Implementation and Effectiveness Monitoring Plan for the Hydraulic Project Approval Program


George Wilhere, Kirk Krueger, Phillip Dionne, and Timothy Quinn

Draft, February 2014

Washington Department of FISH and WILDLIFE

For more information contact:

George Wilhere
Habitat Program
Washington Department of Fish and Wildlife
Olympia, WA 98501
e-mail: george.wilhere@dfw.wa.gov
Phone: 360-902-2369

Preferred citation:
Wilhere, G.F., K. Krueger, P. Dionne, and T. Quinn. 2014. Implementation and Effectiveness Monitoring Plan for the Hydraulic Project Approval Program. Habitat Program, Washington Department Fish and Wildlife, Olympia, Washington.

Table of Contents

Introduction

One of the main responsibilities of WDFW's Habitat Program is protecting fish habitats through the administration and enforcement of the hydraulic code rules (Chapter 220-110 Washington Administrative Code)[1]. Through these rules WDFW regulates the construction of hydraulic projects or the performance of other work that will use, divert, obstruct, or change the natural flow or bed of any of the salt or fresh waters of the state. The rules set forth procedures for obtaining a hydraulic project approval (HPA, i.e., a permit), and the rules incorporate criteria used by WDFW for project review and conditioning HPAs. Furthermore, the hydraulic code rules reflect the best available science and practices related to the protection of fish life, and WDFW will incorporate new information into the rules as it becomes available (WAC 220-110-010).

The immense importance of the Habitat Program's responsibility to protect fish habitats was recently highlighted by the "Culvert Case." In 2007, Federal District Judge Ricardo S. Martinez ruled that the treaties with Indian tribes forbade Washington State from constructing and maintaining highway culverts that blocked salmon migration (Blumm and Steadman 2009, Morisset and Summers 2009). In 2013 the same judge ruled on the remedy to the plaintiffs complaint – the State must replace all state-owned culverts in western Washington (Lovaas 2013). There are hundreds of state-owned culverts blocking passage to hundreds of miles of in-stream fish habitats. The estimated cost for this remedy is over $1 billion. State-owned culverts are but a fraction of the thousands of state, county, city, and private culverts that block fish passage. The cost of repairing or replacing these culverts will be billions of dollars. The enormous cost of restoring fish habitat lost to impassable culverts makes plainly evident the absolute necessity that hydraulic projects be compliant with current rules and that current rules be effective at protecting fish habitats.

To help ensure that hydraulic projects are compliant with current rules and that current rules effectively protect fish habitats, WDFW is developing a hydraulic project monitoring program. The first systematic review of hydraulic project compliance was conducted by the Habitat Program in 2005 (Quinn et al. 2006). That review found that over 60% of permitted projects were not fully compliant with the hydraulic code rules. A later study by the Habitat Program (Price et al. 2010) found that 30% of culverts (23 of 77) permitted under the HPA process for fish passage were barriers to fish movement, and that most culvert failures were due to noncompliance with permit provisions, particularly culvert slope. The Habitat Program's first programmatic compliance monitoring effort was initiated in 2010 (Habitat Program 2011). The first year of monitoring evaluated compliance for permits issued in 2008 and 2009 for water crossing structures, freshwater bank protection, marine shoreline bank protection, and marine overwater structures. The results for culverts were similar to previous studies – at least 35% of culverts were noncompliant for at least one structural dimension of the culvert (Table 1). The next year of monitoring (Habitat Program 2012) found higher rates of compliance than previous studies: about 81% for projects completed in 2010. The lowest rates of compliance were found for freshwater bank protection (about 72%) and the highest rates for marine shoreline bank protection (about 91%).

The rates of compliance found by Quinn et al. (2006), Price et al. (2010), and Habitat Program (2010) for HPAs are disconcerting, and the most recent compliance study (Habitat Program (2011) estimated that one out of five hydraulic projects are noncompliant. These studies used different approaches to investigate different aspects of HPA compliance, and hence, are not strictly comparable. However, they all indicate that compliance with HPA permits has been unacceptably low. Increasing compliance will rely, in part, on better information with which to track and improve WDFW's performance with respect to HPAs. The main purpose of HPA monitoring is provide information with which to improve over time

[1] The hydraulic project statutes may be the oldest environmental protection laws in Washington State; the first rules were established in 1943.

both compliance with current hydraulic code rules and the effectiveness of those rules at protecting fish habitats. Specifically, the purpose of monitoring the HPA program is to provide reliable, useful information that describes:

1. opportunities to improve WDFW's process for issuing HPA permits,
2. opportunities to improve compliance by permittees,
3. success and failure rates of protecting fish habitats by hydraulic projects that are compliant with HPA permits, and
4. the characteristics of compliant hydraulic projects that are commonly associated with failures to protect fish habitats.

Table 1. Results of compliance monitoring conducted in 2010 for culverts completed in 2008 or 2009 (Habitat Program 2011). Includes no-slope and stream simulation culverts.

Structural Dimension of Culvert	Sample Size	Percent Noncompliant
length	23	35
slope	24	25
countersinking at inlet	19	42*
countersinking at outlet	17	29*
width compared to width in HPA permit	24	8
width compared to width in WDFW's design guidelines	24	33†

* These values may be invalid because prior to measurement 85% of culverts had been exposed to high water events which may have removed material from the culvert.

† Some noncompliant culverts may be caused by imprecise measurements of bankfull width.

Adaptive Management

Monitoring implies adaptive management. If there is no intention to improve management activities then it is pointless to monitor them. Monitoring is the feedback loop providing information for management decisions in the adaptive management process (Figure 1). The adaptive management process is a continual cycle consisting of planning, action, monitoring, evaluation, and adjustment (Bormann et al. 1994, Wilhere 2002). Monitoring and evaluation form the foundation for adaptive management, however, a "monitoring plan" alone cannot fulfill the goal of adaptive management – i.e., the continual improvement of management. Ultimately, the Habitat Program will need a comprehensive "HPA Adaptive Management Plan" that integrates every phase of adaptive management and considers how information collected through monitoring will lead to changes in the HPA process or hydraulic code rules. The Habitat Program will develop an adaptive management plan in the near future, and developing that plan will be a collective endeavor of policy makers, managers, field staff, engineers, and scientists.

