# Data Analysis Plan and Study Design

## Draft 10/2/15

Ecology RSMP Effectiveness Study
Stormwater Source Control Data Assessment
City of Lakewood and Cardno consultants

**INTRODUCTION**

This document provides a description of the anticipated data analysis plan for the Ecology Regional Stormwater Monitoring (RSMP) Source Control effectiveness study. The main activity of the study will result in an assessment of source control inspection data from NPDES permittees (Phase I and Phase II) throughout western Washington. The objective of the assessment is provided below along with a description of how each source control effectiveness question driving the study will be addressed.

Initial considerations for a data analysis design are proposed here based on the type, quality, and quantity of data thought to be available. This draft data analysis plan will be updated as data are obtained, reviewed, and as the project team becomes more familiar with the compiled dataset.

**OBJECTIVE OF DATA ASSESSMENT**

The source control effectiveness study will provide NPDES Municipal permittees with information from a regional data assessment of municipal business inspection program data. The analysis and results are intended to help NPDES jurisdictions improve inspection programs, especially ones that include inspection of stormwater source control effectiveness and best management practices (BMP). The assessment will come from analyzing available data in the context of addressing and answering the source control effectiveness questions developed by the Stormwater Work Group (SWG) for Puget Sound and NPDES permittee representatives on the Effectiveness subgroup. Question topics include inspection frequency of stormwater treatment and control facilities, effectiveness of focusing on property owners versus business owners, compliance rates, use of BMPs, and barriers to BMP implementation.

**SOURCE CONTROL EFFECTIVENESS QUESTIONS**

Eight effectiveness questions in total, related to source control at existing sites, were selected by the Effectiveness subgroup and approved by the SWG. The questions were prioritized from a much larger list of questions raised by permittees for consideration in developing effectiveness studies. The first seven source control questions are being addressed in this phase of the project (data assessment) and the last question (number 6 in the list below) is expected to be addressed during a later phase of the project intended to develop a coordinated inspection framework.

1. What is the optimum frequency of inspections to maintain the functionality of stormwater treatment and control facilities and ensure the proper use of source control best management practices (BMPs) at businesses?
2. Which is more effective for specific high value BMPs: focusing on the property owners or

focusing on the business owners, or a combination of the two?
    a. Target both structural and operational BMP types, and situations where a business owner is and is not cooperative and willing.
3. Which required BMPs were implemented based upon follow up inspection? Which optional BMPs were installed based upon follow up inspection?
4. What were the primary barriers to not adopting or installing BMPs?
5. Address the connection between in-person visits and source control BMPs, and identify situations where technical assistance and/or follow-up inspections are needed to ensure required BMPs are implemented.
    a. Gather data about percent compliance. Partner with Ecology Local Source Control program to do this study.
6. Are stormwater source control inspections more effective if combined with other types of inspections? How can coordination of inspections be improved or better organized regionally for referral of issues to the correct entity?

## DATA ACQUISITION

Data for this assessment will be solicited primarily from NPDES municipal permittees in western Washington. The solicitation will be by a data request letter, data variables list, and survey emailed to NDPES permit managers. In 2014, a preliminary survey and comment form was sent to a handful of permittees during the proposal stage of this project. The survey inquired about the availability of data that could be used to address the effectiveness questions, existing data analysis efforts that jurisdictions may have already done using their own data, and an open question asking for suggestions to improve or add to the study design of compiling and analyzing regional data. Information in this data analysis design document incorporates answers and information that permittees provided in response to that proposal-stage survey.

A ubiquitous comment by permittees in that survey indicated that for some questions existing data could be used as-is (e.g. dates of inspections to calculate inspection frequency) and for other questions, data would need to be combined, parsed, qualified, or quantified (e.g. business owners versus property owners in context of high value BMPs). The amount of data preparation needed will be determined after receiving data and is likely to depend on data condition, availability of municipal staff, and consideration of the scope and budget for this assessment. The extent to which the jurisdictions and/or the project team will do such data preparation is unknown at this time and will likely be discussed with each jurisdiction based on data condition.

