Monitoring Framework to Assess Salmonid Responses to Enhancement Actions and Climate Change

March 1, 2017

Prepared by

BioAnalysts, Inc.
Yakama Nation

Prepared for

Yakama Nation

Citation: Hillman, T., B. Rogers, K. Murdoch, T. Iverson, and B. Rose. 2017. Monitoring framework to assess salmonid response to enhancement actions and climate change. Report to the Yakama Nation, Toppenish, WA.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

Appendix 1: Common Metrics Associated with Measured Physical and Biologic Parameters

LIST OF ACRONYMS AND ABBREVIATIONS

ANOVA	Analysis of variance
AREMP	Aquatic and Riparian Effectiveness Monitoring Program
BA	Before-After study design
BACI	Before-After Control-Impact study design
BPA	Bonneville Power Administration
CHaMP	Columbia Habitat Monitoring Program
DART	Columbia River Data Access in Real Time
EDT	Ecosystem Diagnosis and Treatment
EMAP	Environmental Monitoring and Assessment Program
EPT	Extensive Post-Treatment study design
ESA	Endangered Species Act
GRTS	Generalized random-tessellation stratified sample
IMW	Intensively Monitored Watershed
IPT	Intensive Post-Treatment study design
ISEMP	Integrated Status and Effectiveness Monitoring Program
km	Kilometer
LWD	Large woody debris
m	Meter
MERR	Monitoring, Evaluation, Research, and Reporting plan
NOAA	National Oceanic and Atmospheric Administration
OBMEP	Okanogan Basin Monitoring and Evaluation Program
OWEB	Oregon Watershed Enhancement Board
PIBO	Pacfish/Infish Biological Opinion
PIT tag	Passive Integrated Transponder tag
PTAGIS	PIT Tag Information System
PNAMP	Pacific Northwest Aquatic Monitoring Partnership
PUD	Public Utility District
RHIP	Regional Habitat Indicator Project
RME	Research, Monitoring, and Evaluation
SRFB	Salmon Recovery Funding Board
VSP	Viable Salmonid Population
YN	Yakama Nation

1 Introduction

Climate change is predicted to have significant effects on Pacific salmon and their habitat in the Pacific Northwest (Crozier 2014). Summer air temperatures in the Pacific West have increased by roughly 0.2°C/decade, and 99% of streams have increased by about 0.1°C/decade since the mid-1960s (Isaak and Young 2015). Current projections for Chelan County, Washington, using the mean model indicate that by 2050-2074 annual average maximum air temperature will increase 3.2°C, average precipitation will increase 6.2 mm/month, snowfall will decrease -99.3 mm, and soil storage will decrease -6.4 mm (see https://www2.usgs.gov/climate_landuse/clu_rd/nccv/viewer.asp). These changes in climatic conditions are expected to reduce holding, spawning, and rearing habitat for salmonids and other species such as lamprey (Keefer et al. 2009). As such, it is important to implement actions that ameliorate climate change effects.

Several enhancement techniques have been developed to improve tributary habitat quantity and quality (e.g., Roni and Beechie 2013). The literature indicates that these techniques are generally effective at improving habitat conditions for salmonids (see review by Hillman et al. 2017). It is largely unknown, however, if these actions can fully ameliorate the effects of climate change. Beechie et al. (2013) recently published a paper on restoring habitat under a changing climate. They identified common techniques that are expected to ameliorate climate change effects on peak flows, low flows, and stream temperatures. Those techniques include projects that increase longitudinal, lateral, and vertical connectivity; restore natural stream flows; and improve riparian functions. Although placement of instream structures is not considered an action that improves stream flows or temperature regimes, if it is constructed and placed correctly, large wood and boulder structures can increase the number and extent of cold-water refugia for salmonids (see Hester et al. 2009).

Managers in the Columbia River basin are working diligently to implement enhancement actions in tributaries that will improve habitat quality and quantity for salmonids. This work is designed not only to address current factors limiting salmonid production, but also to address projected changes in habitat conditions due to climate change. Because of the wide array of actions that can be implemented to address a given limiting factor and the cost of those actions, it is important to determine which actions provide the greatest benefit per unit cost. The purpose of this paper is to develop a framework for monitoring the effectiveness of enhancement actions on salmonids in the mid- and upper-Columbia regions. This framework is unlike most other monitoring plans or frameworks in that it is set up to track a small suite of metrics directly applicable to the action being implemented. Thus, it provides a simple, cost-effective approach to determining the effectiveness of enhancement actions on salmonids and their habitat, and it tracks important metrics that are sensitive to climate change.

This framework is intended to guide development of monitoring plans that address the primary and long-term question, *Are the habitat restoration actions providing environmental resilience within the stream/riparian areas and do they translate into meaningful and measurable effects on the productivity and survival of juvenile and adult salmonids?* To that end, in this report, we (1) identify a framework for monitoring tributary enhancement projects at the project or reach scale; (2) identify a framework for monitoring tributary enhancement projects at the watershed or population scale; and (3) describe the implementation of enhancement and monitoring under an adaptive management approach. Under the first two sections, we describe what questions can be answered with effectiveness monitoring, identify appropriate monitoring designs, identify appropriate parameters to measure, identify possible measuring methods, and describe implementation and coordination.

Importantly, this framework is not a “new” approach that needs to be tested. Rather, it draws heavily from existing monitoring programs that have been implemented successfully within the Pacific Northwest. Thus, the elements included in this document have been proven to work and are consistent with the National Oceanic and Atmospheric Administration (NOAA) Research, Monitoring, and Evaluation (RME) Guidance for Endangered Species Act (ESA) listed salmon (Crawford and Rumsey 2011) and the Council’s Monitoring, Evaluation, Research, and Reporting (MERR) Plan. This work also supports and builds upon the climate change vulnerability assessments developed by NOAA Fisheries (Crozier et al. in prep). The monitoring framework proposed here will help evaluate the four key exposure factors identified by NOAA Fisheries.[1] Finally, the work of Roni (2005) informed the development of this document.

---

[1] The four salmon-specific freshwater exposure attributes include: **hydrologic regime** (identify watersheds where the fundamental hydrologic regime may shift from snow-dominated to transitional or from transitional to rain-dominated); **storm frequency/flood magnitude** (characterize a change in exposure to flood events that can scour eggs out of the gravel, input fine sediment that buries eggs, or modify stream structural habitat); **stream temperature in August** (use spatially explicit projections of warming stream temperature to quantify the increase in thermal stress for the ESU); and **climate water deficit** (represent the risk of climate change on summer water availability for streams and ecosystems) (from Crozier et al. in prep).

2 Project-Scale Effectiveness Monitoring

Most habitat projects implemented by the Yakama Nation (YN) are implemented at a small scale, with the goal of enhancing or improving habitat features or habitat-forming processes. Project or reach-scale effectiveness monitoring[2] is concerned with measuring local habitat and fish parameters to determine whether the actions implemented were effective in creating a desired change in habitat condition and fish performance. In this case, the focus is on assessing the effects of habitat actions on habitat characteristics and fish abundance, biomass, growth, movement, habitat use, and/or distribution (e.g., Polivka et al. 2015). What follows is a framework for developing specific monitoring plans to assess the effectiveness of tributary habitat enhancement actions at the project or reach scale.

2.1 Monitoring Questions

Because enhancement projects, like many management actions, are experiments, they should be implemented according to standard rules of experimental design. This means that the goal of the project and the monitoring objectives and questions must be identified clearly. It is beyond the scope of this document to identify all the objectives and questions for each type of enhancement action. Therefore, in this document, we identify general questions that can be made more specific depending on the type of enhancement action implemented. What follows are “general” questions that can be addressed with project or reach-scale effectiveness monitoring.

What are the effects of specific habitat actions on habitat parameters (physical/water quality) that were the target of the action?

> Depending on reach conditions, geomorphic processes, and current limiting factors, different types of habitat actions can be implemented to address specific ecological concerns. This question addresses which of those actions effected a positive change in fish passage, fish entrainment, instream structure (e.g., pool area, large wood debris, habitat complexity), off-channel habitat, fine sediments, water quality, stream flows, nutrient enhancement, or riparian habitat. If the action did not produce a positive change in the habitat parameter, it will not translate into a positive fish response.

What are the effects of specific habitat actions on fish distribution and abundance or biomass at the project scale?

> The purpose of this question is to determine if the fish responded favorably to the habitat action. Specifically, this question determines if actions to improve fish passage, fish entrainment, instream structure, off-channel habitat, fine sediments, water quality, stream

---

[2] In this document, we do not differentiate between project and reach-scale effectiveness monitoring. Strictly speaking, project monitoring occurs at a smaller scale than reach-scale monitoring; however, the same approaches are used to measure treatment effects. Therefore, we do not provide a separate monitoring framework for the two scales of monitoring.

flows, nutrient enhancement, or riparian habitat increased fish abundance or biomass and distribution.

What are the relationships between fish response (abundance or biomass) to habitat enhancement and the range of specific habitat parameters (e.g., number of pools, volume of woody debris, etc.)?

The monitoring designs identified in this document (see next section) will allow the collection of fish data under a wide range of habitat conditions and seasons. With a high level of replication, one can examine differences among sites and seasons, and therefore develop relationships between fish abundance or biomass and differing levels of key habitat parameters. For example, previous studies have found a strong correlation between the change in pool area or pieces of pool-forming wood and coho and steelhead response to enhancement, indicating that those sites where wood or boulder placement was the most intense and increased pool area the most had the largest increase in fish densities (Roni and Quinn 2001; Roni et al. 2006).

