Quality Assurance Project Plan

Assessment of Regional Project Effectiveness in Puget Sound

June 18, 2014

Publication Information

This project is funded by the U.S. Environmental Protection Agency's National Estuary Program (NEP) to the Puget Sound Partnership (PSP). The plan describes the objectives of the study and the procedures to be followed to achieve those objectives.

Data for this project are curated by the entities, agencies and jurisdictions that are responsible for collecting, reporting, and entering the data. No data were collected by the Puget Sound Partnership for this project.

The contents of this document do not necessarily reflect the views and policies of the EPA, nor does mention of trade names or commercial products constitute endorsement or recommendation for use. Copies of this Quality Assurance Project Plan (QAPP) and the final project report will be available on request from the PSP.

Author and Contact Information

Leska Fore
Puget Sound Partnership
326 East D Street,
Tacoma, WA 98421
206-708-5048
Leska.fore@psp.wa.gov

1.0 Title Page, Table of Contents, and Distribution List

Quality Assurance Project Plan

Assessment of Regional Project Effectiveness in Puget Sound

May 2014

Approved by:

Signature _____ Date 6/4/14

Ken Dzinbal, Senior Monitoring Program Manager, PSP

Signature _____ Date 6-18-14

Heather Kibbey, PSEMP Steering Committee

Signature _____ Date 6/2/14

Alana Knaster, Deputy Director, PSP

Signature _____ Date 6/2/14

Tracy Collier, Science Director, PSP

Signature _____ Date 5/28/14

Bill Kammin, Ecology QA Officer

PSP = Puget Sound Partnership
PSEMP = Puget Sound Ecosystem Monitoring Program

Table of Contents

Distribution List

Heather Kibbey
Chair, PSEMP Steering Committee
City of Everett, WA
3200 Cedar Street, Everett, WA 98201
425.257.8800
hkibbey@everettwa.gov

Ken Dzinbal
Senior Monitoring Program Manager
Staff to PSEMP Steering Committee
Puget Sound Partnership
326 East D Street, Tacoma, WA 98421
360.464.1222
ken.dzinbal@psp.wa.gov

Angela Bonafaci
Performance Management, Puget Sound Team Lead
EPA, Region 10
1200 Sixth Avenue, Seattle, WA 98101
206.553.0332
bonifaci.angela@epa.gov

Leska Fore
Monitoring Program Coordinator
Puget Sound Partnership
326 East D Street, Tacoma, WA 98421
206.708.5048
leska.fore@psp.wa.gov

Constance Sullivan
Research Scientist
Puget Sound Institute, University of Washington
326 East D Street, Tacoma, WA 98421
253.254.7030
casull@uw.edu

Haley Harguth
WA Sea Grant Marc Hershman Marine Policy Fellow
Puget Sound Partnership
326 East D Street, Tacoma, WA 98421
253.666.6289
haley.harguth@psp.wa.gov

Tom Gries, NEP QA Coordinator
Washington Department of Ecology
300 Desmond Drive SE
Lacey, WA 98503
360-407-6327
Tgri461@ecy.wa.gov

PSEMP Steering Committee email list

2.0 Abstract

The goal of this study is to summarize the effectiveness of selected ecosystem recovery projects and compare the relative impact of restoration actions across the Puget Sound basin. This study is part of a planned larger framework for effectiveness monitoring and will demonstrate how project level data may be rolled up for watershed or regional assessment. The Puget Sound Partnership and Puget Sound Institute (University of Washington, Tacoma) will identify studies that have collected data before and after restoration or management actions that are designed to improve site condition. Using existing data, we will calculate summary statistics derived from published sources. New summary statistics will be calculated to compare the effectiveness of different types of studies. We will also evaluate the responsiveness of various indicators of change for projects across the Puget Sound region. We will use statistical meta-analysis and power analysis to evaluate data sets. Meta-analysis is a formal statistical analysis used to compare the results of a number of empirical studies that have tested, or can be used to test, the same hypothesis. Meta-analysis will allow us to calculate the mean response to experimental treatment across studies and to discover key variables that may explain any inconsistencies in the results of different studies (Harrison, 2011). We will evaluate and comment on the meta-analysis methods and their relevance for restoration and management programs in Puget Sound. Power analysis evaluates the sensitivity of a variable's response (Gurevitch and Hedges, 1999). Characterization and analysis will include calculation of description statistics and analysis across projects using meta-analysis methods, e.g., Cohen's *d* as a test statistic (DeCoster, 2009).

Data will be compiled from a variety of sources, such as published TMDL effectiveness studies completed by the Washington Department of Ecology, Kitsap and Island Counties. Other data sources include measurements of toxic compounds in the environment and in organisms before and after a management response. Toxic compounds may include pesticides in water or metals or organics in organism tissue. An additional component of this study will be a power analysis and statistical evaluation of existing water quality and shoreline measurement data collected by partner agencies and volunteer organizations.

3.0 Background

Effectiveness Monitoring Framework

The Puget Sound Action Agenda is a framework developed by regional partners to prioritize the most important actions for restoring Puget Sound (PSP, 2013). The Action Agenda defines measurable goals for 21 indicators and includes 70 substrategies and plans for over 200 actions. How do we measure whether the actions are working?