In addition to data from NPDES municipal permittees, relevant data sources are also available from Ecology's Local Source Control program and possibly from non-governmental organizations, such as Environmental Coalition of South Seattle (ECOSS). Data from these entities will also be requested and prepared as needed for the assessment.

## DATA ANALYSIS OVERVIEW

The data analysis design for the assessment comprises compiling and analyzing data from western Washington NPDES permittees' existing businesses and commercial property inspection programs. Inspections of stormwater facilities at businesses and on commercial properties has been a required part of each permittee's Stormwater Management Program since the previous municipal permit period

began in 2007. Thus, we expect some amount of data to be available from most primary permittees and possibly some secondary permittees. We also expect several robust data sources to be available from jurisdictions that have long-standing programs, have relatively large jurisdictions to manage, and/or have relatively more sophisticated programs and resources available. Indeed, some jurisdictions have already performed some level of data analysis or assessment (personal communications, L. Tomchick, King County Envirostars and E. Stewart, City of Seattle) and representatives from some of those jurisdictions will be part of the project advisory committee.

As a post-hoc evaluation of existing data, it is recognized that the extent of the data analysis used to address some questions may be limited because of the availability of data. Some elements of a conventional data analysis design are not included here, such as data quality objectives, because the data for this study is observational; no experiments are being done with treatments and controls where data collection is included.

In order to design the data assessment to address the effectiveness questions, the following section provides an explanation for each of the effectiveness questions, the planned steps to prepare the data received from permittees, and the initial evaluation methods. For many of the effectiveness questions the same data will be evaluated, with the main differences attributed to how the data are parsed. Thus, there are overlaps in the types of analysis methods that may be used to address the effectiveness questions given the data. An overview of these specific types of analysis methods is provided in the last section of this document.

1. **What is the optimum frequency of inspections to maintain the functionality of stormwater treatment and control facilities and ensure the proper use of source control best management practices (BMPs) at businesses?**

   In the context of stormwater BMP inspections, frequency is understood to mean the average time between inspections expressed in units such as inspections per quarter or per year. For some BMPs or businesses, a simple frequency of inspection dates at a business may be appropriate, such as for routine catch basin cleaning and inspection. But for other BMPs or certain types of businesses, inspection frequency can be dictated by factors other than routine maintenance, including risk or potential of pollution, complexity of inspection, construction activities on property, industrial processes and discharges, public outreach efforts, spills/illicit discharges, and special BMP maintenance, among others.

   It is expected that many data sources will include multiple inspections at a given business for a variety of reasons. A simple frequency calculation may not distinguish what drives an inspection or how the frequency relates to stormwater BMP functionality. Thus, data used to address this question will be assigned a reason code to represent the type of inspection in order to compare similar inspections across jurisdictions and calculate frequencies of inspections related to BMP functionality. This will improve the comparability of data sources, and the evaluation of inspection frequency may result in several types of frequencies calculated for different types of BMPs or businesses, or other factors that influence inspections.

   For purposes of this data assessment, the term inspection is generally considered to be an on-site visual evaluation. However, communications, reporting, and desk research before and after on-site inspections contribute to the resolution of issues and desired compliance. Thus, it is expected that some data sources will include the dates of contact and correspondence as part of the record leading to compliance. It may be useful to consider some of these activities to inform

the determination of inspection frequency for the complete inspection cycle, which may include multiple onsite inspections or other activities related to BMP functionality, such as initial inspection, education and outreach, follow-up inspection, and compliance status evaluation.

In addition, there are two parts to the question, with the first part asking about maintaining BMP functionality and the second part asking about ensuring proper BMP usage. BMP functionality is related to how a BMP works, its effectiveness, and if it's working correctly. BMP usage is related to how a BMP is operated and maintained. The extent to which this distinction can be evaluated will be based on the available data and the types of BMPs included in the data sources.

Finally, it is understood that "at businesses" applies to both parts of the question and the intent is to inform inspections of BMPs located at businesses where activities are covered under the NDPES Municipal permit. In this context, "at businesses" is not understood to refer to large quantity generators or other entities with individual NPDES discharge permits, including construction sites, boat yards, and industrial businesses with permitted discharges.