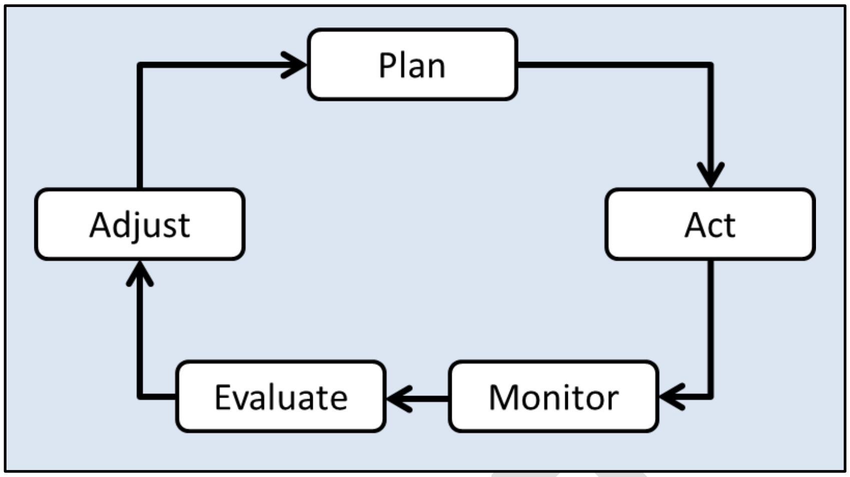


Figure 1. The adaptive management cycle (modified from Bormann et al. 1994).

Monitoring Versus Research

Monitoring and research are scientific activities that contribute to adaptive management. Both can entail the systematic, objective, empirical testing of hypotheses. The differences between monitoring and research often lie in their goals, purposes, and their relationships to management operations. As part of the feedback loop in adaptive management, the primary goal of monitoring is to provide information about regular management operations. The word "monitor" is derived from the Latin word *monēre*, which means to warn. Hence, monitoring provides "a warning" when management practices are not achieving desired objectives. The primary goal of research is to acquire fundamental knowledge about natural phenomena (basic research) or about innovative management practices (applied research). Research leads to the facts and principles upon which regular management operations are based. Monitoring can also acquire fundamental knowledge and basic research can also provide information on regular operations, but these goals are secondary. Because monitoring collects essential information in the adaptive management cycle, monitoring should be integrated into regular management operations. In many cases research is separate from regular management operations. However, under a comprehensive adaptive management system, applied research might test innovative practices through integration into regular management operations.

An important purpose of monitoring is protection of natural resources – water quality, fish, wildlife, etc. With this purpose in mind, monitoring focuses on human impacts to those resources. The science of environmental monitoring began with the measurement of air and water pollution for the purposes of protecting air and water quality. From these beginnings, the classical monitoring question became, "Are human activities damaging the environment?" In contrast, the ultimate purpose of basic research is to understand and explain the environment. Unlike monitoring which focuses on human impacts to the environment, research typically focuses on the environment itself. The differences in purpose lead to different perspectives on error avoidance.

Research prefers false negative claims to false positive claims, and monitoring prefers false positive to false negative claims (Table 2). That is, research scientists try to avoid claiming a hypothesis is true when it is actually false. Advancing invalid hypotheses can harm science. Hence, research scientists are willing to accept some false negatives – claims that a hypothesis is false when it is actually true. Monitoring should avoid false negatives − situations where no impact was detected but a significant environmental impact actually occurred. When monitoring environmental impacts, conscientious managers should be willing to tolerate some false positives – situations where a significant impact was detected but an impact did not actually occur. Since avoiding all errors is impossible, scientists must settle for some rate of error. Research scientists often impose a rather arbitrary standard of 5 percent error for false negatives. Meeting a lower error rate costs more money because it requires a larger sample size (i.e., more site visits). Monitoring often feels the pinch of fiscal constraints, so designers of a monitoring plan must deal with a compromise between error rate and budget. For the purposes of monitoring, higher error rates for false positives may be acceptable. This compromise is more fully explained in the monitoring design section.

Table 2. Possible outcomes of statistical inference. Inference can result in Type I or Type II error. The state of reality "no change occurred" is typically known as the null hypothesis (H_0).

Inference from Data	**Reality**	
	No Change Occurred	**Change Occurred**
No Change occurred	Correct	False Negative (*Type II Error*)
Change Occurred	False Positive (*Type I Error*)	Correct

Types of Monitoring

A hierarchy consisting of three types of monitoring – implementation, effectiveness, and validation – has become a common organizational framework for monitoring programs in natural resource management (MacDonald et al. 1991, USDA and USDI 1994). All three types of monitoring provide essential feedback for adaptive management. Implementation monitoring simply determines whether or not hydraulic projects are implemented properly. Implementation monitoring is not compliance monitoring, although the two are related. Compliance monitoring focuses exclusively on the performance of a permittee. Implementation monitoring is broader in scope; it monitors the performance of both the permittee and permitor. Using the information collected through implementation monitoring, the entire HPA process may be improved to achieve a higher level of compliance.

Effectiveness monitoring is done to determine whether or not hydraulic projects are yielding the desired habitat conditions. For water crossing structures, the desired condition is "no-net-loss of productive capacity of fish and shellfish habitat" (WAC 220-110-070). "Net loss" refers to the "net loss of habitat functions necessary to sustain fish life" and the "loss of area by habitat type" (220-110-020(68)). For marine shoreline armoring the desired condition is no "permanent loss of critical food fish or shellfish habitat" (WAC 220-110-285). Effectiveness monitoring compares consequent habitat conditions resulting from a hydraulic project with the desired habitat conditions. Assuming that HPA provisions were implemented correctly, repeated failures to achieve desired conditions at multiple sites would suggest that hydraulic code rules are not protecting fish habitats. Effectiveness monitoring, for example, might track physical changes at a many separate beaches following the construction of marine shoreline armoring and compare changes over time at beaches with HPA compliant bulkheads to physically similar beaches without bulkheads.

Validation monitoring tests the validity of our assumptions regarding the association of a species to the desired habitat conditions. In other words, validation monitoring measures a species response to the desired habitat conditions, and it is done to determine whether or not particular a species responds to the desired habitat conditions as anticipated. For example, validation monitoring might test the hypothesis that new bulkheads compliant with hydraulic code rules will not significantly alter the density or survival of forage fish eggs on adjacent beaches, or validation monitoring might test the hypothesis that the stream simulation culvert design is passible by all life stages of all native fish species during stream flows when fish are known to be moving through the stream network.

This hierarchal organization of implementation, effectiveness, and validation monitoring implies that the quality of inferences made at one level depend on the results at a lower level. That is, the inferences of effectiveness monitoring will be in error if the results of implementation monitoring are in error. For instance, measuring the effectiveness of stream simulation culverts that are thought to be implemented improperly but actually were not would lead to erroneous conclusions about the behavior of stream simulation culverts. Data collected through implementation monitoring will be used to screen which HPAs are compliant, and hence, can be used for effectiveness monitoring. Problems will most certainly arise if data collection at a higher level of monitoring is not coordinated with data collection at a lower level. Therefore, during development of an overall monitoring program the teams charged with developing each type of monitoring will communicate regularly and cooperate on monitoring activities.