These general questions can be transcribed into specific questions depending on the type of enhancement project implemented. Answers to specific questions will give managers the information they need to identify cost-effective actions for addressing specific habitat impairments and ecological concerns under the effects of climate change. That is, data from action effectiveness monitoring is designed to reveal which actions are most appropriate for addressing the habitat impairments.

2.2 Monitoring Designs

There are many potential study designs for monitoring enhancement actions at the project or reach scale. Although none is ideal for all situations, “before-after” (BA), “before-after control-impact” (BACI), and “post-treatment” designs are appropriate for monitoring the effects of enhancement actions at the reach or project scale. A BA study refers to a design where data are collected both before and after treatment. Data collected before treatment serve as pre-treatment or control[3] data (temporal control), while data collected after treatment serve as post-treatment data. If there is a treatment effect, the post-treatment score should be more desirable than the pre-treatment score. Two or more years of pre-treatment data are needed to adequately assess project effects. This design is appropriate if spatial controls are not available.

A BACI study is a BA design with one or more spatial control sites. If the spatial controls are selected appropriately (i.e., they are closely matched with the treatment site but spatially independent)[4], a BACI is more powerful statistically than a simple BA design. Under the BACI study design, a spatial control site is evaluated over the same time period as the treatment site. The addition of a spatial control site to the BA study design is meant to account for environmental variability and temporal trends found in both the control and treatment areas and, thus, increase the ability to differentiate treatment effects from natural variability. Adding more than one control site further increases the probability of detecting a treatment effect. The BACI study design is the preferred design for most project or reach-scale monitoring. O’Neal et al. (2016) used this design to evaluate reach-scale physical and biological

---

[3] It is important to point out that “control” and “reference” sites are not the same thing. A “control” site is defined as being similar to the treatment site before the treatment site is treated; a “reference” site is defined as the ideal or pristine state, with conditions unaltered by the treatment or other human activities (Downes et al. 2002).

[4] Spatial control sites need to be as similar as possible to the treatment sites. The design does not require exact pairing; parameters simply need to “track” each other. False conclusions can occur if pretreatment trends in the parameters of interest are not similar between treatment and control sites.

effectiveness of different types of enhancement actions. They noted that two or more years of pre-treatment data are necessary to adequately determine the effectiveness of project types, especially for fish.

In some situations, collecting data before treatment is not possible. In these situations, the treated area is compared to control areas thought to be similar in the absence of enhancement activities. These studies are replicated spatially rather than temporally (space for time substitution). There are two types of post-treatment designs: "intensive post-treatment" (IPT) designs in which multiple years of data are collected at one or a few paired control and treatment sites; and "extensive post-treatment" (EPT) designs in which paired treatment and control sites are each sampled once over a one to three-year period. Which post-treatment design is used depends on the number of paired control and treatment sites available and the amount of time needed to assess treatment effects.

Which of the various types of designs should be used depends on the "specific" monitoring question, experimental setting (i.e., are spatial controls available, are pre-treatment data available, etc.), and available resources. For example, if the goal is to understand the effects of enhancement in one site, an intensive BA or BACI design is appropriate. On the other hand, if the goal is to understand the effects in multiple sites, an extensive BA, BACI, or EPT design is appropriate. Table 1 describes the advantages and disadvantages of each design type.

Table 1. Summary of advantages and disadvantages of different types of study designs commonly used to assess habitat enhancement projects (modified from Roni et al. 2005). Intensive study designs generally include sampling at one or two sites; extensive study designs sample multiple sites. Years of monitoring needed to detect a fish response are general estimates based on juvenile salmonid studies, and extensive study designs assume more than 10 sites are sampled; thus, fewer years of monitoring are needed. BACI = before-after control-impact.

ATTRIBUTE	STUDY DESIGNS				
	BEFORE AND AFTER			POST-TREATMENT	
	INTENSIVE	EXTENSIVE	BACI	INTENSIVE	EXTENSIVE
Includes collection of pre-treatment data	Yes	Yes	Yes	No	No
Ability to assess inter-annual variation	Yes	Yes	Yes	Yes	No
Ability to detect short-term response	Yes	Yes	Yes	No	Yes
Ability to detect long-term response	Yes	No	Yes	Yes	Yes
Ability to assess interaction of physical setting and treatment effects	Low	High	Low	Low	High
Applicability of results	Limited	Broad	Limited	Limited	Broad
Potential bias due to small number of sites	Yes	No	Yes	Yes	No
Years of monitoring needed to detect a fish response	10+	1-3	10+	5+	1-3

2.3 Monitoring Parameters

Identifying which physical and/or biological parameters to measure depends on the goals and key questions, selection of a monitoring design, and the availability of monitoring tools and protocols. Monitoring parameters should be relevant to the questions asked, strongly associated with the enhancement action, ecologically significant, and efficient to measure. Moreover, parameters must change in a measurable way in response to treatment, must be directly related to the resource of concern, and must have limited variability and not likely to be confounded by temporal or spatial factors. Table 2 identifies potential physical parameters to measure when monitoring different types of enhancement actions. These parameters are directly linked to the ecological concerns identified by Hamm (2012).

Table 2. Rankings of the usefulness of physical parameters to monitoring effects of different tributary habitat actions. Rankings vary from 1 = highly likely to be useful; 2 = moderately likely to be useful; and 3 = unlikely to be useful or little relationship. Table is modified from Hillman (2006). Appendix 1 identifies common metrics calculated from the measured parameters.

PHYSICAL PARAMETERS	TYPES OF HABITAT ENHANCEMENT ACTIONS							
	DIVERSION SCREENS	BARRIER REMOVAL	SEDIMENT REDUCTION	IMPROVE WATER QUALITY	INSTREAM FLOWS	FLOODPLAIN HABITAT	RIPARIAN HABITAT	INSTREAM STRUCTURES
Temperature	3	2	3	1	1	1	1	2
Nitrogen	3	3	3	1	3	2	2	3
Phosphorus	3	3	3	1	3	2	2	3
Migr. Barriers	3	1	3	3	2	3	3	3
Substrate	3	2	1	3	1-2	1-2	2	1-2
Embeddedness	3	1-2	1	1-2	1-2	1-2	2	1-2
Fines	3	1-2	1	1-2	2	1-2	2	1-2
Woody debris	3	3	3	3	2	1	1	1
Pool density	3	1-2	2	3	1-2	1	1-2	1
Pool depths	3	1-2	1	3	1	1	1-2	1
Fish cover	3	2	1	1-2	1	1	1-2	1
Reach length	3	1	2	3	2	1	1	1
Sinuosity	3	3	2	3	1	1	3	1
Wetted width	3	1-2	1-2	3	1-2	1-2	1-2	1
Bankful width	3	1-2	1-2	3	1-2	1-2	1-2	1
Bank erosion	3	2	1-2	3	2	1	1	1
Riparian struct.	3	2	2	2-3	2	1	1	1-2
Riparian disturb.	3	2	2	2-3	2	1	1	1-2
Canopy cover	3	2	2	2-3	2	1	1	1-2
Streamflow	3	2	3	3	1	2	2	2

As with physical parameters, biological parameters to be measured differ among types of enhancement actions (Table 3). Biological parameters are divided into four major categories; adults, redds, juveniles/parr, and smolts. Adults, redds, and juveniles are the primary parameters used to measure success of different types of actions. Smolts are difficult to measure at the project or reach scale and therefore are not identified as an important parameter to measure at this scale. However, the measurement of smolts is important at the watershed or population scale.

Table 3. **Rankings of the usefulness of biological parameters to monitoring effects of different tributary habitat actions. Rankings vary from 1 = highly likely to be useful; 2 = moderately likely to be useful; and 3 = unlikely to be useful or little relationship.**

BIOLOGICAL PARAMETERS	TYPES OF HABITAT ENHANCEMENT ACTIONS							
	DIVERSION SCREENS	BARRIER REMOVAL	SEDIMENT REDUCTION	IMPROVE WATER QUALITY	INSTREAM FLOWS	FLOODPLAIN HABITAT	RIPARIAN HABITAT	INSTREAM STRUCTURES
Adult number	3	1	1	1	1	1	2	1
Adult origin	3	1	2	2	2	2	2	2
Redd number	3	1	1	1	1	1	1	1
Redd distribution	3	1	1	1	1	1	1	1
Fry number	3	2	1	1	1	1	2	1
Fry distribution	3	2	1	1	1	1	2	1
Juv. number	3	1	1	1	1	1	1	1
Juv. distribution	3	1	1	1	1	1	1	1
Juv. size	3	2	2	1	2	1	3	1
Juv. movement	3	1	3	3	2	1	3	2

2.4 Replication

The size of the enhancement action, the variability of the parameters of interest, and the amount of spatial and temporal replication will determine how well a monitoring program can detect change. A "power analysis" is generally used to determine the number of spatial and/or temporal replicates needed to detect a meaningful change in biological and physical conditions. A power analysis requires the following quantities:

1. Variance – Estimate of the amount of variability in a parameter.
2. Power ($1 - \beta$) – The probability of detecting a difference or change if it does exist (the probability of rejecting the null hypothesis when it is in fact false; probability of a type II error).
3. Effect Size – The difference that one would like to detect between groups being compared.
4. Significance Level (α) – The probability of detecting a difference when it does not exist (probability of rejecting the null hypothesis when it is true; probability of a type I error).