Regional monitoring programs recognize the emerging importance of effectiveness monitoring (Whiteway et al., 2010; Bowler et al., 2012). According to the Puget Sound Nearshore Ecosystem Restoration Project (PSNERP), effectiveness monitoring evaluates whether or not restoration or management actions are achieving their stated objectives and is designed to answer the question, "Did the actions achieve their intended outcomes?" (Brandon et al., 2013).

PSEMP will develop an Effectiveness Monitoring Framework that builds primarily from the Puget Sound Partnership's Action Agenda as the key starting point for identifying current, relevant effectiveness questions. The framework will:

1) List the overall goals and objectives for monitoring the effectiveness of actions and recovery efforts in Puget Sound,
2) Identify the specific questions associated with priority recovery strategies, initiatives, and actions described in the PSP Action Agenda, and
3) Relate those questions to existing monitoring efforts or programs, identify data streams and reporting vehicles, and/or identify gaps in effectiveness monitoring, data assessment, and reporting associated with the Action Agenda.

The statistical meta-analysis proposed is part of a larger effort by the Puget Sound Partnership to develop a regional framework for effectiveness monitoring a) leverages existing programs and resources, b) recommends opportunities to improve monitoring focus and design, data sharing, and combined analysis, and c) identifies opportunities and pathways to improve transfer of information from monitoring programs into policy and adaptive management processes.

4.0 Project Description

Many projects and organizations have implemented management and restoration actions that are designed to improve the condition of Puget Sound (Puget Sound Partnership Project Atlas; Bash and Ryan, 2002). Since the 1980s, hundreds of projects have been completed (Figure 1), and for many of these projects data have been collected that measure the condition of the project sites before and after actions were completed. Some of the projects have been evaluated for their impact on water quality, habitat or biota. Other projects have been monitored, but change over time has not been assessed. Still other projects were only monitored initially when projects were implemented, but not since. No studies have looked at the collective impact of individual projects across the Puget Sound watershed.

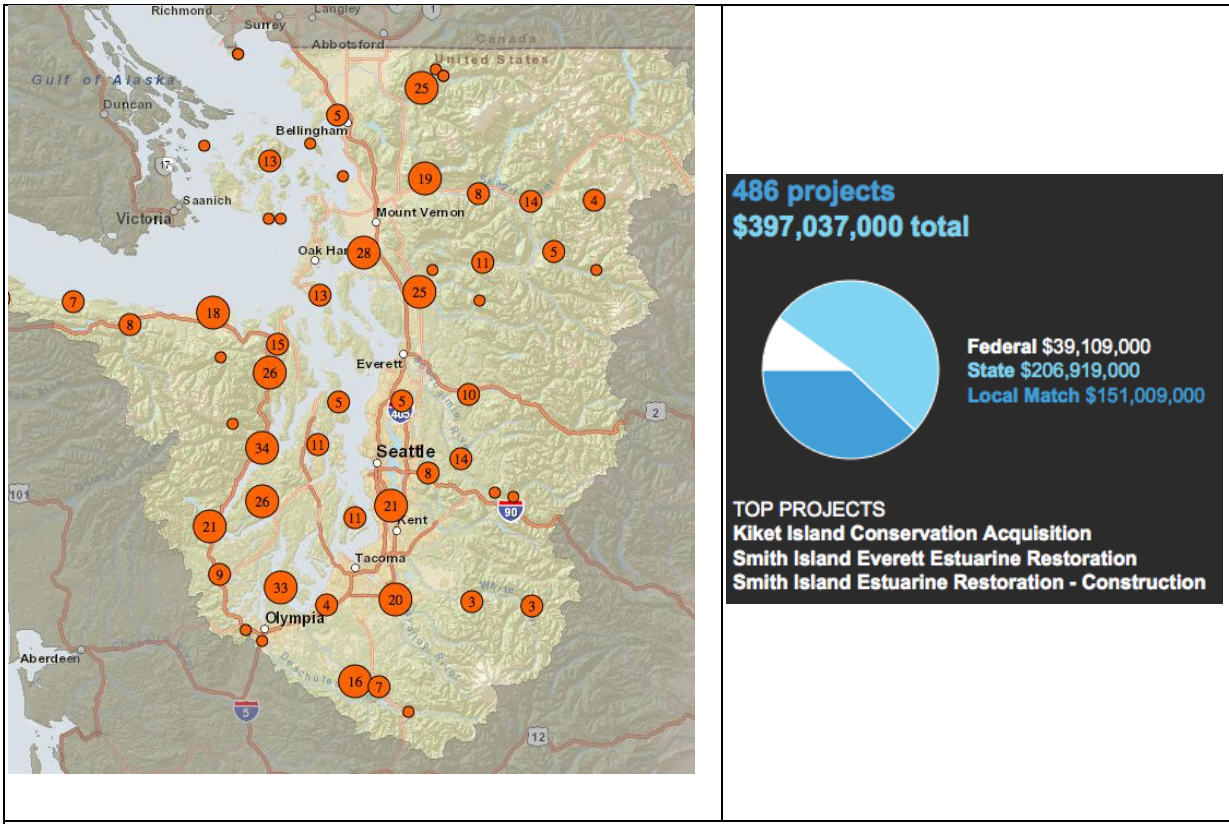

Figure 1. Restoration projects in Puget Sound (PSP Project Atlas).