Hydraulic Project Monitoring

Hydraulic project monitoring will be integrated into the HPA process (Figure 2). For example, because implementation monitoring measures the immediate outcome of the permitting and construction processes, implementation monitoring should occur before the structure is subjected to strong natural disturbances that could alter the dimensions of the hydraulic structure (i.e., shortly after project completion). Therefore, implementation monitoring will be triggered by notification from the permittee that a project is complete. Such notification will be required through a provision in every HPA. In addition, data collected by habitat biologists during pre-permit site visits and post-construction compliance inspections and recorded in the Aquatic Protection Permitting System (APPS) may supplement data collected through HPA monitoring. Because HPA monitoring will rarely collect pre-construction data at project sites, measurements taken during pre-permit site visits (e.g., bankfull width, bulkhead location) may be particularly useful.

Implementation Monitoring

Current funding levels limit the number of hydraulic projects that can be monitored. Consequently, HPA monitoring will focus on those types of hydraulic projects that have the potential to cause the greatest adverse impacts to fish habitats. Implementation monitoring will address only the following types of projects:

- culverts, both new and replacement on fish bearing streams
- marine shoreline bank protection, with an emphasis on new armoring

Implementation monitoring of freshwater bank protection (rivers and streams only) and marine overwater structures will be conducted when more resources become available.

Implementation monitoring will not cover all provisions of an HPA permit. Implementation monitoring will focus on those provisions that can: 1) be evaluated post-construction; 2) be objectively measured, and hence, do not require the specialized expertise of a habitat biologist; and 3) require only one site visit. For instance, provisions related to construction timing[2] or equipment[3] will not be evaluated because they

[2] Example of timing provision: Work below the ordinary high water line shall only occur between [ADD DATE HERE] and [ADD DATE HERE].
[3] Example of equipment provision: Equipment crossings of the stream are not authorized by this HPA.

cannot be reliably evaluated post-construction, and provisions related to re-vegetation[4] will not be evaluated because they require either a subjective expert judgment or more than one site visit. In other words, implementation monitoring will focus on the hydraulic structures covered under a permit – culverts and bank armoring – because relative to other activities regulated under the HPA permit (e.g., re-vegetation), hydraulic structures have the greatest potential to adversely impact fish habitats and are the principal activity being regulated.

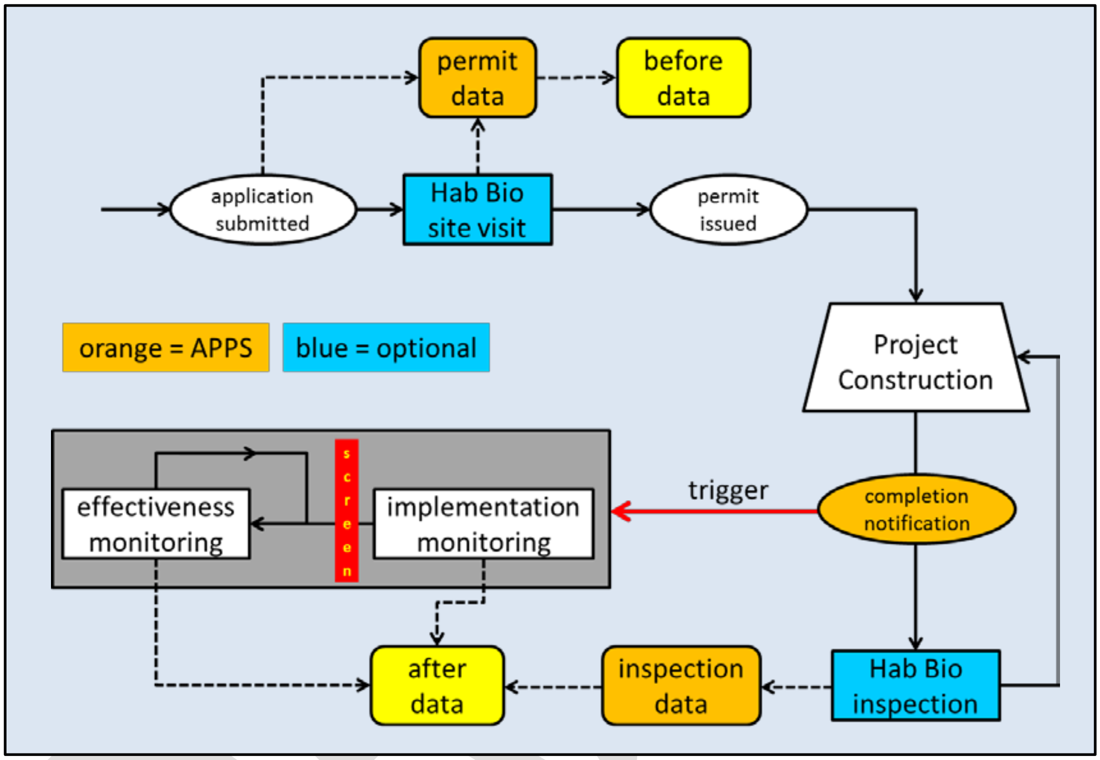


Figure 2. Relationship of hydraulic project monitoring to the HPA process. Solid lines represent process flow and dashed lines represent data flow. APPS is the Aquatic Protection Permitting System. Notification of project completion triggers monitoring at the hydraulic project site. A screen is applied between implementation and effectiveness monitoring. Projects with failures detected through implementation monitoring will not be considered for effectiveness monitoring.

Implementation monitoring determines whether HPAs were implemented properly. Implementation monitoring examines completed hydraulic projects in order to document successes and failures in the permitting or construction processes. Hydraulic projects are implemented by people. Implementation monitoring focuses on the performance of people engaged in the permitting and construction processes – permittors and permittees. Hence, we define two forms of implementation failure: 1) permittor failure in which an HPA permit fails to meet or exceed design standards in WAC 220-110 or WDFW's design

[4] Example of re-vegetation provision: Alteration or disturbance of the bank and bank vegetation shall be limited to that necessary to construct the project. Within seven calendar days of project completion, all disturbed areas shall be protected from erosion using vegetation or other means. Within one year of project completion, the banks, including riprap areas, shall be re-vegetated with native or other approved woody species. Vegetative cuttings shall be planted at a maximum interval of three feet (on center) and maintained as necessary for three years to ensure 80 percent survival.

guidance (e.g., Bates et al. 2003, Cramer et al. 2003, Johannessen et al. 2013); and 2) permittee failure in which a completed project fails to meet or exceed design specifications stipulated in an HPA permit (Figure 3). A hydraulic project can present one or both forms of failure.

Permittor Failure. There are 3 types of permittor failure: a) a permit does not comply with WAC 220-110; b) a permit approves a structure that is inappropriate for the project site, i.e., the approved structure does not follow WDFW's design guidance; and c) the inaccuracy or absence of key measurements necessary for project design and submitted with the plans result in a structure that is inappropriate for the project site.