For example, these quantities are needed to determine the sample size (number of replicates) necessary to have an 80% probability of detecting a 50% physical or biological change, given $\alpha = 0.05$ and variance of X. Fewer replicates are needed to detect a change if the variance of the parameter is small and the effect size (magnitude of the treatment) is large. The number of replicates is also determined by the significance or alpha-level (α) set for the test and the level of statistical power. The investigator determines these levels before monitoring is initiated.

The number of replicates needed to detect the response of different parameters can be highly variable. Nevertheless, for BA studies (including BACI), O'Neal et al. (2016) found that two or more years of pre-treatment data are necessary to adequately determine the effectiveness of project types. As a starting point, we recommend at least two years of pretreatment data, three consecutive years of post-treatment data, and then sampling every five years for 15 years. Thus, for BA designs (including BACI), the investigator collects a total of eight temporal replicates: two pretreatment replicates and six post-treatment replicates. For IPT designs, as a starting point, we recommend at least three consecutive years of sampling and then sampling every five years for 15 years.

Both BACI and post-treatment designs (IPT and EPT) require matched treatment and control sites (project scale) or reaches (reach scale). Inadequate pairing of control and treatment sites can lead to greater variability and therefore less statistical power. We recommend using the stream classification approach used for Ecosystem Diagnosis and Treatment (EDT) modeling. Stream reaches and assessment units throughout the upper Columbia region will be classified using this approach. It is important that the treatment and spatial control sites be independent of each other. That is, the treatment cannot affect the environmental and biological parameters in the spatial control site or reach. Thus, we recommend that spatial controls be located upstream of treatment sites.

2.5 Sampling Scheme

Once parameters and the number of spatial and/or temporal replicates have been determined, one must determine the spatial allocation of sampling within a site or study reach. The allocation of sampling depends on the size of the reach or project and the parameters being measured. If the site or reach is a short segment in a small stream, one may survey the entire site or reach (a census). However, for long treatment sites or reaches that cannot be surveyed with a complete census (because of limited resources or time), one will need to identify a small number of areas within the reach to sample. These sampling areas must be selected in such a way as to be representative of the entire reach (unbiased samples). Common sampling methods include simple random sampling, stratified random sampling, systematic sampling, cluster sampling, multistage sampling, double sampling, generalized random-tessellation stratified sample (GRTS), and capture-recapture sampling (Thompson 1992; Stevens 2002).

We recommend that sampling sites for effectiveness monitoring be selected according to a stratified random sampling design. Given that stream classification work will be used to help select appropriate control sites or reaches, the stratification process should already be complete (see Replication Section above). Thus, one only needs to select randomly survey sites within each reach. The number of sites selected randomly within each reach should be proportional to the size (length) of the stratum. That is, a longer stratum would receive more sites than a shorter stratum. As a starting point, we recommend that up to 30% of the stratum be sampled. In addition, sampling sites should vary in size according to the mean bankful width. That is, a sampling site should be 20 times the mean bankful width, but not less than 150-m long or longer than 500 m.[5]

---

[5] This reach length differs from Simonson et al. (1994) and Reynolds et al. (2003), which use 40x the wetted width. The use of 20x the bankful width is consistent with AREMP and PIBO protocols. This protocol also allows one to assess channel conditions even if the channel is dry.

As noted above, sampling also depends on the parameters being measured. Biological parameters related to adults and redds should be measured throughout the entire site or reach regardless of the length of the site or reach (complete census). Physical parameters such as temperature, water quality, and flow can be measured at the upstream and downstream ends of the site or reach. Juvenile parameters and the other physical/environmental parameters should be measured within the randomly selected sites.

Most of the parameters identified in Tables 2 and 3 are measured annually during low flow conditions. Others, such as temperature and stream flows, should be measured nearly continuously (hourly). Both parameters are projected to change with climate change and therefore it is important to see how they vary temporally and with enhancement actions. Biological parameters related to adults and redds are measured only during periods when those life-stages are present in the system. Juvenile parameters, on the other hand, need to be measured seasonally to see how the treatments affect their seasonal abundance, distribution, habitat use, and movement. Measuring juvenile parameters only during low flow conditions may miss an important bottleneck in the life-cycle of the fish. For example, a given enhancement action may improve habitat conditions for fish during the summer low-flow period, but a bottleneck during winter may erase any benefits accrued during summer (Mason 1976). As a result, smolt production does not change even though parr production increased during the summer.

2.6 Monitoring Methods

Methods for measuring the physical and biological parameters have evolved considerably over time. There are dozens of protocols available for monitoring physical and biological parameters. We recommend that the methods for measuring parameters be based largely on the Salmon Recovery Funding Board (SRFB) approach[6] (O’Neal 2007; Crawford 2011a-e; also see www.monitoringmethods.org). The reason for this is because the SRFB approach has been used successfully in evaluating habitat actions at the project and reach scales (O’Neal et al. 2016). In addition, this approach is consistent with the programmatic approach for the Columbia Basin Fish and Wildlife Program funded by Bonneville Power Administration (BPA) and the Oregon Watershed Enhancement Board (OWEB) approach. Consistency among monitoring programs allows the results from different programs to be evaluated in concert without the concern that different monitoring protocols influence the results. This increases the number of BA, BACI, and EPT replicates, which allows for a more robust evaluation of different enhancement action types. That is, studies conducted by the YN can be pooled with studies conducted by other monitoring programs. This not only increases replicates, but improves cost efficiencies.

As a final note, as indicated in Table 3, we recommend the evaluation of juvenile fish movement as it relates to habitat enhancement. We believe it is important to understand how enhancement actions influence the movement and residence times of juvenile salmonids. It is also important to determine if enhancement actions increase fish abundance or simply redistribute juvenile fish among sites. To the extent possible, we recommend mark-recapture techniques based on passive integrated transponder (PIT) tags. Analysis of PIT tags allows assessment of seasonal habitat use, movements, residence times, and survival. This is important in IPT and intensive BA and BACI studies, and is currently lacking in most effectiveness monitoring programs.

---

[6] The SRFB approach was largely adapted from the U.S. Environmental Protection Agency’s Environmental Monitoring and Assessment Program (EMAP) (Peck et al. 2003) and the Integrated Status and Effectiveness Monitoring Program (ISEMP 2013).

2.7 Data Management, Quality Assurance, and Analysis

Data management is a key part of any monitoring program. Because we recommend data collection that closely follows the SRFB approach, it is appropriate to use their field data sheets for recording data. To minimize transcription errors, we recommend the use of computer clipboards and electronic data loggers. These tools greatly simplify entering data into a database. In addition, feedback loops can be programmed into the recording devices to make sure all data fields are populated with the appropriate data.

The database needs to be set up to handle large volumes of raw data and calculate common metrics associated with each measured parameter (see Appendix 1). As a starting point, we recommend that the database be consistent with the SRFB database. Additional functions may be needed to store and process mark-recapture data; although, regional databases such as PIT Tag Information System (PTAGIS) and Columbia River Data Access in Real Time (DART) are available to help store and process PIT-tag information. In addition, the database should be set up to calculate estimates of precision and variability for each metric.

The metrics calculated within the database are then used in statistical analyses. These analyses are used to determine if the enhancement actions at the reach or project scale had a beneficial effect on the fish and their habitat. There are several different analytical approaches that can be used to evaluate effectiveness monitoring data, including parametric approaches, non-parametric approaches, multivariate approaches, randomization or Monte Carlo techniques, and Bayesian analysis. Which one to use will depend on the monitoring design, degree of pseudoreplication, data type (continuous, discrete, ratio, interval, ordinal, or nominal), and statistical hypothesis. For example, paired treatment and control data collected from post-treatment studies can be analyzed with paired *t*-tests (or its non-parametric alternative, Wilcoxon test). Regression techniques can be used to relate differences in response variables among sites with other measured parameters. Analysis of variance (ANOVA) models or time series analyses are often used to analyze data collected under BACI designs (Downes et al. 2002). Regardless of the statistical analyses used, all effectiveness monitoring data need to be analyzed graphically. Graphical analysis places more emphasis on “practical” or “biological” significance and less on “statistical” significance. A simple time series plot with error bars reveals much about the effectiveness of enhancement actions.

2.8 Reporting

Given the cost of enhancement actions, it is important to communicate the effectiveness of those actions to both the scientific community and the public. Regardless if an enhancement project is a success or failure, it is important to report the findings. Often failures go unreported. This is unfortunate given the money spent on enhancement work. To avoid making the same mistakes in the future, it is just as important to report failed efforts as successes. To that end, we recommend the preparation of annual reports, which describe the status of the project and identify any shortcomings in the design, sampling methods, and analyses that may need to be revisited and corrected. Annual reports should also provide necessary adaptive feedback on the effectiveness of the enhancement action. Tetra Tech (2016)[7] provides an example of annual reporting on project effectiveness. Once the project is complete, a final report should be generated describing results and recommendations.

---

[7] http://www.rco.wa.gov/documents/monitoring/2015AnnualProgressReport.pdf

Results need to be reported in a format interpreted easily by scientists, managers, and policy makers. Managers and policy makers tend to favor graphical displays and analyses, while scientists and technical experts prefer statistical analyses. Thus, we recommend a combination of the two reporting styles because it is the most effective way to convey results to a broad interdisciplinary audience.

