	1 mg/L
Specific conductance	Cool, ≤6 °C	28 days	P, FP, G	SM 2510B	±0.1 us/cm
Hardness as $CaCO_3$	HNO_3 to pH <2, Store at 4°C	6 months	P, G	EPA Method 200.7	1 mg/L
VSS	Cool, ≤6 °C	7 days	P, G	EPA Method 160.7	1 mg/L

Water Quality Indicator	Preservative	Sample Holding Time	Type of Container[1]	Recommended Analytical Method	Detection Limit
Particle Size Distribution	Cool, ≤4 °C	48 hours	P	Coulter Counter, Laser diffraction, or comparable method	N/A
TDS	Cool, ≤6 °C	7 days	P, G	SM 2540C-1997	1 mg/L
Temperature	None Required.	Analyze	P, FP, G	SM 2550B	±0.1 °C
Turbidity	Cool, ≤6 °C	48 hours	P, FP, G	EPA 180.1; SM 2130B-2001	0.5 NTU
Metals					
Aluminum	HNO_3 to pH <2, Store at 4°C	6 months	P, FP, G	EPA 200.8	25 µg/L
Cd, Cu, Pb, Zn				EPA 200.8	1 µg/L
Nutrients					
Nitrate	Cool, ≤6 °C	48 hours	P, FP, G	EPA 300.0, Rev 2.1 (1993) and 300.1-1, Rev 1.0 (1997); SM 500-NO3-E-2000	0.1 mg/L
Nitrate-Nitrite	Cool, ≤6 °C, H_2SO_4 to pH <2	28 days	P, FP, G	EPA 350.2; EPA 350.3	0.1 mg/L
Nitrite	Cool, ≤6 °C	28 days	P, FP, G	EPA 300.0, Rev 2.1 (1993) and 300.1-1, Rev 1.0 (1997)	0.1 mg/L
Orthophosphate	Cool, to ≤6 °C	Filter within 15 minutes; Analyze within 48 hours.	P, FP, G	EPA 365.3; 300.1-1, Rev 1.0 (1997)	0.03 mg/L
Oxygen, Dissolved Probe	None Required.	Analyze within 15 minutes.	G, Bottle and top	SM 4500-O G-2001	None
Phenols	Cool, ≤6 °C, H_2SO_4 to pH <2	28 days	G	EPA 420.1[1](Rev. 1978)	2.5 µg/L
Phosphorus, Total	Cool, ≤6 °C, H_2SO_4 to pH <2	28 days	P, FP, G	EPA 365.2	0.03 mg/L

Water Quality Indicator	Preservative	Sample Holding Time	Type of Container[1]	Recommended Analytical Method	Detection Limit
Organic					
Benzidines	Cool, ≤6 °C, 0.008% $Na_2S_2O_3$[5]	7 days until extraction	G, FP-lined cap	EPA 625/1625B	10 µg/L
PCBs	Cool, ≤6 °C	1 year until extraction, 1 year after extraction.	G, FP-lined cap	EPA 608/625	Not Determined
Polycyclic Aromatic Hydrocarbons	Cool, ≤6 °C, Sodium thiosulfate in presence of chlorine	extract - 7 days; analyze – 40 days	G	EPA 8310	0.05 µg/L
Herbicides					
Diuron/Glyphosate	Cool, ≤6 °	7 days	Glass-amber	EPA 8151A	Check with lab
Total Petroleum Hydrocarbons (TPH)					
TPH (Gasoline)	Cool, ≤6 HCL to pH < 2	14 days	VOA Vial	EPA 8015M; EPA 8260	50 µg/L
TPH (Diesel)	Cool, ≤6 HCL to pH < 2	extract - 7 days; analyze – 40 days	G	EPA 8015M	50 µg/L
TPH (Motor Oil)	Cool, ≤6 HCL to pH < 2	extract - 7 days; analyze – 40 days	G	EPA 8015M	50 µg/L
Oil and Grease	Cool to ≤6 °C, HCl or H_2SO_4 to pH <2	28 days	G	EPA 1664	5 mg/L

[1]"P" is polyethylene; "FP" is fluoropolymer (polytetrafluoroethylene (PTFE); Teflon®), or other fluoropolymer, unless stated otherwise in this Table II; "G" is glass; "PA" is any plastic that is made of a sterilizable material (polypropylene or other autoclavable plastic); "LDPE" is low density polyethylene (CFR 2017).