To evaluate the first type of permittor failure, the provisions stipulated in the HPA and the construction plans attached to the HPA will be compared to design standards in WAC 220-110.[5] For example, permits or plans for no-slope culverts must specify: 1) culvert slope of approximately 0%; 2) culvert countersinking of at least 20%; 3) minimum culvert width at the stream bed elevation; and 4) average width of the stream bed. The third item must be greater than or equal to the fourth item. We assume that hydraulic project construction follows the HPA permit and the plans attached to it. Therefore, if any of the four specifications do not comply with WAC 220-110, then the first type of permittor failure has occurred. If such a failure occurs, then the hydraulic project is dropped from further monitoring – i.e., no data collection at the site will occur − but the project will be referred to the local habitat biologist for further action.

To evaluate the second type of permittor failure, the provisions stipulated in the HPA and the construction plans attached to the HPA will be compared to design guidelines (e.g., Bates et al. 2003, Cramer et al. 2003, Johannessen et al. 2013). For a no-slope culvert, for example, if the channel gradient at the project site is greater than 3%, then a no-slope design is inappropriate (Bates et al. 2003). If a no-slope is approved at a site with channel gradient greater than 3%, then an implementation failure has occurred and that hydraulic project is dropped from further monitoring. If such a failure is detected simply by reading the plans, then no data collection will occur at the site.

The third type of permittor failure occurs when the HPA and attached plans comply with WAC 220-110, the approved structure appears to be appropriate for the project site, but inaccurate or incomplete measurement of site conditions resulted in a structure that is actually inappropriate for the project site. For example, plans submitted with an application for a no-slope culvert could show that the channel gradient is 2.5%. The maximum gradient for a no-slope culvert is 3%. If that measurement is correct, then the no-slope design is appropriate for that site. However, if the channel slope is actually 3.5%, then an implementation failure has occurred. Key measurements necessary for project design should be submitted with the plans, and through implementation monitoring we will determine the accuracy of key measurements. If key measurements do not meet a minimum level of accuracy (e.g., ± 5% error) and the inaccuracy results in a structure that is inappropriate for the project site, then an implementation failure has occurred. The measurement error was committed by applicant/permittee, but ultimate responsibility for validity of the permit rests with the permittor, hence, this type of failure is a permittor failure.

Permittee Failure. To evaluate the second form of implementation failure, the constructed hydraulic project will be compared to the provisions in the HPA and the plans attached to the HPA. For the no-slope culvert example, if culvert slope, culvert counter sinking, or culvert width do not meet or exceed the provisions of the permit (including the attached plans), then the second form of implementation failure has occurred. If such a failure occurs, then the hydraulic project is dropped from further monitoring.

[5] Recall that implementation monitoring checks only those provisions associated with the hydraulic structure covered under the permit.

Permittee failure can be subject to action by WDFW's Enforcement Program, however, the purpose of implementation monitoring is not rule enforcement. Potential rule violations detected through implementation monitoring will be reported to the habitat biologist responsible for the HPA. Compliance inspections done by habitat biologists have been and will continue to be the mechanism for determining violations of the hydraulic code rules.

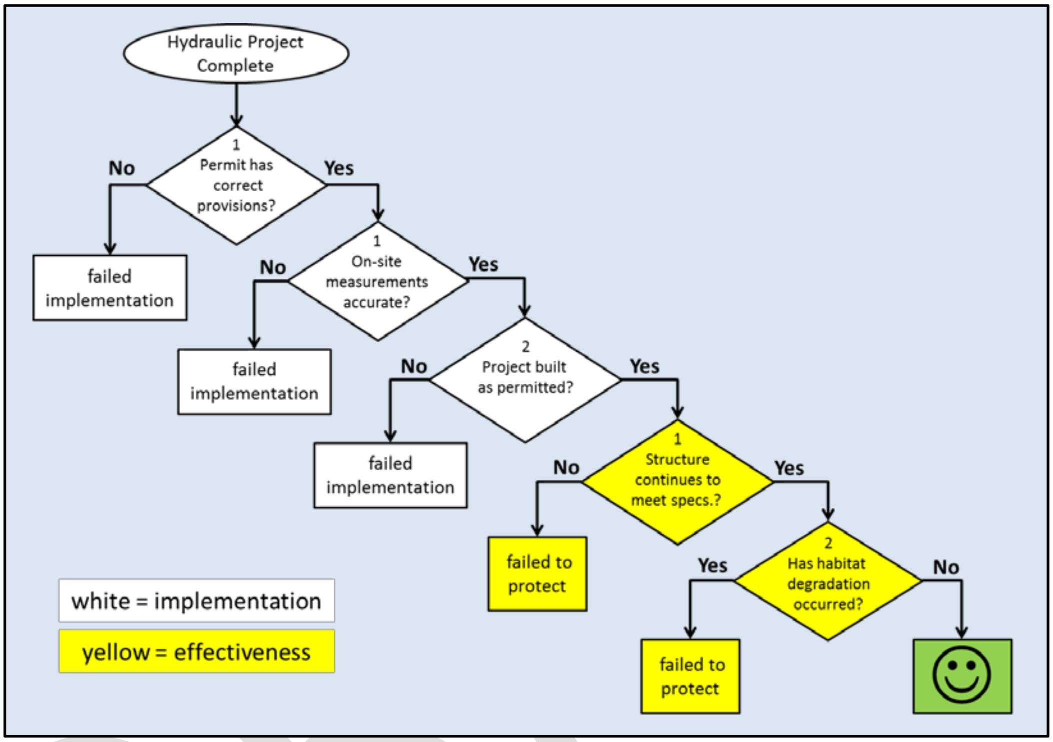


Figure 3. Decision tree for implementation and effectiveness monitoring at a single site. HPA monitoring answers a sequence of questions. A failure for any question terminates monitoring of that hydraulic project. There are two types of implementation failure: 1 indicates permittor error and 2 indicates permittee error. There are two types of effectiveness failure: 1 indicates hydraulic structure failure and 2 indicates habitat degradation.

Accuracy, Precision, and Tolerance. Accuracy is how close a measurement is to the "true" value of a quantity. Precision is how close repeated measurements of the same quantity are to each other. Many physical measurements associated with hydraulic projects – bankfull width, channel gradient, location of ordinary high water line – tend to be inaccurate and/or imprecise. Measurements such as bankfull width require challenging subjective judgments that result in high inter-observer variability and hinder objective quantification. In other words, such measurements are inherently imprecise and have no "true" value.

Implementation monitoring must take into account the inexact nature of the measurements associated with the permitting and construction processes. This is done through measurement and engineering tolerances. A tolerance is the maximum acceptable difference between the actual value of a quantity and the value specified for it. For our purposes, measurement tolerances refer to the key measurements necessary for project design such as bankfull width, channel gradient, and location of ordinary high water line. An engineering tolerance is the acceptable difference between the actual physical dimension of a constructed structure and the dimension specified on the construction project's plan. In implementation monitoring, engineering tolerances will be applied to dimensions such as culvert gradient, culvert counter sinking, and bulkhead length.