Because the approach for monitoring effectiveness in this framework document is consistent with other effectiveness monitoring programs in the region, it is important to collaborate with those other programs, share information, and report findings of combined studies. As noted above, this increases samples size (replicates), allows for interpretation of varying results (e.g., why a given action worked in one place but not in another), and increases the transferability of results (i.e., the results can be generalized to other areas)[8].

Finally, to the degree possible, project or reach-scale results need to be presented in a way that they can be incorporated into higher-level assessments (e.g., population status assessments and vulnerability assessments). This is accomplished by reporting changes in physical and biological metrics at the project or reach scale. Physical metrics inform vulnerability assessments (see Crozier et al. in prep), while fish metrics (adult and juvenile abundance, density, and distribution) help inform viable salmonid population (VSP) parameters such as abundance and spatial structure. Reporting movement and migration characteristics of juvenile fish as they relate to enhancement actions informs population diversity parameters. Lastly, any survival information collected during mark-recapture studies can be used to inform population productivity. Importantly, data collected at the project or reach scale only inform population parameters, they do not replace them.[9] VSP parameters are measured at the population scale.

2.9 Coordination

As discussed throughout this document, it is important to coordinate monitoring activities with other monitoring programs. This monitoring framework is designed to integrate with and complement other effectiveness monitoring programs such as SRFB and OWEB effectiveness monitoring programs, BPA's Programmatic monitoring program, Bureau of Reclamation monitoring program, and the Northwest Fisheries Science Center monitoring program (e.g., Integrated Status and Effectiveness Monitoring Program or ISEMP). This will lead to efficiencies in cost and effort. A major benefit of integrating and coordinating with the other effectiveness monitoring programs is the larger sample size for many enhancement types. The increased sample size will improve ability to detect significant changes in habitat and fish due to enhancement actions. Close coordination with other programs will also allow the YN to evaluate enhancement actions that are not adequately covered in other programs.

---

[8] A study that evaluates the effectiveness of a given action within a specific location is referred to as a "case study." If implemented correctly, a case study can tell us much about the effectiveness of that action in that location (in statistical terms, this type of study has high "internal" validity). However, the results from that study may not be transferable to other locations (the study has low "external" validity). If the study is repeated in many locations with similar results, the action under investigation can be generalized to other locations without the need to monitor it (high external validity). Comparing results among different programs requires that the studies were conducted using similar protocols.

[9] Researchers need to be careful not to over-extrapolate results from project and reach-scale effectiveness monitoring studies to help estimate VSP and vulnerability parameters. Rarely are enhancement sites selected randomly and therefore they may not be representative of the larger assessment unit or population. Thus, the results collected at the project or reach scale should not be expanded to the entire assessment unit or population.

It is important to point out that there are several large-scale status and trend monitoring programs and hatchery evaluation programs within the upper Columbia River basin. These include the Columbia Habitat Monitoring Program (CHaMP), Okanogan Basin Monitoring and Evaluation Program (OBMEP), Aquatic and Riparian Effectiveness Monitoring Program (AREMP), Pacfish/Infish Biological Opinion (PIBO), and Public Utility District (PUD), Tribal, and Federal-funded hatchery monitoring and evaluation programs. These programs collect information on several of the parameters called for in this document. In addition, these programs may be collecting data within one or more of the reaches in which project or reach-scale enhancement monitoring will occur. Therefore, it is important to coordinate with these programs to share information, evaluate cost-sharing opportunities, coordinate field work, and collect like information.

As a final important point on coordination, the Pacific Northwest Aquatic Monitoring Partnership (PNAMP) Regional Habitat Indicator Project (RHIP) is working to understand what management questions are important to organizations throughout the Pacific Northwest in order to focus discussions about indicators, data, and data sharing (https://www.pnamp.org/project/3149). They have identified four topic areas as a focus for their initial efforts, including stream flow, macroinvertebrates, stream temperature, and water quality. With the development of common habitat indicators among the monitoring partners in the Columbia River Basin, monitoring data can be shared to support multiple uses. Much of the data necessary for the YN monitoring program may be available through existing partner monitoring efforts, and partners may rely on YN efforts to augment their own programs.

2.10 Implementation

The preceding sections serve notice that considerable care must be put into the methods and logic structure of an effectiveness monitoring plan. The intent of this section is to distill the information presented above into a concise outline that one can follow to develop a plan to assess the effectiveness of specific enhancement actions at the project or reach scale. For convenience, we offer this summary as a checklist of steps that will aid in the development of valid monitoring plans. Although these steps are generic, each one should be completed to develop a valid effectiveness monitoring plan.

2.10.1 PROGRAM SETUP CRITERIA

In order to setup a monitoring plan, it is important to follow a logical sequence of steps. These steps should aid in implementing a valid monitoring plan that reduces duplication of sampling effort and reduces overall costs.

Setup Steps:

1. Identify the populations and/or subpopulations of interest (e.g., spring Chinook, coho salmon, steelhead, bull trout, etc.).
2. Identify the geographic boundaries (areas of the populations or sub-groups of interest.
3. Describe the purpose for selecting these populations or sub-groups (what are the concerns?).
4. Identify the objectives for monitoring.
5. Identify and review existing monitoring and research programs in the area of interest.
6. Determine if those programs satisfy the objectives of the proposed plan.

7. If data gaps exist, implement the appropriate monitoring approach by following the criteria outlined in Section 2.10.2 below.
8. Classify the landscape and streams in the area of interest (see Section 2.4).
9. Describe how data collection efforts will be shared among different entities.
10. Identify the database for storing biological and physical data.
11. Estimate costs of implementing the plan.
12. Identify cost-sharing opportunities.

2.10.2 PROJECT/REACH-SCALE EFFECTIVENESS MONITORING CRITERIA

If the objective of the monitoring program is to assess the effectiveness of tributary habitat actions at the project or reach scale, then the following steps will help in designing a valid effectiveness monitoring plan.

Problem Statement and Overarching Issues:

1. Identify and describe the problem to be improved or corrected by the action being monitored.
2. Describe current environmental conditions at the project site.
3. Describe factors or threats contributing to current conditions (e.g., road crossing causing increased siltation).
4. Identify and describe the habitat action(s) (treatments) to be undertaken to improve existing conditions.
5. Describe the goal or purpose of the habitat action(s).
6. Identify the hypotheses to be tested.
7. Identify the independent variables in the study.

Statistical Design (see Sections 2.2, 2.4, & 2.7):

1. Describe the statistical design to be used (e.g., BA, BACI, IPT, EPT design).
2. Determine if the study will include “true” replicates or subsamples.
3. Describe how temporal and/or spatial controls will be used and how many of each type will be sampled and for how long.
4. Describe the independence of treatment and control sites (are control sites completely unaffected by habitat actions?).
5. Identify covariates and their importance to the study.
6. If this is a pilot study, explain why it is needed (e.g., to collect information to be used in power analyses).
7. Describe descriptive and inferential statistics to be used and how precision of statistical estimates will be calculated.
8. Describe graphical methods to be used to demonstrate treatment effects.

Sampling Design (see Sections 2.5):

1. Describe the statistical population(s) to be sampled.
2. Define and describe sampling units.
3. Describe the number of sampling units (both treatment and control sites) that make up the sampling frame.

4. Describe how sampling units will be selected (e.g., random, stratified, systematic, etc.).
5. Define "practical significance" (e.g., environmental or biological effects of the action) for this study.
6. Describe how effect size(s) will be detected.
7. Describe the variability or estimated variability of the statistical population(s).
8. Define Type I and II errors to be used in statistical tests (we recommend no less than 0.80 power).

Measurements (see Sections 2.3, 2.6, & 2.7):

1. Identify and describe the parameters (dependent variables) to be measured.
2. Describe methods and instruments to be used to measure parameters.
3. Describe the precision of measuring instrument(s).
4. Describe possible effects of measuring instruments on sampling units (e.g., electrofishing and fish handling may affect their movement and survival). If such effects are expected, describe how the study will deal with this.
5. Describe steps to be taken to minimize systematic errors.
6. Describe QA/QC plan, if any.
7. Describe sampling frequency for field measurements.

Results (see Section 2.8):

1. Explain how the results of this study will yield information relevant to management decisions.

These steps should be considered when designing a monitoring plan to assess the effectiveness of any habitat action, regardless of how simple the proposed enhancement action may be. In some cases, it may not be possible to address all steps with a high degree of certainty, because adequate information does not exist. For example, the investigator may lack information on population variability, effect size, "practical significance," or instrument precision, which will make it difficult to design studies and estimate sample sizes. In this case, one can address the statements with the best available information, even if it is based on professional opinion, or design a pilot study to answer the questions.

3 Watershed-Scale Effectiveness Monitoring

Because the overall goal of the enhancement actions implemented by the YN are to increase fish performance in various populations, it is therefore important to assess the cumulative effects of habitat actions on fish performance (e.g., abundance, distribution, productivity, and survival) at the watershed or population scale. Population/watershed-scale effectiveness monitoring, also known as intensively monitored watershed (IMW), is designed to measure benefits associated with tributary habitat actions at the watershed or population scale.[10] Although this level of effectiveness monitoring measures benefits at the watershed or population scale, unless coupled with reach-scale effectiveness monitoring, population/watershed-scale effectiveness monitoring cannot by itself determine the effects of specific habitat actions on populations if more than one type of action is implemented. Therefore, wherever possible, reach-scale effectiveness monitoring should occur within populations in which watershed-scale effectiveness monitoring also occurs. Using causal-comparative approaches, one can then determine which actions had the greatest effect on fish performance at the population scale.