4.1 Project Goals

The goal of the current work is to evaluate how measurements at the project scale can be summarized at the watershed or regional scale. This will allow users to assess the effectiveness of different types of projects and the sensitivity of various measures to detect change associated with restoration or management actions.

Two types of analyses will be used in this project: 1) Meta-analysis of project level data to assess the impact of individual projects at a regional level and 2) statistical power analysis to evaluate existing data sets to determine how much change in condition is detectable over time.

4.2 Project Objectives

The objectives for this study are to (DeCoster, 2009):

- Define the relationship of interest to be the effect of restoration and management actions on the recovery or improvement of Puget Sound habitat and water quality affecting stream and nearshore biota.
- Gather data from relevant sources with before and after comparisons after recovery treatments or actions.
- Collect data across studies into a single data file, calculate statistics for comparison, record details of studies related to location, duration of treatment, types of treatment or other relevant factors.

- Compare the effectiveness of different types of projects; identify any factors that are associated with more effective projects, and identify indicators that are the most sensitive to change.
- Vet results with project sponsors and local experts.

This initial meta-analysis will cover a broad range of variables, management actions, locations, and scales (temporal and spatial). Variables will include water quality parameters (dissolved oxygen, nutrients, and fecal coliform bacteria), organic contaminants in biota, pesticides in water, habitat restoration, and biological effects of organic contaminants. Management actions will include TMDL implementations, which involve a variety of actions such as BMPs, treatment of waterways (i.e., alum treatments to counteract high levels of phosphorous), and diversion of runoff. Other management actions will involve full or partial bans on chemicals, habitat restoration, capping of a contaminated site, clean up of a contaminated site, and PIC programs.

5.0 Schedule and Organization

Task Description and Schedule

This project is supported by a grant from EPA and includes tasks in addition to the statistical analysis. PSP has contracted PSI to help with the statistical analysis and other aspects of this project (Table 1).

This project integrates across three sections of the Puget Sound Partnership: Science, Performance Management, and Soundwide and Functional Programs (Planning) (Table 2). The science team will lead on gathering data and meta-analysis. Performance management and Soundwide Programs will play an advisory role with interpretation of data analysis and connecting results to appropriate audiences. Performance management will support the project with GIS analysis as needed.

Table 1.Tasks contracted with Puget Sound Institute, deliverables, and proposed due dates.

Tasks are numbered consistent with PSP's contract with PSI.

Task Description	Deliverables	Due Date
9. Support development of PSEMP effectiveness monitoring framework		
9a. PSI will provide staff to work under the direction of the PSP Project Leader and coordinate with others as needed to: **1)** Identify specific monitoring questions associated with priority recovery strategies, initiatives, and actions described in the PSP Action Agenda (NOTE: the Action Agenda and 2012/2013 Appendices describe numerous strategies, sub-strategies, near-term actions, performance measures, and responsible parties. Lacking (in many but not all cases), are clearly articulated and specific monitoring	An expanded table (e.g. Appendix C to the 2012/2013 Action Agenda) delivered to PSEMP staff at PSP listing the specific monitoring questions and known data sources/reporting mechanisms associated with line-items in the Action Agenda. Note – the final table format will be developed as part of this task.	April 30, 2014

Task Description	Deliverables	Due Date
questions and indicators, or clearly identified data sources, to guide the development or organization of an overall monitoring and reporting program that can inform managers with regard to the success or effectiveness of those actions. **2)** Relate those questions to existing monitoring efforts or programs, and identify data streams and reporting vehicles.		
9b. Select a sub-set of priority Strategies or Actions that have measureable outcomes and that could potentially be evaluated for their effectiveness. Identify data sources that could contribute to an evaluation of the effectiveness of the selected NTAs. Conduct a meta-analysis of change in indicators for existing data associated with the selected sub-set.	• List of NTAs, programs and program elements with description of selection criteria • Description of potential data sets; downloads of data • Presentation for PSEMP and data authors' review	March 1, 2014
9c. • Work with partners to vet results of meta-analyses and with the Executive Oversight Committee's subcommittee to determine how results should be shared with external partners. • Coordinate with data authors such as agencies, organizations, and government units that are responsible for implementing and/or reporting on actions as relevant and necessary.	Revised analysis and interpretation to address comments, issues and concerns raised by partners	May 2014
9d. • Assist in the development of reports on the effectiveness monitoring framework • Document approach to develop an effectiveness monitoring prototype for a subset of NTAs or program elements. PSI provides: • Data sets relevant to NTAs and other related effectiveness monitoring efforts • Data file of statistics derived from projects assessing change over time • Report describing data sources, analysis and results	• Contributions to draft report delivered to PSEMP staff at PSP for review and approval • Contributions of revised material for final report delivered to PSEMP staff at PSP for review and approval • Final contributions delivered to PSEMP staff at PSP	June 27, 2014

Project Staff

Table 2. Project staff and clients and their responsibilities.