Implementation failure rates will be sensitive to our error tolerances, and therefore, tolerances should be realistic and fair. Tolerance values are difficult to specify *a priori*, and hence, our preliminary measurement tolerances (Table A1) may change as we learn more about the capabilities of habitat biologists and HPA permittees or their contractors. Preliminary engineering tolerances for most structural dimensions will be ± 5% error (D. Ponder, WDFW, pers. comm.). One exception to this rule-of-thumb is no-slope culverts. No slope culverts should be installed at 0% slope, but according to Barnard et al. (2013) culvert slope should be no greater than 1% (using survey grade equipment) and a culvert installed at greater than 2% slope culvert is noncompliant. Like the slope tolerance for no-slope culverts specified by Barnard et al. (2013),WDFW engineers should specify engineering tolerances for critical dimensions associated with various types of hydraulic projects.

Effectiveness Monitoring

Current funding levels limit the number of hydraulic projects that can be monitored. Hence, effectiveness monitoring will address only the following types of projects:

- culverts, both new and replacement on fish bearing waters streams
- marine shoreline bank protection, new only

An effective hydraulic project is defined as a properly implemented hydraulic project that causes no net loss of fish habitats for the intended lifetime of the structure. Effectiveness failure occurs when a properly implemented hydraulic project causes a net loss of fish habitat. We define two forms of effectiveness failure: *structure failure* in which a hydraulic structure (e.g., culvert, bulkhead) that met design standards in WAC 220-110 or design criteria in WDFW's guidelines (e.g., Bates et al. 2003) immediately post-construction no longer meets those standards or criteria; and *habitat failure* in which a hydraulic structure causes a net loss of fish habitats. A hydraulic project can present one or both types of failure. A hydraulic project that experiences structure failure will be dropped from further effectiveness monitoring.

Structure Failure. An implicit assumption of the hydraulic code rules and WDFW's design guidelines is that properly implemented hydraulic projects result in no net loss of fish habitats. Effectiveness monitoring for structure failure will make the same assumption. That is, we will assume if a no-slope culvert, for instance, maintains the slope, counter-sinking, and bed width to bankfull width relationship that were specified in the HPA permit (including the attached plans), then that culvert has caused no net loss of fish habitats. In other words, we will assume that a no-slope culvert that continues to meet the provisions specified in its HPA permit continues to pass fish.

For culverts we define three levels of structure failure: 1) the culvert no longer meets the design standards in WAC 220-110 or the design criteria in WDFW's guidelines (e.g., Bates et al. 2003) that were stipulated in the HPA permit; 2) the culvert does not pass a level B barrier assessment (WDFW 2009); and 3) the culvert does not pass a level A barrier assessment, which is the worst type of failure. For marine shoreline armoring, structure failure occurs when the structure no longer meets design standards in WAC 220-110.

Habitat Failure. The hydraulic code rules implicitly assume that permitted hydraulic structures cause no net loss of fish habitats. Water crossing structures shall cause "no-net-loss of productive capacity of fish and shellfish habitat" (WAC 220-110-070), where "net loss" refers to the "net loss of habitat functions necessary to sustain fish life" and the "loss of area by habitat type" (220-110-020(68)). Marine shoreline armoring shall "not result in permanent loss of critical food fish or shellfish habitat" (WAC 220-110-285).

Hydraulic projects, such as culverts and shoreline bulkheads, may have adverse effects on fish habitats immediately downstream or upstream (or downdrift and updrift) of the permitted structure. Certain culverts, for instance, may pass fish but at high flows cause downstream channel scour or upstream fine sediment deposition which both degrade fish habitats. Ideally, effectiveness monitoring would determine whether a properly implemented hydraulic project has caused a net loss of fish habitats. However, due to natural variability in the physical structure of stream channels and beaches, and consequently, natural variability in fish habitats as well, determining whether a hydraulic structure caused a net loss of fish habitats is technically challenging because changes to habitats in the immediate vicinity of a hydraulic structure may: 1) not be caused by the structure, and 2) not constitute a net loss of habitat.

The main technical challenge faced by effectiveness monitoring is distinguishing habitat loss caused by permitted hydraulic structures from habitat loss caused by other factors. These other factors include natural stream channel variability, natural catastrophic events, the legacy of prior hydraulic structures, the effects of upstream land use changes, or long-term climate trends. Confounding factors such as these are particularly acute in aquatic ecosystems. Our task is analogous to a radio receiver. Change caused by a hydraulic structure is a "signal" and changes caused by all other factors are "noise." The challenge is to detect and accurately differentiate a signal which is being broadcast in an extremely noisy environment.

Comparing the monitoring of hydraulic structures to other types of environmental monitoring reveals serious technical challenges. Historically, much of the "science" of environmental monitoring was developed for monitoring industrial facilities that emit air or water pollutants. The problems faced by air and water quality monitoring appear minor when compared to effectiveness monitoring of hydraulic structures. First, industrial facilities are point sources of pollutants. In contrast, habitat change at a hydraulic structure can be caused by the cumulative effects of many activities occurring upstream or updrift. Hence, sorting out which activities are actually responsible for habitat loss is complicated. Second, when monitoring pollutants, separating anthropogenic changes from natural variation is obvious. Pollutants are entirely artificial and many have no analog in nature. In contrast, hydraulic structures cause changes that are quite similar to natural changes: bed scour, sediment deposition, and channel migration. Third, prior to start-up of an industrial facility, "before" conditions for air or water quality can be accurately determined and used to establish reliable reference conditions. When monitoring hydraulic structures, "before" data could be collected, but such data are unlikely to serve as a reliable reference; the "reference" data can be confounded by the legacy of prior hydraulic structures or by conditions resulting from past natural disturbances which have yet to equilibrate. Fourth, industrial facilities can precisely regulate their impacts – i.e., effluent production is under their control. Regulating the impacts of hydraulic structures is impossible because the degree of impact is ultimately a function of processes beyond our control – i.e., weather-related events, such as flooding or mass wasting. Finally, in environmental monitoring, a failure to control pollutant emissions occurs when the concentration of a pollutant exceeds a pre-determined numeric threshold. In effectiveness monitoring of hydraulic structures, a failure occurs when a structure causes a net loss of fish habitats, however, we have no pre-determined numeric threshold for "habitat loss", and, in fact, such a threshold may be undeterminable. In short, effectiveness monitoring of hydraulic structures is much more complicated than other types of environmental monitoring and will require a long-term commitment to data collection in order to obtain reliable information.

Extreme events in which culverts are catastrophically "blown out" during flood flows cause severe scouring of the channel bed, and obviously result in loss of downstream fish habitats. Effectiveness monitoring would classify such events as *structure failures*. Effectiveness monitoring for *habitat failure* is concerned with more subtle or gradual changes in channel or beach morphology. Loss of fish habitats caused by hydraulic structures will be indicated by changes in channel or beach morphology, however, changes in morphology are not necessarily a net loss of fish habitat. That is, change is not necessarily bad.