Although many of the elements described in Section 2 apply to watershed-scale effectiveness monitoring, there are significant challenges of monitoring or evaluating enhancement actions at the watershed scale that are not as important at the reach or project scale. The greatest challenges in developing watershed-scale monitoring programs are design, site selection, selecting appropriate monitoring parameters, and coordination (aka procedural challenges; Reid 2001).[11] Fortunately, there are published guidelines to consider when designing a watershed-scale monitoring program (Roni et al. 2015). What follows is a framework for developing specific monitoring plans to assess the effectiveness of tributary habitat enhancement actions at the project or reach scale.

3.1 Monitoring Questions

At the population/watershed-scale, the YN is interested in answering the following questions.

What are the effects of habitat actions on overall watershed habitat quality and quantity?

> Implementing habitat actions at the reach scale throughout a watershed should result in improvements in the overall quality and capacity of the habitat within the watershed or population. This question addresses the cumulative effect of all habitat actions implemented within the watershed or population. If the suite of actions implemented within the watershed or population do not improve overall habitat quality, it is unlikely that survival or capacity will increase at a watershed scale.

---

[10] In this document, we use population/watershed-scale effectiveness monitoring and intensively monitored watershed (IMW) interchangeably.

[11] Although there are some important differences between project or reach-scale and watershed-scale effectiveness monitoring programs, we encourage readers to first read Section 2 before reading this section. Many of the terms used in this section are described in Section 2 and those descriptions are not repeated here.

What are the effects of habitat actions on fish performance at a watershed or population level?

> Answers to this question will determine if fish abundance, distribution, and/or survival increased as a result of the implementation of habitat actions within the watershed or population. Success is determined by comparing fish metrics to a control or reference watershed, baseline conditions, or desired future conditions. These comparisons are needed to determine whether it was the implementation of habitat actions and not some other extraneous factor that caused the increase in fish performance.

Which actions contributed most to the increase in fish performance at the watershed scale?

> As noted earlier, without reach-scale effectiveness monitoring, population/watershed-scale effectiveness monitoring cannot determine which action contributed most to a change in survival if more than one habitat action type was implemented (Roni et al. 2015).

What are the relationships between fish performance and overall habitat quality?

> The monitoring designs identified in this document will allow the collection of fish performance data under a wide range of habitat conditions. Thus, data from watershed-scale effectiveness monitoring should provide data that can be used to develop relationships between changes in habitat quality and fish performance.

Results from both reach-scale and population/watershed-scale effectiveness monitoring efforts will help determine if the habitat actions implemented are providing the expected fish benefits and which actions contributed most to the benefits. It is possible that one may find a positive benefit at the reach scale, but not at the watershed or population scale. This may be due to other confounding factors including other uncontrolled or unaccounted for management activities, the scale of the action, the amount of enhancement, or other factors limiting the population at a broader scale. Under this scenario, one needs to evaluate results from population and habitat status and trend monitoring. By considering all the monitoring data in concert, one should be able to determine if other factors confounded the effects of habitat actions. Alternatively, reach-scale benefits may not translate into population-scale benefits because the amount of enhancement that occurred in the watershed was not large enough to produce a population-level response. Under this scenario, if only a small percentage of the watershed has been improved, fish performance may have increased at a reach scale, but when rolled up to a watershed scale the increase in performance is not large enough to be measured at a population scale. In fact, one study indicated that more than 20% of a watershed would need to be improved to measure a population/watershed-scale response to enhancement (Roni et al. 2010). Finally, it is possible that some reaches could show positive increases in fish, but a bottleneck or limiting factor exists in a downstream reach that causes mortality.

The general questions above can be transcribed into specific questions depending on the goals of the watershed-enhancement goals. Answers to specific questions will provide managers the information they need to identify the most effective actions for addressing habitat impairments at the watershed or population scale.

3.2 Monitoring Design

Evaluation of watershed-scale changes in fish and habitat generally requires BA, BACI, or post-treatment designs (EPT or IPT), or some combination of these. As noted in Section 2.2, BA designs require collection of fish and habitat data before and after implementation of enhancement actions, while BACI designs not only require collection of data before and after implementation of enhancement actions within the treated watershed, but also within a control or reference watershed. The control watershed should be as similar as possible to the treated watershed. For example, in the John Day River basin, Bouwes et al. (2016a) used River Styles (Brierley and Fryirs 2005; O'Brien and Wheaton 2015) to match Murderers Creek (control watershed) with Bridge Creek (treatment watershed). Using a BACI design, they demonstrated that treatments in Bridge Creek increased juvenile steelhead production (product of density, growth, and survival) in Bridge Creek by 175% relative to the control watershed.

More complex versions of BA designs such as staircase designs or hierarchical-staircase designs (Walters et al. 1988) can be used to assess reach-scale and watershed-scale responses to habitat enhancement projects. These designs require the sequential treatment of several study reaches or watersheds over time. Pre-treatment data are collected on all reaches or watersheds, then each reach or watershed is treated sequentially over time, with multiple reaches or watersheds serving as untreated controls. Data from these designs are generally analyzed using general linear models (Walters et al. 1988). The staircase design is useful when applied at large scales and is currently being used in the Asotin IMW, Bridge Creek IMW, and the Entiat IMW (ISEMP 2013).

3.3 Monitoring Parameters

As with project or reach-scale effectiveness monitoring, which physical and biological parameters need to be measured depends on goals and key questions, monitoring design, and available tools and resources to measure parameters. We recommend that researchers select a small but comprehensive set of parameters to monitor basin-scale enhancement. Metrics derived from the parameters need to be diagnostic in nature and capable of detecting which aspects of the watershed have improved by different enhancement actions. These metrics should represent physical and biological endpoints of enhancement and they should also capture landscape and watershed processes that form and sustain stream ecosystems (Beechie et al. 2009).

Following Roni et al. (2015), we recommend a short list of parameters to consider when designing a watershed-scale effectiveness monitoring plan (Table 4). We identify parameters for key watershed processes such as hydrology, sediment delivery, riparian parameters, and biological parameters. We also identify the scales at which these parameters are measured and analyzed. It is important to point out that not all the parameters identified in Table 4 should be measured in every watershed-scale evaluation. Rather, this list represents the core set of parameters to consider initially. However, parameters such as temperature, discharge, and riparian vegetation need to be measured in all treatment and control watersheds, because these parameters are expected to change in response to climate changes. The parameters identified in Table 4 are directly linked to the ecological concerns identified by Hamm (2012).

Table 4. Suggested list of parameters to monitor watershed-scale responses to enhancement actions and the appropriate level of monitoring and inference. Although detection of watershed-scale changes generally requires randomly or systematic sampling, selecting sites to evaluate project or reach-scale enhancement activities typically requires non-random selection of treatment and control sites (see Section 2). NA = not applicable. This table is from Roni et al. (2015) with modifications.

CATEGORY	PARAMETER	SAMPLING SCALE	SAMPLE SITE EXTENT	SAMPLING FREQUENCY	LEVEL OF INFERENCE
Hydrology	Discharge	Target reaches affected by enhancement actions	Point sample	Continuous gauging	Watershed
Coarse and fine sediment	Sediment supply (sediment budget)	Entire watershed	Entire watershed	5-10 year intervals	Watershed
Fine sediment	Plot-scale sediment yield	Stratified random or stratified systematic reaches	Road segment or field plot	Continuous	Watershed, sites
Riparian conditions	Species composition, stem density, size distribution	Stratified random reaches	Reach	Decadal for watershed-scale trends, annual for newly planted sites	Watershed, reach
Channel morphology	Pool frequency	Randomly selected stream reaches	Reach	Annually	Reach or watershed
Channel morphology	Wood abundance and volume	Randomly selected stream reaches	Reach	Annually	Reach or watershed
Channel morphology	Residual pool depth	Target response reaches downstream of enhancement actions	All reaches downstream of enhancement work	Annual or semi-annual	Watershed
Water quality	Temperature	Randomly selected reaches	Reaches	Hourly (data loggers)	Reach
Fish	Juvenile	Stratified or systematic random reaches	Reach	Seasonal	Reach and watershed
Fish	Smolts	Entire watershed	Point sample downstream of enhancement work	Seasonal (during smolt migration)	Watershed
Fish	Adults or redds	Entire watershed or stratified or systematic random reaches	Entire watershed or reach	Seasonal (during adult spawning)	Watershed or reach

3.4 Replication

As described in Section 2.4, a power analysis is generally used to determine the number of spatial and/or temporal replicates needed to detect a meaningful change in biological and physical conditions. This is also true at the watershed scale. In general, watershed-scale effectiveness monitoring programs are long-term endeavors. Roni et al. (2015) reported that for salmonids, more than 10 years of data are needed to detect changes in fish abundance of 25% or more for BACI studies in paired watersheds. However as little as 4 years of monitoring (2 years before and 2 years after treatment) could be needed for a study using a BACI design with extensive spatial replication (20 watersheds). Bouwes et al. (2016a) was able to detect significant effects on steelhead performance at the watershed scale based on a BACI design with 3 years of pre-treatment data and 4 years of post-treatment data. As a starting point, we recommend that researchers collect 2 or more years of pre-treatment data for BA and BACI designs. Depending on the extent, timing, and types of the treatments, 10 or more years of data may need to be collected post-treatment. As data are collected during the early stages of the monitoring program, those data can be used in power analyses or sample size calculations to determine the appropriate level of replication needed.