Data variables needed to calculate frequencies of inspections include: date of inspection; date of follow-up activities; type of inspection; reason for inspection; type of stormwater BMPs inspected; the level of BMP evaluation (functionality and/or usage); date of non-compliance, compliance, or similar benchmark parameter; size and type of business.

Data evaluation is expected to include calculation of inspection frequencies for a variety of inspection activities, business types, BMP types, and reasons motivating the inspection. Data will be graphed for visual comparison and basic trend analysis. Statistical evaluation of inspection frequencies may include summary statistics such as averages, minima, and maxima. Inspection frequencies may also be evaluated with respect to compliance rates. Thus, data evaluation is expected to include a summary of associations between inspection frequencies and compliance rates.

**2. Which is more effective for specific high value BMPs: focusing on the property owners or focusing on the business owners, or a combination of the two?**

**a. Target both structural and operational BMP types, and situations where a business owner is and is not cooperative and willing.**

Question two refers to high value BMPs and distinguishes between property owners and business owners. The second part of the question is a clarifying statement to address both structural and operational BMPs and business managers or property owners who are not cooperative. "High value BMPs" is understood to refer to those BMPs that have the greatest positive effect on stormwater source control and treatment. For some businesses, this may be a structural BMP such as catch basins or expensive proprietary treatment technology, and for other businesses a high value BMP may be an operational BMP such as material transfer or equipment maintenance. Thus, data needed to address high value BMPs should include the type of BMP and the value provided by the BMP (specific treatment or function). The data review and preparation will include identifying or assigning appropriate BMP types and values to permittee data as needed.

Because the high value BMPs in question are presumably located on commercial property, their effective operation may be affected by the property owner or business owner and the type of

business activities. The distinction between property owners and business owners is important as some municipal inspectors have found that working with one or the other party or a combination of both is preferable at some businesses or properties. The data needed to evaluate this distinction includes knowing if the contact person for an inspection was the owner of the business, property, both, or neither. In addition, the data records would need to indicate if communicating with one party or the other was more effective. The extent of the data evaluation for this question will largely depend on whether or not inspection data includes these distinctions: the ability to distinguish between high value (and not high value) BMPs from inspection data, and if the contact person was the property owner or business owner.

It is expected that most existing data sources will not include these fine point distinctions about BMPs and contact persons. Parameters to distinguish these factors will be identified or added to the data set as possible. Following the clarifying statement of the question, it may also be possible to identify (or have jurisdictions identify) and evaluate a subset of inspections data where the business was not cooperative. If such a subset of data is available, it may be possible to compare the effectiveness and functionality of a variety of BMPs, including those that are high value, structural, and operational.

Data variables needed to evaluate the effectiveness of focusing on property owners versus business owners for high value BMPs include: inspection frequency variables from Question 1, value of BMPs inspected, role of the contact person at the business, and information on how effective the communication was especially in context of high value BMPs.

Data evaluation is expected to be similar to that for Question 1, with added comparisons based on the values of various BMPs, compliance rates of high value BMPs among business types, and comparison of some parameter that represents effective communication among the different roles of contact persons at businesses.

**3. Which required BMPs were implemented based upon follow-up inspection? Which optional BMPs were installed based upon follow-up inspection?**

Question three refers to BMP implementation after a follow-up inspection and distinguishes between required and optional BMPs. Follow-up inspection is understood to refer to inspection activities related to issues identified during previous inspections. In this context, BMPs implemented “based on” follow-up inspections is understood to mean simply after a follow-up inspection and not necessarily *caused* by the follow-up inspection. In practice, follow-up activities only sometimes include revisiting a site to perform onsite inspection. Common follow-up activities may include emails, calls, and letter correspondence to ask for verbal or written confirmation of BMP implementation. While it may be possible to evaluate data of a variety of follow-up activities depending on available data, the question is inquiring about follow-up inspections. Thus, the data analysis for this question will consider just those records that include, among other activities, follow-up inspections linked to issues identified during a previous inspection.