Evaluating the effectiveness of culverts is further complicated by the fact that culverts, especially no-slope culverts, are *expected* to cause changes in channel morphology both inside and outside the pipe. For no-slope culverts, WDFW's design guidelines (Barnard et al. 2013) state, "width of the bed inside the culvert is equal to the prevailing bankfull width." In other words, a no-slope culvert is designed to accommodate the bankfull flow. Bankfull flow may be functionally defined as the flow at which the stream is about to overtop its banks (Dunn and Leopold 1978). Bankfull flow is reported to occur every 1 to 2 years with a mean recurrence interval of about 1.5 years (Rosgen 1994, Mulvihill et al. 2009). Hence, a no-slope culvert should accommodate all flows with a recurrence interval of about 2 years or less (Figure 4). When a stream overtops its banks, a no-slope culvert cannot accommodate the entire stream flow. The culvert constrains stream flow which causes slower flow velocities immediately upstream and faster flow velocities immediately downstream of the culvert. Changes in stream flow velocity cause changes in sediment movement – greater deposition occurs immediately upstream and greater mobilization occurs immediately downstream of the culvert. As stream flows becomes larger the effects of the culvert become greater. Therefore, we expect to see changes in channel morphology immediately upstream and downstream of no-slope culverts subjected to large stream flow events, i.e., 5, 10, 20, 50-year events. The same logic is valid for stream simulation culverts. Stream simulation culverts are wider, and therefore can accommodate larger flows. However, there is a limit to the flows which stream simulation culverts can accommodate, and therefore, extreme stream flow events, i.e., 50 and 100-year events, are expected to alter stream channel morphology in the vicinity of the culvert. For example, Barnard et al. (2013) showed that streambeds inside stream simulation culverts respond differently to extreme flood events than natural streambeds outside the culvert.

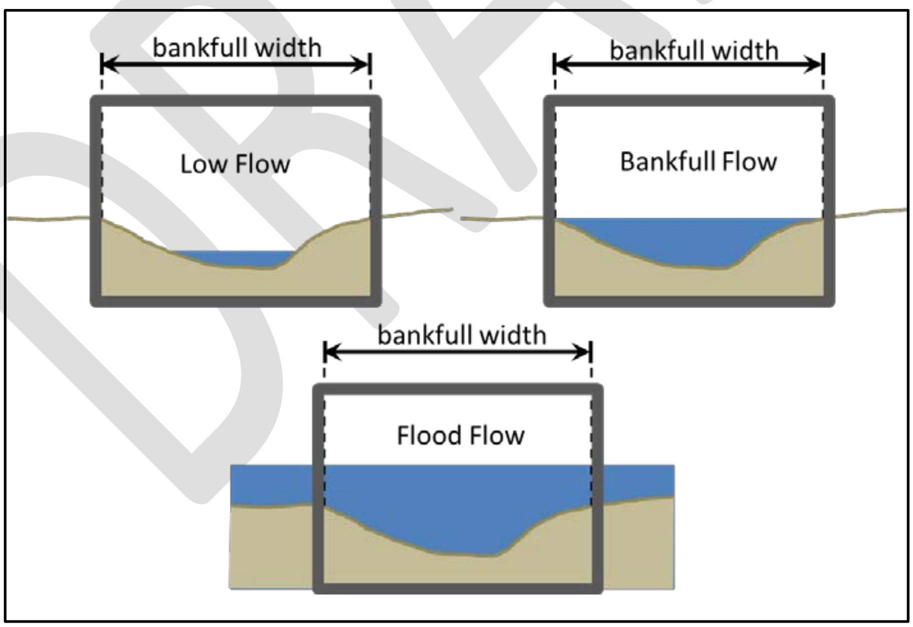


Figure 4. A no-slope box (i.e., rectangular) culvert at different stream flows. Width of the bed inside the culvert is equal to the bankfull width. Hence, the culvert can fully accommodate the bankfull flow. When a stream overtops its banks, the culvert cannot accommodate the flow which changes stream flow velocities immediately upstream and downstream of the culvert.

Based on theoretical grounds, culverts are expected to cause changes in stream channel morphology. However, we cannot predict *a priori* the exact nature or degree of those changes. Effectiveness monitoring will quantify the changes to channel morphology caused by culverts or to beach morphology caused by bulkheads. Effectiveness monitoring may detect statistically significant changes. If those changes are thought to be biologically significant, i.e., to result in a net loss of fish habitats, then validation monitoring will be conducted to determine whether those changes result in a net loss of fish or shellfish habitats.

Validation Monitoring

Limited fiscal resources limit the number of hydraulic projects that can be monitored. Hence, validation monitoring may address only marine shoreline bank protection at new armoring structures only. See the section on monitoring design for marine shoreline armoring for a description.

Monitoring Design for Culverts

Monitoring design addresses objectives, the sampling frame, controls, replication, stratification, site selection, and sample size. Implementation and effectiveness monitoring of culverts are essentially descriptive (i.e., observational) studies. That is, the main purpose of implementation monitoring is to describe over time failure (or success) rates for implementation of hydraulic projects, and the main purpose of effectiveness monitoring is to describe over time changes in fish habitats. Implementation and effectiveness monitoring of culverts are not experimental or manipulative studies, and therefore, cannot establish cause-and-effect relationships. Descriptive studies can, however, estimate correlational relationships.

Challenges

Monitoring design for hydraulic project implementation and effectiveness monitoring is challenging because the permitting and construction processes are completely independent of monitoring. That is, compliance and effectiveness monitoring have no control over the time, place, design, or conditions of hydraulic projects. Monitoring of culverts faces several challenges: 1) timing and duration of field season, 2) notification of culvert completion, and 4) obtaining access to sites on private lands,

Timing and Duration of Field Season. The exact timing and duration of the field season for culvert monitoring are unknown. The field season begins when we learn of that year's first completed culvert, sometime in July, and ends when high stream flows arrive, typically sometime between mid-October to early November. The uncertain nature of the field season creates challenges for meeting minimum sample size requirements, scheduling, and staffing.

In order to protect fish life, HPA permits include provisions that restrict hydraulic project construction to certain time periods. Permittees must conduct in-channel construction activities during the construction window specified on their permit. Timing restrictions vary by county and by individual streams (Table 3). For some streams the construction window may be only one month. In western Washington, the construction window generally runs from July 1 to about September 30. Therefore, we expect the first culverts to be completed about mid-July.