3.5 Sampling Scheme

The parameters to be measured will largely determine the spatial allocation of sampling within the watersheds. Table 4 indicates the scheme, scale, extent, and frequency for sampling each parameter that could be measured to assess the effectiveness of enhancement actions at the watershed scale. Because samples collected within a watershed need to be representative of the watershed and will ultimately be rolled up to the watershed scale, they must be collected using unbiased sampling techniques. Depending on the parameter, simple random sampling, stratified random sampling, systematic sampling, or combinations of these (e.g., GRTS) can be used to collect representative samples. As described in Section 2, selected sites should vary in size according to the mean bankful width. That is, a sampling site should be 20 times the mean bankful width, but not less than 150-m long or longer than 500 m. When evaluating overall watershed conditions, we urge caution in using data collected from project or reach-scale effectiveness monitoring programs. Although including these data increases the overall sample size for describing watershed conditions, these data are rarely collected from randomly selected sites. That is, sites selected for evaluating project or reach-scale enhancement actions are rarely selected randomly and therefore cannot be used easily to describe overall watershed conditions.[12]

Except for discharge and temperature, most physical parameters only need to be measured once per year (or less frequently; see Table 4). Biological parameters, on the other hand, may require more frequent sampling during the year. Both adults (redds) and smolts are measured during the period when they are present. Juveniles need to be measured seasonally to see how the treatments affect their seasonal abundance, distribution, and habitat use at the watershed scale. Sampling juveniles only during low flows may miss an important bottleneck in the life-cycle of the fish. For example, a given enhancement action may improve habitat conditions for fish during the summer low-flow period, but a bottleneck during winter may erase any benefits accrued during summer. As a result, smolt production does not change even though parr production increased during the summer.

3.6 Monitoring Methods

As indicated in Section 2.6, there are numerous protocols available for monitoring physical and biological parameters at the watershed scale. For consistency, we recommend that most of the methods described in Section 2.6 be used to assess the effectiveness of habitat actions at the watershed scale. Different methods, however, are needed to measure sediment yield and sediment supply at the

---

[12] Importantly, however, results from project or reach-scale effectiveness monitoring sites can be used to help infer which enhancement actions affected watershed-scale changes (see Section 3.1 and 3.7).

watershed scale.[13] Consistency between project or reach-scale monitoring and watershed-scale monitoring, and among other monitoring programs, allows results from different programs to be evaluated in concert without the concern that different monitoring protocols influence the results. This also increases the number of replicates, which allows for a more robust evaluation of different enhancement action types and improves cost efficiencies. Detailed descriptions of monitoring methods can be found at www.monitoringmethods.org.

3.7 Data Management, Quality Assurance, and Analysis

Regardless of the scale of effectiveness monitoring, data management is a critical part of monitoring. Because the methods used to measure parameters at the watershed scale are mostly consistent with those used at the project or reach scale, we recommend a common database for both programs. For all scales of monitoring, we recommend the use of computer clipboards and electronic data loggers to minimize transcription errors. These tools greatly simplify entering data into a database. In addition, feedback loops can be programmed into the recording devices to make sure all data fields are populated with the appropriate data.

As described in Section 2.7, the database needs to be set up to handle large volumes of raw data and calculate common metrics associated with each measured parameter (see Appendix 1). Additional functions may be needed to store and process mark-recapture data (e.g., DART and PTAGIS). Importantly, the database should be set up to calculate estimates of precision and variability for each metric.

The metrics calculated within the database are then used in statistical analyses. These analyses are used to determine if the enhancement actions at the watershed scale had a beneficial effect on the fish and their habitat. The statistical methods described in Section 2.7 can be used to evaluate effectiveness monitoring data. Which one to use will depend on the monitoring design, degree of pseudoreplication, data type (continuous, discrete, ratio, interval, ordinal, or nominal), and statistical hypothesis. Regardless of the statistical analyses used, all effectiveness monitoring data need to be analyzed graphically. Simple graphical analysis may be more informative especially when conveying results to managers and executives.

As we stated earlier, it is important to link project or reach-scale effectiveness monitoring with watershed scale monitoring. Unless there is only one action type implemented within a watershed, it will be difficult to determine which action or actions caused a measurable response at the watershed scale. By implementing project or reach-scale effectiveness monitoring, one can determine which actions contributed most to any measurable change at the watershed scale. A great deal of coordination is needed to tease out treatment effects if there are many different types of actions (including non-habitat-related activities such as harvest and hatcheries) implemented within the treated watershed (see Section 3.9 below).

3.8 Reporting

It is critical to report the results from effectiveness monitoring to the scientific community and the public. This is true even if habitat enhancement work was a failure. As with project or reach-scale effectiveness monitoring, we recommend the preparation of annual reports describing the status of the

---

[13] Reid and Dunne (1996) provide methods for developing sediment budgets.

enhancement work at the watershed scale. The report should identify any shortcomings in the implementation design, monitoring design, sampling methods, and analyses that may need to be revised or corrected. Annual reports should also provide necessary adaptive feedback on the effectiveness of the enhancement action. The reports should include both statistical analysis and graphical analysis. Once the project is complete, a final report should be generated describing results and recommendations.

Because watershed-scale effectiveness monitoring provides physical and biological information at the watershed or population scale, these data or metrics need to be presented so they can be used in population status assessments and vulnerability assessments. This is accomplished by reporting biological metrics associated with abundance, productivity, diversity, and spatial structure (VSP parameters) and physical metrics associated with vulnerability assessments. Reporting movement and migration characteristics of juveniles, smolts, and adults as they relate to enhancement actions informs population diversity parameters. Adult abundance, age structure, and origin (hatchery or wild) data can be used to estimate population abundance and productivity metrics. In addition, survival information collected during mark-recapture studies can be used to inform population productivity. Spawning and juvenile fish distribution, as they relate to habitat enhancement work, can be used to inform spatial structure.

NOAA Fisheries is completing regionwide climate-change vulnerability assessments for Pacific salmon (Crozier et al., in prep). They have identified four salmon-specific freshwater exposure attributes that will most likely challenge salmon populations because of changing climate. These include hydrologic regime, storm frequency/flood magnitude, stream temperature in August, and climate water deficit. Tracking changes in physical metrics at the watershed scale, especially temperature, discharge, and riparian conditions, can be used to evaluate the effects of climate change and habitat actions on fish habitat in the larger context of NOAA's regionwide vulnerability assessments.

3.9 Coordination

Undoubtedly the most difficult aspect of watershed effectiveness monitoring is coordination. Monitoring enhancement actions across a watershed requires extensive coordination of management actions – a procedural problem rather than a technical problem. Hillman et al. (2017) evaluated IMWs throughout the Pacific Northwest and found that procedural problems often limit the success of watershed monitoring programs. In general, the more organizations involved in implementing monitoring and enhancement work, the more challenging is the coordination of efforts and sharing of data. In addition, there are often delays in summarizing and analyzing data for annual reports. If data are not analyzed and summarized on an annual basis, a project can go on for many years before critical errors in design or data collection are detected (Roni et al. 2015).

The Entiat IMW is an example of a watershed-scale effectiveness monitoring program that is challenged by coordination among restoration, monitoring, funding entities, and landowners. The Entiat IMW relies on a staircase design, which has been largely compromised because of landowner issues. This has resulted in the interruption of the placement and timing of enhancement actions within the Entiat River. In contrast, the Asotin and Bridge Creek IMWs, which also rely on staircase designs, have been quite successful, because, in part, there are few entities involved with the programs (Bennett et al. 2016). In addition, these are relatively small watersheds. The larger the watershed and the more complex the actions, the more critical coordination becomes and the more time that must be dedicated to it.

Within the upper Columbia River basin, there are several large-scale monitoring programs and hatchery evaluation programs evaluating habitat conditions and populations. These include CHaMP, ISEMP, OBMEP, AREMP, PIBO, and PUD, Tribal, and Federal-funded hatchery monitoring and evaluation programs. These programs collect information at the watershed scale on several of the parameters called for in this document. In fact, most of the subbasins within the upper Columbia have some form of large-scale monitoring. Thus, any watershed-scale effectiveness monitoring program implemented within the upper Columbia must coordinate closely with existing monitoring programs.

As described earlier, the Pacific Northwest Aquatic Monitoring Partnership (PNAMP) Regional Habitat Indicator Project (RHIP) is working to understand what management questions are important to organizations throughout the Pacific Northwest (https://www.pnamp.org/project/3149). With an understanding of management questions, they will be able to focus discussions on indicators, data, and data sharing. They have identified four topic areas as a focus for their initial efforts, including stream flow, macroinvertebrates, stream temperature, and water quality. With the development of common habitat indicators among the monitoring partners in the Columbia River Basin, monitoring data can be shared to support multiple uses. Much of the data necessary for the YN monitoring program may be available through existing partner monitoring efforts, and partners may rely on YN efforts to augment their own programs.