Project Team Members	Responsibilities
Leska Fore (206) 708-5048 Monitoring Program Coordinator, PSP	Project Lead. Prepares QAPP, reviews all products, statistical analysis, editor and writer of final report, engages regional experts
Constance Sullivan Puget Sound Institute (PSI)	Reviews QAPP, provides analysis, engages regional experts, writes final report, lead on peer-review article
Haley Harguth (360) 666-6289 Hershman Marine Policy Fellow	Statistical analysis of existing data for LIO partners
Ken Dzinbal (360) 464-1222 Monitoring Program Manager, PSP	Reviews and approves QAPP, reviews final report, engages regional experts
Tracy Collier (360) 464-2008 Science Director, PSP	Approves QAPP and Final report
Thomas Gries NEP Quality Assurance Coordinator Dept. of Ecology	Reviews and comments on drafts of QAPP and project report. Recommends QAPP approval.
Bill Kammin Quality Assurance Officer Dept. of Ecology	Reviews and approves final QAPP.
Heather Kibbey Chair, PSEMP Steering Committee	Client for report
Katherine Boyd (360) 339-4627 Performance Manager	Reviews final report and interim products
Jim Bolger (360) 464-2014 Director Soundwide and Functional Programs	Reviews final report and interim products
Brian Walsh (360) 464-1234 Planning Manager	Reviews final report and interim products

PSP = Puget Sound Partnership
PSEMP = Puget Sound Ecosystem Monitoring Program

Leska Fore has designed and reviewed monitoring plans and holds a MS in statistics. Constance Sullivan has worked on a variety of projects as an analyst and laboratory scientist and holds a MS in Marine Science. Haley Harguth has an MS in Marine Policy and familiarity with water quality issues.

6.0 Quality Objectives

6.1 Decision Quality Objectives

Not applicable. Decision quality objectives describe the quality of data needed to make specific decisions, such as whether environmental conditions at a site or in an area exceed a regulatory standard. Results from this project may provide strong evidence for recommending policy changes, for example, but additional lines of evidence will likely be needed for decision-making.

6.2 Data Quality Objectives

Data Quality Objectives (DQOs) are the quantitative and qualitative terms used to describe how good data need to be to meet the project's objectives (USEPA, 1996). A key part of this study will be to find and collect data from various sources. Part of this meta-analysis will be documenting the criteria for inclusion of the various data sources. In general, data that have been published, and unpublished data that have undergone a rigorous quality assurance review as part of a report or study, will be reliable. Other sources will be considered after their quality has been verified by data authors.

6.3 Measurement Quality Objectives

Measurement Quality Objectives (MQOs) specify how good the field and laboratory measurement data must be in order to meet the objectives of the project. MQOs are commonly described in terms of the following quality indicators: precision, accuracy/bias, representativeness, comparability, completeness, and sensitivity. Not all of these apply to this project as they would to one involving environmental sampling and lab analysis. However, the following sections describe targets for these quality indicators in the context of this project.

6.3.1 Targets for Precision, Bias, and Sensitivity

We assume, and will verify as possible, that the existing data from individual studies used in this meta-analysis meet MQOs for precision, bias, and sensitivity defined for each project.

Precision

Not applicable. This project will address the inherent variability of environmental results reported in most studies by 1) averaging values from multiple samples to summarize before and after conditions for each study, and by 2) including a large number of studies for each meta-analysis.

Bias

Bias is defined as systematic error associated with an assessment or analysis. We assume the data for each study used in this meta-analysis reflect acceptable levels of bias. This will be confirmed by consulting data authors and by via peer review conducted following PSP's Guidelines for Scientific Review (Hamel and Currens, 2012). But systematic bias is a major concern for meta-analysis studies. When using published literature, bias can result from choosing only certain types of studies for the analysis and from choosing studies that report only positive results. We do not believe these will be issues of concern because of the peer review that will occur and because of the types of data sources that will be used, e.g., TMDL projects implemented in response to regulatory requirements.

Sensitivity

Sensitivity, as commonly used to judge data quality, does not apply to this project. However, a primary goal of this study is to identify which variables are most sensitive to change in site condition resulting from restoration actions. By standardizing variables using Cohen's d, we can compare sensitivity of variables across projects.

6.3.2 Targets for Comparability, Representativeness, and Completeness

Comparability

Comparability is the degree to which different methods, data sets, and decisions agree or are similar. Standard meta-analysis does not require data derived from different projects to be collected using the same methods. However, within a project, meta-analysis does require that data be collected using the same methods before and after the management action being evaluated. Data must be comparable within projects but not necessarily across projects. The mean difference between before and after conditions is calculated for all variables and standardized by dividing by the pooled variance. This change statistic enables comparison across projects.

Representativeness

For this project, representativeness plays a similar role to bias. We cannot guarantee that the projects included in the meta-analysis will fairly represent the universe of all possible projects. For this reason, peer reviewers will be asked to comment on whether the projects are representative of watershed or regional conditions.

Completeness

Completeness refers to the amount of data obtained compared to what was planned. By the end of this project, we expect to gain a rough estimate of how much effectiveness monitoring data is available. Our target for completeness is to conduct a meta-analysis on as many different types of studies as possible, including toxics reduction, TMDL recovery, habitat restoration, and increase in biological diversity.

7.0 Study Design

This project will focus solely on data that have been previously collected and reported. Existing data will be compiled, evaluated, and analyzed using Microsoft Excel and an appropriate statistical program such as R.

Existing data will be provided from the following sources:

- Final reports from agencies
- Final reports and/or journal articles from the Principal Investigator
- Final data sets from the Principal Investigator

First, the available before and after data for studies conducted in specific projects will be assembled. Next, the project will apply to those data standard meta-analysis procedures and power analysis, to address the following questions:

- What specific management actions, or combination of management actions, are effective at the watershed level at achieving their stated goals?
- What environmental variables (e.g., water quality parameters, habitat characteristics, individual species) respond to each management action or combination of management actions?
- Which environmental variables demonstrate the strongest response?