The climate of western Washington exhibits significant increases in rainfall around late October or early November (Figure 5) with a consequent rise in stream flows. Extreme rainfall events (10 to 50-year floods) could alter the structure of recently completed culverts. Countersinking depth and culvert width at streambed are particularly sensitive to high flows. For this reason, implementation monitoring after extreme rainfall events is invalid and will be avoided. Hence, the field season for implementation monitoring will end after the first heavy rains in October or November. Because extreme events are unpredictable, implementation monitoring of a hydraulic projects should occur as soon as practicable after WDFW receives notification of project completion. Culvert monitoring during high stream flows is also difficult and potentially dangerous, which are other good reasons to terminate the field season after the first heavy autumn rains.

Notification of Culvert Completion. Scheduling site visits for implementation and effectiveness monitoring is challenging primarily because scheduling depends on communication and coordination with people who are not invested in the success of the HPA program, namely, HPA permittees.

Until very recently, there was no automatic permit provision to notify WDFW when culvert construction was completed. Even when a notification provision is included in an HPA permit, permittees do not always comply with it. Hence, at present, we must repeatedly contact permittees to learn when their culvert project is done. This is a very time consuming and frustrating task. In November 2013 a standard notification was added to APPS . This should help immensely, but there is still a two to four year backlog of permits that do not contain the new notification provision.

Lack of timely notification of culvert completion may hinder our ability to meet minimum sample size objectives. Our field season may be shorter than expected. It can begin later than expected because we may learn of the year's first completed culvert later than we should, and it may end earlier than expected because we will not learn of some completed culverts until after high stream flows have arrived. Lack of notification of culvert completion may also affect the quality of our random sample. Implementation failure rates for the population of culverts for which we do not receive notification may be different than implementation failure rates for the overall population of culverts

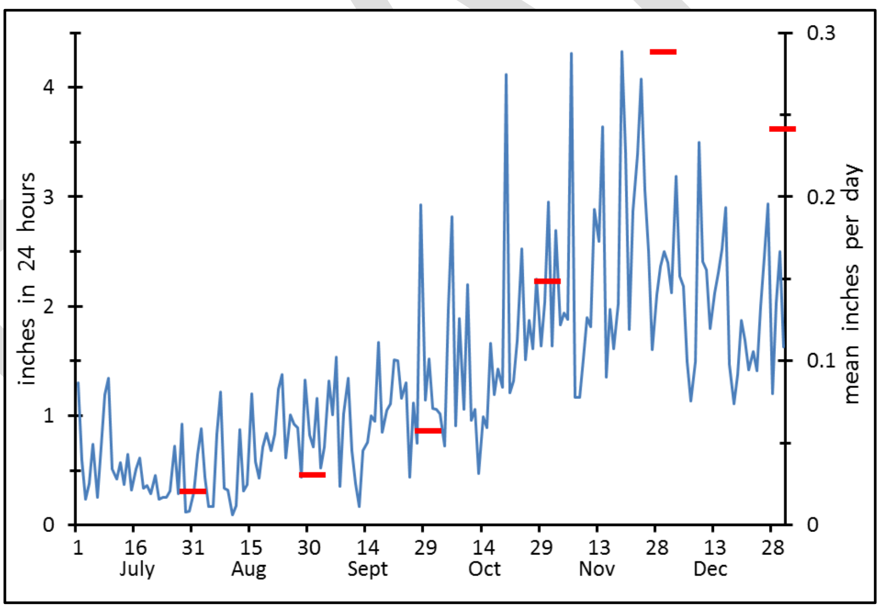


Figure 5. Long-term precipitation statistics for July 1 to December 31 for data recorded at Olympia Regional Airport. Red bars indicate mean daily precipitation for each month. Blue line shows 24-hour maximum rainfall records for each day. The tall peaks around September 29 were recorded in 2013. Some culverts visited for implementation and effectiveness monitoring after these record rains exhibited changes that were obviously attributable to high stream flows. Consequently, the culvert and stream channel conditions immediately after culvert installation are unknown and unknowable.

Table 3. Salmonid spawning and incubation times used to guide timing restrictions incorporated into HPA permit provisions in Washington. Within each county, particular streams have their own specific estimates of salmonid spawning and incubation times.

County	Period when Spawning or Incubation are Least Likely to be Present	
	Begin	**End**
western Washington		
Island	June 16	Oct 15
San Juan	July 1	August 31
Whatcom	July 16	August 15
Pierce	July 16	August 31
Clallam	July 16	Sept 15
Snohomish	July 16	Sept 15
Thurston	July 16	Sept 15
Wahkiakum	July 16	Sept 15
Clark	July 16	Sept 30
Cowlitz	July 16	Sept 30
King	July 16	Sept 30
Grays Harbor	July 16	Oct 15
Kitsap	July 16	Oct 15
Jefferson	July 16	Oct 31
Skagit	August 1	Sept 15
Lewis	August 1	Sept 30
Pacific	August 1	Sept 30
Mason	August 1	Oct 15
eastern Washington		
Benton	June 1	Sept 30
Franklin	June 1	Sept 30
Spokane	June 16	August 31
Lincoln	June 16	Feb 28
Okanogan	July 1	August 15
Ferry	July 1	August 31
Pend Oreille	July 1	August 31
Douglas	July 1	Sept 30
Kittitas	July 1	Sept 30
Adams	July 1	Oct 31
Grant	July 1	Oct 31
Chelan	July 16	August 15
Stevens	July 16	August 31
Asotin	July 16	Sept 15
Skamania	July 16	Sept 15
Columbia	July 16	Sept 30
Garfield	July 16	Sept 30
Klickitat	July 16	Sept 30
Whitman	July 16	Dec 15

Obtaining Access to Sites on Private Lands. To obtain an unbiased sample of culverts we must have access to all randomly selected culverts. Implementation monitoring needs access to culverts, and to measure bankfull width, their adjacent stream channel. Effectiveness monitoring needs access to culverts *and* to up and downstream reaches of the stream channel. Culverts under public roads are readily accessible, but accessing culverts under private roads requires landowner permission. Accessing up and downstream reaches of the stream channel almost always requires landowner permission.

HPA applicants must complete a Washington State Joint Aquatic Resources Permit Application (JARPA) Form. At the end of that form the applicant agrees to the following:

> "I consent to the permitting agencies entering the property where the project is located to inspect the project site or any work. These inspections shall occur at reasonable times and, if practical, with prior notice to the landowner."

For the purposes of monitoring and based on the JARPA consent above, we have developed a protocol for notifying landowners and requesting permission to obtain access to culverts on private land (Appendix D). This landowner/permittee contact protocol is valid only while the JARPA is valid. Notwithstanding an HPA applicant's signature on the JARPA form, our protocol respects a landowner's decision to deny access to their private land. The landowner/permittee contact protocol is very similar to the current draft of WDFW's Policy 5212-*Monitoring Compliance with the State Hydraulic Code*.

To obtain access to up and downstream reaches of the stream channel located on private lands. We will attempt to identify the residences associated with those private lands, visit that residence, and attempt to make in-person contact with the landowner to request permission.