The question also asks about required versus optional BMPs. In the context of municipal NPDES-driven inspections, required BMPs are those that are needed to treat or control stormwater at the source and any potential pollutants it may come in contact with. The options for BMPs are typically selected from the Stormwater Management Manual for Western Washington (Ecology 2012, especially volume III (Flow Control BMPs) and volume IV (Source Control BMPs)), from

local ordinances and engineering standards, and some by state and federal regulations, such as the transportation and transfer of hazardous materials. The list of required BMPs varies according the business activities, the potentially hazardous materials used, and a business' waste generator status (small, medium, or large quantity generator).

An additional element to this question and other to be addressed this assessment is the concept of compliance versus non-compliance for designating the status of a BMP or a business overall. While overall compliance with the NPDES permit is, presumably, the goal and requirement of every permittee, the concept of compliance in the context of stormwater and environmental inspections at businesses is used and emphasized differently in different settings. Many jurisdictions emphasize outreach, education, incentives, and technical assistance for using BMPs correctly and deemphasize a strictly compliance oriented approach for routine inspections. However, there is a time and place for compliance focus and enforcement actions and it is expected that data sources will include both approaches. Regardless of a jurisdiction's emphasis on compliance, a working presumption for the data analysis design is that issues identified during previous inspections typically trigger follow-up inspections. Thus a compliance status or otherwise articulated outcome of an inspection cycle is necessary to define different inspections on the cycle, such as initial, follow-up, spill response, etc.

It is expected that most existing data sources will include some degree of distinguishing between initial and follow-up inspections, and a desired compliance status or inspection outcome. Jurisdictions will be asked how they define compliance and link follow-up inspection data to specific issues in order to complement this information. The extent to which the data sources will have this link is unknown. It may be possible to add fields to data records to represent the type of inspection (screening, initial, follow-up, spill response, etc.) and associated compliance status.

Likewise, it is not known to what extent the data sources will explicitly note if BMPs were required or optional. It may be possible to code data records with this information based on the BMPs inspected; this would require assigning required versus optional values to the data in the context of applicable regulations and inspection history.

Data variables needed to evaluate whether required or optional BMPs were implemented after a follow-up inspection are: date of inspection; date of follow-up activities; type of inspection (initial or follow-up); the BMP requirements (required or optional); and many of the inspection frequency variables from Question 1. Data evaluation is expected to summarize compliance rates by BMP type for records that include follow-up inspections.

**4. What were the primary barriers to not adopting or installing BMPs?**

Question four inquires about barriers to using BMPs. This question is understood to refer to the stormwater BMPs inspected as part of municipal inspections and includes operational, structural, high value, and other BMPs evaluated in other questions of the assessment. It is expected that data needed to evaluate the question of barriers to adopting or using BMPs would be mostly in descriptive form via written or anecdotal accounts from inspection staff. It is unknown whether the data sources will include comments on barriers to BMP use. Assuming some data regarding BMP barriers is available, this question will be evaluated by identifying if BMPs were used or not and by assigning reason codes or descriptions of the barriers. Reasons given for barriers to BMP use within and across jurisdictions will be summarized. Examples of

BMP barriers are expected to include cost, complexity of installation or operation, English comprehension, clarity of communication by municipal inspectors about the necessity of BMPs, attitudes of businesses toward environmental protection and government, and procrastination.

Data variables needed to evaluate barriers to adopting or installing BMPs include: date of inspection; date of follow-up activities; type of inspection; reason for inspection; type of stormwater BMP; the level of BMP evaluation (functionality and/or usage); date of non-compliance, compliance, or similar benchmark parameter; size and type of business; and comments and anecdotal information from inspectors. Data evaluation is expected to summarize this data.

**5. Address the connection between in-person visits and source control BMPs, and identify situations where technical assistance and/or follow-up inspections are needed to ensure required BMPs are implemented.**

**a. Gather data about percent compliance. Partner with Ecology Local Source Control program to do this study.**

Question five asks about the connection between in-person visits (i.e. inspections) and (stormwater) source control BMPs. Specifically, the question refers to situations that require technical assistance or follow-up inspections to make sure BMPs that are required are used. Additionally, the question includes instructional comments about using data on percent compliance and about including data from the statewide Ecology Local Source Control program.