- What types of effectiveness monitoring projects can reliably be included in the meta-analysis? I.e., can water quality parameters, habitat restoration, species population monitoring, and biological effects be easily incorporated into a meta-analysis?

Assumptions underlying analysis

Data are collected using the same methods before and after the restoration or management action. We assume that data have been collected and reported using standard QA/QC methods appropriate for the study. Data collected at a project scale can suggest or point to patterns at a regional scale, but cannot be assumed to apply in other areas not sampled. Over time, as more studies are summarized and compared, confidence in regional patterns will emerge.

8.0 Sampling Procedures

Not applicable - see Section 9.0 below.

9.0 Measurement Procedures

Not applicable. There will be no field samples or original measurements collected for the project in this QAPP. This project will focus solely on the use of existing data from various agency or academic peer-reviewed reports and datasets. The methods of collection for the different types of data are provided in the original reports, which will be documented in our cited literature. Our underlying assumption is that the datasets have been reviewed and approved by internal QA/QC procedures, so we do not need to do further evaluation of their suitability for use. During the vetting process, we will discuss outliers and results with the data authors to ensure a full understanding of the data we use and any associated caveats. We will include datasets that acknowledge QA/QC procedures and report how sampling methods may have been modified over time and how those modifications are expected to impact results.

Parameters of interest include, but are not limited to:

- Fecal coliform counts
- Dissolved oxygen
- Temperature
- Turbidity
- Total suspended solids
- pH
- DDT
- PCBs
- Metals
- Nitrogen (various forms)
- Organism counts
- Habitat condition

10.0 Quality Control Procedures

Quality Control (QC) refers to the standard operating procedures developed for field and laboratory handling that are intended to control the accuracy of measurement. For this project,

we will determine whether individual TMDL reports and other environmental studies have QAPPs associated with them that detail QA/QC protocols followed. Data sources that do not reflect formal QC methods will be identified as such. Authors of each study report will be asked specifically to address the reliability of their data (Appendix 1). When data are collected and compiled across numerous studies, decisions must be made about how to select which studies and results to include.

11.0 Data Management and Analysis Procedures

11.1 Data Management

Gathering data for meta-analysis

Most of the data for this project are located in other reports or databases. Possible sources of data include:

- TMDL studies by Department of Ecology and other jurisdictions within Puget Sound (see References)
- Other water quality data in EIM that has before and after data related to management or restoration actions
- Status and Trends monitoring by Department of Ecology, 25 sites sampled in 2009 were sampled again in 2013
- Results from effectiveness monitoring studies by counties or other jurisdictions
- The Salmon Recovery Funding Board's effectiveness monitoring studies to evaluate fish passage, riparian planting, and other types of restoration
- Monitoring projects referenced in Habitat Work Schedule
- WA Conservation Reserve Enhancement Program (CREP) work with farmers and ranchers to protect and restore riparian areas
- WA Department of Fish and Wildlife's Puget Sound Nearshore Ecosystem Restoration Project

We will cite these sources as appropriate.

Individual study before-and-after data from these sources will include:

- Project name, location, Lat/long, WRIA, project contact, variables measured, actions, years, method for calculating variance, number of samples used to calculate the mean for the start and end year, mean values for the start and end years, standard deviation for the start and end years.

These data will be transferred or manually entered into one or more Excel spreadsheets, and manually double-checked for accuracy. The spreadsheets will be managed (e.g., stored, backed up, and archived) at the Puget Sound Institute (PSI) according to its Quality Management Plan (PSI 2013).

11.2 Data Analysis

Before and after data for a restoration or management action (or suite of actions) reported by each author will be organized in Excel files. We will use Excel and "R" to calculate the following statistics (R Core Team, 2013):

- Mean
- Standard deviation
- Pooled standard deviation
- Ratio of the start and end year means
- Cohen's d statistic
- Estimates of variance
- Confidence intervals
- t-tests for significant difference

11.2.1 Meta-analysis

Meta-analysis is the statistical analysis of a large collection of studies for the purpose of integrating the findings (DeCoster, 2009). Techniques for meta-analysis were developed primarily in medical, physical and behavioral science and have only recently begun to be applied to ecological data (Gurevitch and Hedges, 1999; Harrison, 2011). Meta-analysis calculates an effect size for each study and derives general conclusions based on the central tendency and variability of the effect sizes (Nakagawa and Cuthill, 2007). The most common types of meta-analyses are based on literature reviews (e.g., Stewart et al., 2006; Miller et al., 2010; Whiteway et al., 2010). In recent years, extensive data bases such as WA Department of Ecology's Environmental Information Management (EIM) database, with 12.5 million records of data for the state, provide new opportunities to more directly evaluate project level data.

The basic purpose of the meta-analysis approach is to provide the same methodological rigor to a data review that would be required for experimental research. Meta-analyses are generally designed to evaluate the relationship between an explanatory and a response variable, that is, "the effect of X on Y." In our case, we will evaluate the effect of restoration or management actions on the physical, chemical, or biological condition of stream or nearshore habitat. Where possible, we will evaluate the effectiveness of individual actions; however, this will often not be possible when a suite of actions are applied across multiple sites or when information about the location of specific actions is not recorded.