3.10 Implementation

The intent of this section is to distill the information presented above into a concise outline that one can follow to develop a plan to assess the effectiveness of specific enhancement actions at the watershed scale. We offer this summary as a checklist of steps that will aid the investigator in developing a valid watershed monitoring plan. These steps are similar to those identified for project or reach-scale effectiveness monitoring (see Section 2.10).

3.10.1 PROGRAM SETUP

In order to setup a monitoring plan, it is important to follow a logical sequence of steps. These steps should aid in implementing a valid monitoring plan that reduces duplication of sampling effort and reduces overall costs.

Setup Steps:

1. Identify the populations and/or subpopulations of interest (e.g., spring Chinook, coho salmon, steelhead, bull trout, etc.).
2. Identify the geographic boundaries (areas of the populations or sub-groups of interest.
3. Describe the purpose for selecting these populations or sub-groups (what are the concerns?).
4. Identify the objectives for monitoring.
5. Identify and review existing monitoring and research programs in the area of interest.
6. Determine if those programs satisfy the objectives of the proposed plan.
7. If data gaps exist, implement the appropriate monitoring approach by following the criteria outlined in Section 3.10.2 below.
8. Classify the landscape and streams in the area of interest (see Section 3.4).

9. Describe how data collection efforts will be shared among different entities.
10. Identify the database for storing biological and physical data.
11. Estimate costs of implementing the plan.
12. Identify cost-sharing opportunities.

3.10.2 WATERSHED-SCALE EFFECTIVENESS MONITORING

If the objective of the monitoring program is to assess the effectiveness of tributary habitat actions at the watershed scale, then the following steps will help in designing a valid effectiveness monitoring plan.

Problem Statement and Overarching Issues:

1. Identify and describe the problems within the watershed to be improved or corrected by the actions being monitored.
2. Describe current environmental conditions at the watershed scale.
3. Describe factors or threats contributing to current conditions (e.g., road density, upland activities, etc.).
4. Identify and describe the habitat action(s) (treatments) to be undertaken to improve existing conditions.
5. Describe the goal or purpose of the habitat action(s).
6. Identify the hypotheses to be tested.
7. Identify the independent variables in the study.

Statistical Design (see Sections 3.2, 3.4, & 3.7):

1. Describe the statistical design to be used (e.g., BA, BACI, etc.).
2. Describe the methods used to identify control watersheds.
3. Describe the independence of treatment and control watersheds (are control watersheds completely unaffected by habitat actions?).
4. Identify covariates and their importance to the study.
5. Describe potential threats to the study (e.g., harvest or hatchery operations that could confound estimation of population metrics).
6. If this is a pilot test, explain why it is needed.
7. Describe descriptive and inferential statistics to be used and how precision of statistical estimates will be calculated.
8. Describe graphical methods to be used to demonstrate treatment effects.

Sampling Design (see Sections 3.5):

1. Describe the statistical population(s) to be sampled.
2. Define and describe sampling units within the watersheds.
3. Describe the number of sampling units that make up the sampling frame.
4. Describe how sampling units will be selected (e.g., random, stratified, systematic, etc.).
5. Define “practical significance” (e.g., environmental or biological effects of the action) for this study.
6. Describe how effect size(s) will be detected.
7. Describe the variability or estimated variability of the statistical population(s).

8. Define Type I and II errors to be used in statistical tests (we recommend no less than 0.80 power).

Measurements (see Sections 3.8):

1. Identify and describe the parameters (dependent variables) to be measured.
2. Describe methods and instruments to be used to measure parameters.
3. Describe the precision of measuring instrument(s).
4. Describe possible effects of measuring instruments on sampling units (e.g., electrofishing and fish handling may affect their movement and survival). If such effects are expected, describe how the study will deal with this.
5. Describe steps to be taken to minimize systematic errors.
6. Describe QA/QC plan, if any.
7. Describe sampling frequency for field measurements.

Results:

1. Explain how the results of this study will yield information relevant to management decisions.

These steps should be carefully considered when designing a monitoring plan to assess the effectiveness of habitat actions at the watershed scale. In some cases, the investigator may not be able to address all steps with a high degree of certainty, because adequate information does not exist. For example, the investigator may lack information on population variability, effect size, "practical significance," or instrument precision, which makes it difficult to design studies and estimate sample sizes. In this case, the investigator can address the statements with the best available information, even if it is based on professional opinion, or design a pilot study to answer the questions.

4 Adaptive Management

We believe "active" adaptive management[14] is a critical part of any enhancement program and therefore effectiveness monitoring plans need to be designed within the context of adaptive management. Enhancement actions applied under an adaptive management framework will be the most efficient way to understand the effectiveness of enhancement work. In short, adaptive management is an iterative process of exploring uncertain outcomes to management actions while making progress toward broader management goals (Walters and Holling 1990). In general, the cycle of adaptive management includes plan, do, evaluate, and learn (Bouwes et al. 2016b). In the context of stream enhancement work, the hallmark of adaptive management is to adjust either the implementation plan or the enhancement actions based on effectiveness monitoring. Without monitoring, adjustments are simply based on trial and error.

An excellent example of applying active adaptive management to watershed-scale enhancement work is described in Bouwes et al. (2016b). They show that each step in the adaptive management cycle is critically important to the overall success of watershed (or reach-scale) restoration. They noted during the "Planning" phase that it is important to use existing watershed assessments, literature reviews, and targeted field studies to identify the "problems" within the watershed. Here, "problems" refer to both limiting factors (aka ecological concerns) and threats. This information is then used to develop enhancement goals and hypotheses, identify enhancement actions, development implementation and monitoring plans, and, if necessary, model potential enhancement scenarios. The "Doing" phase of adaptive management includes the implementation of the monitoring and enhancement plans. Here, coordination among the many organizations (including landowners) is needed, especially if the monitoring plan calls for the collection of pre-treatment data. As noted earlier, information from monitoring is used to assess whether the objectives of the program are being met and if there are any unforeseen consequences causing harm that may need to be adjusted.

Monitoring provides the feedback loop that is used during the "Evaluation" and "Learning" phases of adaptive management. Annual evaluations of monitoring data are used to determine if the right problem was identified, if the enhancement work and plan are achieving the predicted responses, if enhancement work is causing harm to the resources, if the intensity of monitoring is appropriate, and if the most important parameters are being measured.

In the Upper Columbia, there are several planning and assessment processes that have been implemented, including subbasin plans, recovery plans, a biological strategy, watershed assessments, and reach assessments. In addition, modeling efforts such as life-cycle modeling, habitat modeling, food-web modeling, and EDT have been conducted to assess potential effects of habitat enhancement actions in the Upper Columbia. These processes are all part of the "Planning" phase of adaptive management and they are a critical component of this monitoring framework document. By implementing specific monitoring plans developed under the guidance of this framework document, information will be generated that will improve decision making, which completes the adaptive management cycle.

---

[14] "Active" adaptive management implements actions with the goal to maximize learning or reduce uncertainties that inform management actions. It is needed to understand causal mechanisms of responses (Williams 2011). In contrast, "passive" adaptive management uses models and existing knowledge to describe the most likely action to achieve management goals. Learning is an unintended consequence of passive adaptive management.

In conclusion, this document provides a framework for developing effectiveness monitoring plans at the project or reach scale and at the watershed scale. Implementing monitoring plans within an active adaptive management framework will be the most efficient and powerful way to determine if enhancement activities are working. Although adaptive management can be complex and daunting (e.g., see Appendix Q to the Upper Columbia Spring Chinook Salmon and Steelhead Recovery Plan), the approach described above is relatively simple and has led to a successful enhancement program in the Asotin Creek watershed (Bouwes et al. 2016b).

5 References

Beechie, T., G. Pess, M. Pollock, M. Ruckelshaus, and P. Roni. 2009. Pages 695-716 *in* R. Beamish and B. Rothschild, editors. Restoring rivers in the 21st century: science challengers in a management context. The future of fisheries science in North America. Springer-Verlag, Heikelberg.

Beechie, T., H. Imaki, J. Greene, A. Wade, H. Wu, G. Pess, P. Roni, J. Kimball, J. Stanford, P. Kiffney, and N. Mantua. 2013. Restoring salmon habitat for a changing climate. River Research and Applications 29:939-960.

Bennett, S., G. Pess, N. Bouwes, P. Roni, R. Bilby, S. Gallagher, J. Ruzycki, T. Buehrens, K. Krueger, W. Ehinger, J. Anderson, C. Jordan, B. Bowersox, and C. Greene. 2016. Progress and challenges of testing the effectiveness of stream restoration in the Pacific Northwest using intensively monitored watersheds. Fisheries 41:92-103.

Bouwes, N., N. Weber, C. Jordan, C. Saunders, I. Tattam, C. Volk, J. Wheaton, and M. Pollock. 2016a. Ecosystem experiment reveals benefits of natural and simulated beaver dams to a threatened population of steelhead (*Oncorhynchus mykiss*). Science Report 6, 28681; doi: 10.1038/srep28581.

Bouwes, N., S. Bennett, and J. Wheaton. 2016b. Adapting adaptive management for testing the effectiveness of stream restoration: an intensively monitored watershed example. Fisheries 41: 84–91.

Brierley, G., and K. Fryirs. 2005. Geomorphology and River Management: Applications of the River Styles Framework. Blackwell Publishing, Victoria, Australia.