This question requires similar considerations mentioned above under questions one and three about defining compliance or a desired outcome of BMP usage, and about the inspection cycle and what triggers a follow-up inspection. Also, the reference to technical assistance touches on some of the non-inspection activities mentioned above that are often part of follow-up actions. Technical assistance is understood to include education and outreach, connecting with existing resources and information, financial assistance and incentives, and BMP operational assistance, among other activities. In this way, technical assistance is a type of non-structural BMP itself that can help achieve compliance with stormwater source control requirements, both on an individual business level and on a jurisdictional level for overall permit compliance.

The concept of a connection between onsite inspections and BMP usage is understood to refer to the elements of inspections and inspection programs that lead to positive environmental outcomes. It is expected that data sources do not include explicit data addressing a connection between inspections and BMP usage. Instead, the assessment for this question could include discussing the results of the analysis of the other questions in the context of business types, specific BMPs, or inspection techniques that shed light on how inspection and follow-up actions specifically may lead to compliance. Thus, no specific data analysis is anticipated for question five, rather it will be addressed in the discussion of the results of the analysis of the other source control effectiveness questions.

The connection for partnering with the Ecology Local Source Control program has been made, and Jane Dewell who is the lead staff for the program is on the project advisory committee.

## SPECIFIC DATA ANALYSIS METHODS

The raw compiled data set for the Stormwater Source Control Data Analysis will consist of categorical, quantitative and descriptive data. The majority of variables are expected to be categorical (qualitative). Some examples include jurisdiction, type of business, type of inspection, indication of compliance, types of follow-up actions, and type of BMP. The primary quantitative variables that may be used in the analysis are based on time, such as time since last inspection or number of inspections per year, and based on measured parameters associated with BMP maintenance, such as catch basin sediment depth or percent full associated with a clean-out schedule. As such, the primary analysis methods used to evaluate the effectiveness questions will be based on proportions or rates. Some proportions will be summarized directly from the data (e.g., proportions of types of BMPs), while others will be calculated based on other variables (e.g., compliance rates based on discretized frequencies of inspection). Qualitative data summaries specific to the effectiveness questions will be provided in the form of tables and graphics.

Depending on the available data, some effectiveness questions may also be evaluated using more formal statistical summaries, including confidence intervals and hypothesis tests that make comparisons across groups. For example, for question two we may be able to formally test for differences in compliance rates for high-valued BMPs by business type. Formal inference methods, however, require sufficient sample sizes in order to have adequate power to detect differences if they exist. Typical methods for inference for proportions may include z-tests for the difference between two population proportions, chi-square tests for homogeneity to compare a set of more than two proportions between two or more groups, or tests of independence for looking at the association between two quantitative variables.

Some other statistical methods that may be considered include: (a) logistic regression to model the probability of compliance given the time between inspections and (b) classification and regression trees (CART) to estimate the probabilities of compliance given various factors. CART methods may help identify the variables that are the best predictors for compliance. Logistic regression and CART both require relatively large sample sizes, so use of these methods will be highly dependent on available data.

## DATA ANALYSIS PLAN FINALIZATION AND NEXT STEPS

The initial data analysis design described above is proposed based on the type, quality, and quantity of data thought to be available. This design document will be updated as data are obtained, reviewed, and as the project team becomes more familiar with the compiled dataset. An Access® database will be created to store and query the compiled dataset in preparation for data analysis. In addition to a revised and updated version of the data analysis design, other project deliverables to be prepared leading up to the data analysis include the draft database, the compiled raw data provided by permittees, and a technical memo summarizing the process of reviewing and qualifying data and preparing the database. After the data analysis is completed per the scope of work, a draft report will be prepared for review by the project advisory committee, which is composed of permittee representatives. A final report and presentation of the project findings will be prepared as the last steps of this phase of the project.