Because the analysis is dependent on available data, the results of meta-analysis must be reviewed and vetted by the people who originally collected the data or are responsible for its application and interpretation. When data are collected and compiled across numerous studies, decisions must be made about how to select which studies and results to include. These decisions must also be transparent in the analysis procedure and shared with reviewers. Decisions and conclusions drawn from a meta-analysis, thus, are derived from a larger community of practice than just the authors of the study (Stewart et al., 2006).

Measuring the effect of actions on ecosystem health

Individual studies have documented changes in water quality variables as a result of fencing cows from streams, adding riparian cover, and removing sources (Sargeant and Svrjcek; 2008). Effectiveness studies related to Total Maximum Daily Load (TMDL) assessment under the Clean Water Act require monitoring to document whether the actions have resulted in enough improvement to change the status of the water body from impaired to unimpaired (Collyard, 2009). Actions related to salmon recovery such as removal of barriers, placement of wood structures, and increasing habitat connectivity also provide data to evaluate the impact of

actions on habitat condition. The goal for this meta-analysis is to look across a variety of data sets and see which individual actions, or suites of actions, result in measurable outcomes to the habitat, water quality or biota.

Calculating statistics

For each study, the "effect size" will be calculated as the difference between conditions before and after the treatment or restoration divided by the pooled standard deviation. Effect sizes provide a standardized measure of the change observed as a result of an action. Because differences are standardized by the pooled variance, results can be compared across projects and for different response variables. The effect size is a unitless measure of change (Garamszegi, 2006; Blanar et al., 2009).

We will use Cohen's *d,* because it is easy to interpret and widely used as a common estimator of effect size (Rosnow et al., 2000; Nakagawa and Cuthill, 2007; DeCoster, 2009). Cohen's *d* is roughly equivalent to a *Z*-score from a normal distribution and is calculated as:

$$d = \frac{\mu_A - \mu_B}{\sigma}$$

Where d = Cohen's d,

μ_B = Mean value of the variable before the action or treatment,

μ_A = Mean value of the variable after the action or treatment, and

σ = the pooled standard deviation.

The pooled variance is calculated as:

$$SD_{Pooled} = \sqrt{\frac{(N_A - 1)(SD_A)^2 + (N_B - 1)(SD_B)^2}{N_A + N_B - 1}}$$

Where SD_B and SD_A are the standard deviations of the before and after groups, and N_B and N_A are the sample sizes for each group.

To correct for bias due to small or different sample sizes, each study may be weighted as (Miller et al., 2010):

$$J = 1 - \frac{3}{4(N_A + N_B - 2) - 1}$$

The statistic *d* represents the difference in water quality, abundance, or habitat measures before and after restoration or management actions. We will use a 90% confidence interval to test whether *d* is statistically significant from zero. A 90% confidence level is standard for these types of studies but other levels can easily be calculated and reported.

Statistical testing for change

For individual studies, an effect size of 0.8 indicates that the mean condition or value before the action is 0.8 standard deviations (SD) different from the mean condition or value after the action (the response). In general, a difference larger than 0.8 SD is considered a strong effect (Cohen, 1988; Blanar et al., 2009). An observed difference before and after an action that is less than 0.2 SD is considered trivial, a difference 0.2–0.5 SD is considered small, and 0.5–0.8 SD is

considered moderate. These guidelines were developed by Cohen (1988) and are widely cited in the literature on meta-analysis. The relative change in any measure of effect size needs to be interpreted within the context of how much change was possible and expected.

To evaluate effectiveness of all studies, we will test for change in Cohen's *d* across all studies (Nakagawa and Cuthill, 2007). As an example, data from Salmon Creek, WA were used to illustrate how meta-analysis can be used to test for change in fecals and turbidity at eight study sites where agricultural management practices were implemented (Collyard, 2009; Figure 2).

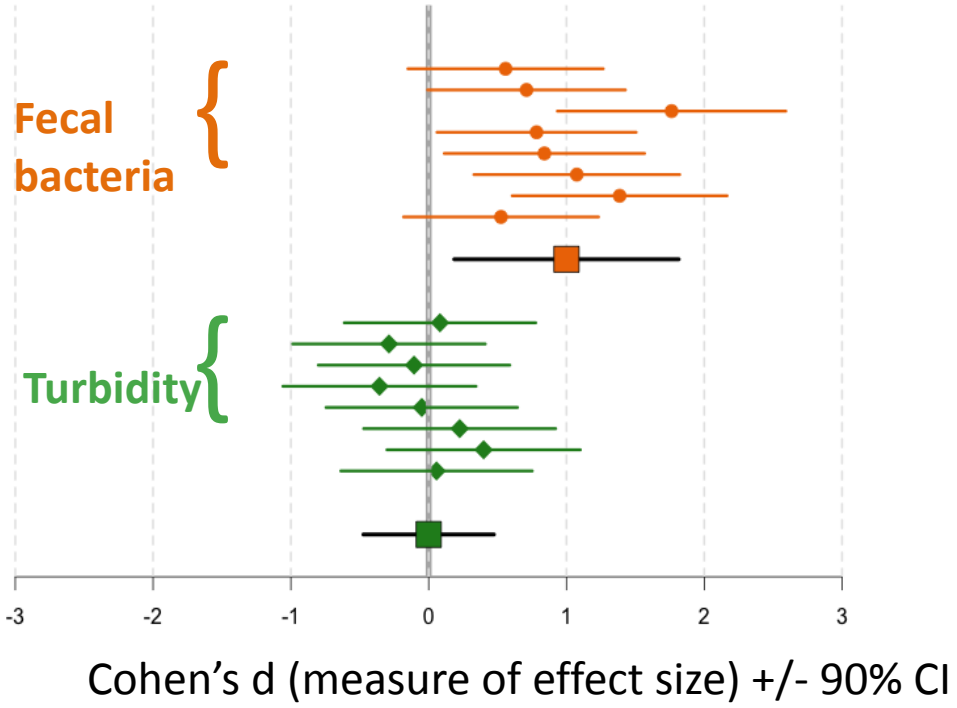

Figure 2. Cohen's *d* for eight sites and two water quality measures.