Crawford, B. 2011a. Protocol for monitoring effectiveness of fish barrier removal projects. Washington Salmon Recovery Funding Board, MC-1, Olympia.

Crawford, B. 2011b. Protocol for monitoring effectiveness of in-stream habitat projects (channel reconfiguration, deflectors, log and rock control weirs, roughened channels, and woody debris removal). Washington Salmon Recovery Funding Board, MC-2, Olympia.

Crawford, B. 2011c. Protocol for monitoring effectiveness of riparian planting projects. Washington Salmon Recovery Funding Board, MC-3, Olympia.

Crawford, B. 2011d. Protocol for monitoring effectiveness of riparian livestock exclusion projects. Washington Salmon Recovery Funding Board, MC-4, Olympia.

Crawford, B. 2011e. Protocol for monitoring effectiveness of floodplain enhancement projects. Washington Salmon Recovery Funding Board, MC-5/6, Olympia.

Crozier, L. 2014. Impacts of climate change on salmon of the Pacific Northwest: a review of the scientific literature published in 2014. Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, WA.

Crozier, L., M. McClure, M. Haltuch, E. Hazen, S. Bograd, and M. Nelson. In prep. West Coast Salmon Climate Vulnerability Assessment. NOAA Fisheries, Seattle, WA.

Downes, B.., L. Barmuta, P. Fairweather, D. Faith, M. Keough, P. Lake, B. Mapstone, and G. Quinn. 2002. Monitoring ecological impacts. Cambridge University Press, Cambridge, UK.

Hamm, D. 2012. Development and evaluation of a data dictionary to standardize salmonid habitat assessments in the Pacific Northwest. Fisheries 37:6-18.

Hester, E., M. Doyle, and G. Poole. 2009. The influence of in-stream structures on summer water temperatures via induced hyporheic exchange. Limnology and Oceanography 54:355-367.

Hillman, T. 2006. Monitoring strategy for the Upper Columbia Basin. BioAnalysts, Inc. Report to the Upper Columbia Regional Technical Team, Upper Columbia Salmon Recovery Board, Wenatchee, WA.

Hillman, T., P. Roni, and J. O’Neal. 2017. Effectiveness of tributary habitat enhancement projects. Report to Bonneville Power Administration, Portland, OR.

Isaak, D. and M. Young. 2015. Identifying, protecting, and enhancing climate refugia for salmonids. United States Forest Service, Rocky Mountain Research Station, Boise, ID.

ISEMP (Integrated Status and Effectiveness Project). 2013. The integrated status and effectiveness monitoring program: lessons learned synthesis report, 2012 update. Report to Bonneville Power Administration, Project 2003-00-017, Portland, OR.

Keefer, M., M. Moser, C. Boggs, W. Daigle, and C. Peery. 2009. Variability in migration timing of adult Pacific lamprey (*Lampetra tridentate*) in the Columbia River, U.S.A. Environmental Biology of Fishes 85:253-264.

Mason, J. 1976. Response of underyearling coho salmon to supplemental feeding in a natural stream. Journal of Wildlife Management 40:775-788.

O'Brien, G., and J. M. Wheaton. 2015. River Styles Report for the Middle Fork John Day Watershed, Oregon - Example Report for Exploring Leveraging the River Styles Framework in Tributary Habitat Management for the Columbia River Basin. Ecogeomorphology and Topographic Analysis Lab, Utah State University, Prepared for Eco Logical Research and the Bonneville Power Administration, Logan, UT.

O’Neal, J. 2007. Snorkel surveys. Pages 325–340 in D. H. Johnson, B. M. Shrier, J. S. O’Neal, J. A. Knutzen, X. Augerot, T. A. O’Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook. American Fisheries Society, Bethesda, Maryland.

O’Neal, J., P. Roni, B. Crawford, A. Ritchie, and A. Shelly. 2016. Comparing stream restoration project effectiveness using a programmatic evaluation of salmonid habitat and fish response. North American Journal of Fisheries Management 36:681-703.

Peck, D. V., J. M. Lazorchak, and D. J. Klemm. 2001. Environmental monitoring and assessment program—surface waters: western pilot study field operations manual for wadeable streams. Draft Report. EPA/XXX/X-XX/XXX, U.S. Environmental Protection Agency, Washington, D.C. Web link:

http://www.epa.gov/emap/html/pubs/docs/groupdocs/surfwatr/field/ewwsm01.html

Polivka, K., E. Steel, and J. Novak. 2015. Juvenile salmon and steelhead occupancy of stream pools treated and not treated with restoration structures, Entiat River, Washington. Canadian Journal of Fisheries and Aquatic Sciences 72:166-174.

Reid, L. 2001. The epidemiology of monitoring. Journal of American Waters Resources Association 34:815-820.

Reid, L. and T. Dunne. 1996. Rapid evaluation of sediment budgets. GeoEcology Paperback, Catena Verlag, Reiskirchen, Germany.

Reynolds, L., A. Herlihy, P. Kaufmann, S. Gregory, and R. Hughes. 2003. Electrofishing effort requirements for assessing species richness and biotic integrity in Western Oregon streams. North American Journal of Fisheries Management 23:450-461.

Roni, P., editor. 2005. Monitoring stream and watershed restoration. American Fisheries Society, Bethesda, Maryland.

Roni, P., and T. Quinn. 2001. Density and size of juvenile salmonids in response to placement of large woody debris in western Oregon and Washington streams. Canadian Journal of Fisheries and Aquatic Sciences 58:282-292.

Roni, P., M. Liermann, C. Jordan, E. Steel. 2005. Steps for designing a monitoring and evaluation program for aquatic restoration. Pages 13-34 *in* P. Roni, editor. Monitoring stream and watershed restoration. American Fisheries Society, Bethesda, Maryland.

Roni, P., S. Morley, P. Garcia, C. Detrick, I. King, and E. Beamer. 2006. Coho salmon smolt production from constructed and natural floodplain habitats. Transactions of the American Fisheries Society 135:1398–1408.

Roni, P., and T. Beechie (eds.). 2013. Stream and watershed restoration: A guide to restoring riverine processes and habitats. Wiley-Blackwell, Chichester, UK.

Roni, P., T. Beechie, C. Jordan, and G. Pess. 2015. Basin scale monitoring of river restoration: recommendations from case studies in the Pacific Northwest USA. American Fisheries Society Symposium 78:73-98.

Simonson, T., J. Lyons, and P. Kanehl. 1994. Quantifying fish habitat in streams: transect spacing, sample size, and a proposed framework. North American Journal of Fisheries Management 14:607-615.

Stevens, D. 2002. Sampling design and statistical analysis methods for the integrated biological and physical monitoring of Oregon streams. Report No. OPSW-ODFW-2002-07, The Oregon Plan for Salmon and Watersheds, Oregon Department of Fish and Wildlife, Corvallis, OR.

Tetra Tech. 2016. 2015 Annual Monitoring Report, Reach-Scale Effectiveness Monitoring Program. Report to the Washington State Recreation and Conservation Office, Salmon Recovery Funding Board, Olympia, WA.

Thompson, S. 1992. Sampling. John Wiley and Sons, New York. N.Y.

Walters, C. and C. Holling. 1990. Large-scale management experiments and learning by doing. Ecology 71:2060-2068.

Walters, C., J. Collie, and T. Webb. 1988. Experimental designs for estimating transient responses to management disturbances. Canadian Journal of Fisheries and Aquatic Sciences 45:530-538.

Williams, B. 2011. Passive and active adaptive management: approaches and an example. Journal of Environmental Management 92:1371-1278.

Appendix 1:

Common Metrics Associated with Measured Physical and Biologic Parameters

PARAMETERS	COMMON METRICS CALCULATED
Physical Parameters and Metrics	
Temperature	Mean, maximum, minimum, range, number of days greater than threshold
Nitrogen	Ammonia, nitrate/nitrites, total nitrogen
Phosphorus	Total phosphorus, total orthophosphates
Migr. Barriers	Number of road crossings, diversion dams, fishways
Substrate	% fines, % of different types, D_{50}, D_{84}
Embeddedness	% embeddedness
Fines	% fines, pool-tail fines
Woody debris	Number/100 m, volume/100 m
Pool density	% pools, % riffles, channel widths/pool
Pool depths	Mean pool depth
Fish cover	% fish cover
Reach length	Total length, total area
Sinuosity	Sinuosity
Wetted width	Mean width, total area
Bankful width	Mean bankful width, width/depth ratio
Bank erosion	% of length of eroding bank
Riparian struct.	% canopy layer, % understory layer, % ground cover
Riparian disturb.	Presence and proximity of disturbance
Canopy cover	% shade
Streamflow	Mean, maximum, minimum, range
Biological Parameters and Metrics	
Adult number	Total number, density
Adult origin	Total number and density of natural-origin and hatchery-origin fish
Redd number	Total number, density
Redd distribution	Spatial location, habitat use

PARAMETERS	COMMON METRICS CALCULATED
Fry number	Total number by species, density by species (fish/m^2)
Fry distribution	Spatial location, habitat use
Juv. number	Total number by species, density by species (fish/m^2)
Juv. distribution	Spatial location, habitat use
Juv. size	Fish length, fish weight
Juv. movement	Rates, timing, residence
Smolt number	Total number by species
Smolt size	Fish length, fish weight, fish condition
Smolt migr. timing	Mean, 10%, 50%, 90%, range, and distribution of migration (Julian days)