Changes in fecal coliform were significantly different from zero indicating the positive impact of agricultural management practices. In contrast, turbidity measured at the same sites was not significantly different from 0, indicating that turbidity is less sensitive to these types of changes.

11.2.2 Statistical Power Analysis

Data collected for other purposes can be used to estimate the sensitivity of a proposed study design. Statistical power provides an analysis framework for testing whether a proposed study design has a good probability of detecting a change should a change truly occur. If a measure is too variable, and the change expected is small, it is unlikely that the measure will be able to detect a difference that is statistically significant.

One approach for estimating statistical power to detect change is the minimum detectable difference (MDD). The MDD can be calculated for a variety of statistical tests. For example, a two-sample *t* test might be used to compare a site before and after a management action (Zar, 1984). The MDD represents the smallest detectable difference between mean values for a measure of interest, e.g., turbidity, before and after the action.

For this statistical model, MDD can be calculated as:

$$\text{MDD} \geq \sqrt{\frac{2s^2}{n}}\left(t_{\alpha(2),\nu} + t_{\beta(1),\nu}\right),$$

Where s^2 = the within-site error variance of the measure of interest,
n = the number of replicate samples at each site,
$t_{\alpha(2),\nu}$ = the *t* value for alpha of 0.1 for a 2-sided test,
$t_{\beta(1),\nu}$ = the *t* value for beta of 0.1 for a 1-sided test, and
$\nu = 2n - 2$.

12.0 Reporting and Vetting

A draft and final report will be prepared. Only summary statistics will be compiled. The draft report and draft data analysis will be presented to relevant partners, and their comments and caveats will be included with the interpretation of the results (Appendix 1). This is a common protocol for meta-analysis studies because it is impossible to vet all the data sources and understand the relative importance of assumptions for each study, only the data authors can do that.

Steps for vetting process:

- Obtain datasets directly from data authors or via web searches
- Perform meta-analysis
- Meet with each data author and give a presentation that includes:
  - Background on project
  - What meta-analysis is and how we are calculating it
  - How to interpret graphs and data outputs
  - Their data in a meta-analysis and our conclusions based on the results
- During the meeting, ask the data authors the following questions:
  - Do the results look as you would expect?
  - Are there any outliers that you feel need to be explained?
  - Is our interpretation of your data in line with your expectations?
  - Is our interpretation of the impact of management actions correct?
- After meeting, return completed data vetting sheets to authors with any revised graphs
- Receive approval for data vetting sheet and final graphs

13.0 Data Verification and Validation

The Puget Sound Partnership and the Puget Sound Institute will not collect new environmental data that will be used for this project. The WA Department of Ecology, local jurisdictions, and other regional partners will be the primary sources of data. The data gathered are considered secondary data, and the original data collectors were responsible for the quality of the data collected in the field and provided in reports or databases. As part of our final report, we will list data sources, reports, and contact information for the data used in the meta-analysis.

All data will be obtained from published sources or will have been used for regulatory purposes. Therefore, the quality and usability of the data will have been evaluated by the program that collected and analyzed them originally. When non-detect data are used, we will use half the

detection limit or some other commonly-used method that will be documented in the data file of summary statistics.

Data validation is not required for this project.

14.0 Data Quality (Usability) Assessment

We will work with data authors and PSEMP experts will use the form in Appendix 2 to confirm the usability of results from each project, discuss any caveats, and vet conclusions of the meta-analysis. The analysis may then be modified, as indicated by the comments/feedback received.

Data authors will be asked specifically to address the reliability of their data and the conclusions from the meta-analysis (Appendix 2).

Reviewing the results of meta-analysis

Any meta-analysis is dependent on the studies that are available for analysis; therefore, interpretation needs to involve people who collected the data and regional experts. We will work with the Puget Sound Ecosystem Monitoring Program (PSEMP) to vet and interpret results from the meta-analysis. PSEMP serves as an independent group of stakeholders and experts guided by a Steering Committee and composed of more than 10 topical Work Groups. The Freshwater, Nearshore, Stormwater, Salmon, Toxics, and Marine Waters Work Groups include the most likely experts for interpreting meta-analysis results. We will engage select members of these groups and record their responses to specific questions (Appendix 2):

- Are you confident that your data are representative?
- Are you confident that the meta-analysis represents your data fairly and accurately?
- Are there site specific differences that influence results?
- What local factors might limit our ability to generalize results to other locations?

The analysis will be modified as needed and reviewers' interpretations will be included in the final report.

15.0 References

Bash, J. S. and C. M. Ryan. 2002. Stream restoration and enhancement projects: Is anyone monitoring? Environmental Management 29(6):877-885.

Blanar, C. A., K. R. Munkittrick,J. E. Houlahan, D. L. MacLatchy, D. J. Marcogliese. 2009. Pollution and parasitism in aquatic animals: A meta-analysis of effect size. Aquatic Toxicology 93: 18-28.

Bowler, D. E., R. Mant, H. Orr, D. M. Hannah, and A. S. Pullin. 2012. What are the effects of wooded riparian zones on stream temperature? Environmental Evidence, 1:3, www.environmentalevidencejournal.org/content/1/1/3.

Brandon, T., N. Gleason, C. Simenstad, and C. Tanner. 2013. Puget Sound Nearshore Ecosystem Restoration Project Monitoring Framework. Prepared for the Puget Sound Nearshore Ecosystem Restoration Project. Published by Washington Department of Fish and Wildlife, Olympia, Washington, and U.S. Army Corps of Engineers, Seattle, Washington. 73 pp.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Lawrence Erlbaum Associates.

DeCoster, J. 2009. Meta-Analysis Notes. Retrieved December, 2013. www.stat-help.com/notes.html.

Garamszegi, L. Z. 2006. Comparing effect sizes across variables: generalization without the need for Bonferroni correction. Behavioral Ecology 17:682-687.

Gurevitch, J and L. V. Hedges. 1999. Statistical issues in ecological meta-analyses. Ecology 80(4): 1142-1149.

Hamel, N. and K. Currens. 2011. The Puget Sound Partnership's Guidelines for Scientific Review. Puget Sound Partnership, Tacoma, WA. www.mypugetsound.net

Harrison, F. 2011. Getting started with meta-analysis. Methods in Ecology and Evolution 2: 1-10.

Miller, S. W., P. Budy, J. C. Schmidt. 2010. Quantifying macroinvertebrate responses to in-stream habitat restoration: Applications of meta-analysis to river restoration. Restoration Ecology 18(1): 8-19.

Nakagawa, S. and I. C. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biological Reviews 82: 591-605.

Puget Sound Partnership. 2013. The 2012/2013 Action Agenda for Puget Sound. http://www.psp.wa.gov/action_agenda_2012-13.php.

Puget Sound Partnership. 2013. 2013 State of the Sound: A Biennial Report on the Recovery of Puget Sound. Tacoma, Washington. http://www.psp.wa.gov/SOS_download.php.

Puget Sound Institute. 2013. Puget Sound Institute, Quality Management Plan. Tacoma, Washington.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rosnow, R. L., R. Rosenthal, and D. B. Rubin. 2000. Contrasts and correlations in effect-size estimation. Psychological Science 11(6): 446-453.

Stewart, G. B., H. R. Bayliss, D. A. Showler, A. S. Pullin, and W. J. Sutherland. 2006. Does the use of in-stream structures and woody debris increase the abundance of salmonids? CEE review 05-006 (SR12). Collaboration for Environmental Evidence: www.environmentalevidence.org/SR12.html.

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

Whiteway, S. L. P. M. Biron, A. Zimmermann, O. Venter, and J. W. A. Grant. 2010. Do in-stream restoration structures enhance salmonid abundance? A meta-analysis. Canadian Journal of Fisheries and Aquatic Sciences 67:831-841.

Zar, J. H. 1984. Biostatistical Analysis, 2nd ed. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Appendix 1 - Data sources for meta-analysis

Washington Department of Ecology. Accessed December 2013. Environmental Information Management. http://www.ecy.wa.gov/eim/.

Bell-McKinnon, M. 2007. Lake Campbell and Lake Erie total phosphorus total maximum daily load; water quality effectiveness monitoring report. Dept. of Ecology Publication No. 07-03-011.

Collyard, S. 2009. Salmon Creek nonpoint source pollution total maximum daily load; water quality effectiveness monitoring report. Dept. of Ecology Publication No. 09-03-042.

Onwumere, G.2002. Effectiveness monitoring for total phosphorus total maximum daily loads in Fenwick and Sawyer Lakes. Dept. of Ecology Publication No. 02-03-054.

Onwumere, G.2003. Effectiveness monitoring for fecal coliform total maximum daily loads in Piper Creek. Dept of Ecology Publication No. 03-03-027.

Sargeant, D., C. Hempleman. 2007. Skokomish River Basin fecal coliform bacteria total maximum daily load study; water quality attainment monitoring report. Dept. of Ecology Publication No. 07-03-054.

Sargeant, D., and R. Svrjcek. 2008. Snoqualmie River Basin fecal coliform bacteria, dissolved oxygen, ammonia-nitrogen, and pH total maximum daily load; water quality effectiveness monitoring report. Dept. of Ecology Publication No. 08-03-005.

Appendix 2 - Template for Vetting Results of Meta-analysis

Questions for PSP analysts:

Project description and variables measured.

What data were used in what matrix?

Were reference sites used?

Were data measured same way before and after action at same sites?

Was the action effective (was Cohen's d significantly different from 0)? Why or why not?

Questions for data authors:

Do the results make sense? Are you confident that your data are representative?

Are you confident that the meta-analysis represents your data fairly and accurately?

Are there site specific differences that influence results?

What local factors might limit our ability to generalize results to other locations?

Other caveats? (e.g., anomalous events or conditions)

May we share these results?