WDFW's Policy 5212 does not apply to implementation and effectiveness monitoring. Policy 5212 says, "This policy applies to employees who enforce and monitor compliance with the State Hydraulic Code." Habitat Program staff engaged in implementation and effectiveness monitoring are not employees who enforce the hydraulic code. Therefore, Policy 5212 does not apply to them.

Implementation Monitoring

Implementation monitoring monitors the performance of people within a regulatory process. As such, it does not have to contend with the extreme natural variation and confounding external factors of effectiveness monitoring. Consequently the design of implementation monitoring is relatively straightforward.

Objectives. The main objective is to detect failures to properly implement the hydraulic code rules, WDFW's design guidelines (Bates et al. 2003), or HPA permits. Reliably detecting failures and accurately quantifying trends in hydraulic project failure rates is a prerequisite for improvement of the HPA program.

Population / Sampling Frame. The population or sampling frame is the collection of objects from which a sample will be drawn. It defines the scope of statistically valid inference. Our idealized sampling frame is all culverts, new or replacement, on fish bearing streams that have an HPA issued by WDFW and were constructed immediately prior or during our field season. At present, we will limit our sampling frame to culverts in WDFW's Regions 4, 5, and 6 (i.e., western Washington). As more resources become available we will expand our sampling frame to eastern Washington.

The idealized sampling frame is compromised by three potential problems: 1) no notification of project completion for some culverts, 2) denial of access to private lands for some culverts, and 3) a premature end to the field season due to weather. These three problems remove culverts from our population of recently completed culverts, and therefore, our sample of culverts may not be representative of the entire population of recently completed culverts.

Unlike effectiveness monitoring, hydraulic projects are visited only once for the purposes of implementation monitoring. In other words, implementation monitoring is a one-time assessment of each randomly selected culvert.

Controls. Implementation monitoring is mainly descriptive. Nevertheless, we are testing a hypothesis – success rates for hydraulic project implementation are improving. The null hypothesis (H_0) is no change in success (or failure) rate between year 1 and subsequent years. Therefore, in effect, our control or reference is the first year of monitoring. Because implementation monitoring will occur over many years, the null hypothesis could be restated as no change in the slope of line describing success (or failure) rates through time.

There is no pre-construction site visit for implementation monitoring of culverts, and therefore, "before" conditions will not serve as a control.

Replication. Replication is the repetition of an experimental treatment in the same or similar conditions. Implementation monitoring is mainly descriptive. It is not testing the effects of different treatments, therefore no replication is necessary. The Habitat Program's Science Division has no control over hydraulic project design, permitting, or construction processes, and consequently, replication is not even possible.

Stratification. Stratification divides a population in strata (or subpopulations) that are thought to be more homogeneous. Populations are stratified to: 1) improve the accuracy of estimates by ensuring that a random sample is representative of the entire population, 2) enhance the precision of estimates, and 3) obtain parameter estimates for each subpopulation (i.e., each stratum). We will divide culverts into permittor or permittee strata and into temporal strata.

Recall that implementation monitoring focuses on the performance of people engaged in the permitting and construction processes − permittors and permittees. If substantial heterogeneity exists within these groups with respect to implementation of the permitting or construction processes then stratification may be warranted. The permittors, i.e., Habitat Program staff, are organized into six WDFW administrative regions. Regional differences in staff could result in regional differences in the HPA permitting process, and therefore, strata could be WDFW administrative regions. In stratified random sampling, the allocation of sampling effort among strata is usually based on the relative size of each stratum (i.e., proportional allocation). Table 4 provides information for proportional allocation of sampling effort by region.

Permittees could be divided into three types: government (state, county, and city), private forest lands, and other private lands. "Private forest lands" are HPAs associated with a forest practices permit for a private landowner. We expect more heterogeneity among permittees than among permittors, and therefore, we will stratify culverts by permitee types only. Assignment of hydraulic projects to permittee strata is determined through information on the HPA permit and may be done *post hoc*. As we learn more about the culvert permitting and construction processes other forms of stratification will be considered.

Timing restrictions stipulated in HPA permits constrain culvert construction to a temporal window extending from approximately July 1 to October 31. One potential source of bias is the time at which culverts are completed. It seems quite plausible that culverts completed near the end of the construction window would have different rates of compliance than culverts completed at the beginning of that window because near the end of the construction window contractors may be under duress. Hence, we want the random sample spread uniformly across the times at which culverts were completed. Temporal stratification will be 2 week periods from July 1 to October 31. Therefore, we have 24 strata: 3 permittee types x 8 time periods.

Table 4. HPA permits issued per year over 5 years from 2008 to 2012. Based on data from WDFW's HPMS database. Note: the number of hydraulic projects permitted each year does not equal the number of projects completed each year.

	WDFW Regions						Projects
Hydraulic Project Type	**1**	**2**	**3**	**4**	**5**	**6**	**/Year***
Freshwater Culverts	59	14	17	202	128	253	674
Marine Bank Protection†	0	0	0	119	2	178	298
Total	59	14	17	321	130	431	972

* mean over 5 years: 2008 to 2012

† Includes modified, replacement, and new structures. Most permits are for replacement.

Site selection. Site selection, i.e., hydraulic project selection, is complicated, especially for culverts. The field season for culvert monitoring is *approximately* July 16 to October 30. "Approximately" is emphasized because in any given year, we do not know when the first culvert will be completed or when the first heavy rains will commence. Roughly 2000 water crossing structures are permitted each year (Table 3), although some permits are never used. We cannot visit all culverts, and therefore, by necessity, failure rates must be an estimate based on a sample. Accurate estimates and valid inferences about an entire population require a random sample.

We want an unbiased random sample. Stratification by time period addresses a major source of potential bias. Random site selection occurs within strata. Assignment of projects to temporal strata is determined by the reported date of project completion. Projects will be assigned to 2 week strata: July 1 - July 15, July 16 - July 31, August 1 - August 15, etc. Ideally, random selection of sites will occur every two weeks from the temporal stratum assembled during the previous two weeks. In other words, the sites visited between August 1 and August 15 will be those for which we received notification of project completion between July 16 and July 31 and were randomly selected on August 1.

We know that the number of completed culverts is unevenly distributed over time – fewer culverts are completed in July than in September. Considering the staff resources we are likely to have, we expect to be able to visit and measure a large proportion of the culverts completed in July and August but a smaller proportion in September. We would like that the proportion of culverts visited and measured to be constant over time (i.e., proportional allocation of samples across temporal strata). Therefore, if most culverts are completed near the end of the field season, then we should allocate more effort (i.e., more staff) to implementation monitoring in September or October than in July. If proportional allocation of effort is not followed, then our sample may not be representative and the nominal α or β (probabilities of Type I and Type II error) will actually be larger than specified.

Sample Size. Selecting sample size entails subjective judgment and compromise. We must make judgments regarding acceptable levels of Type I and Type II errors, but "acceptable levels" may be compromised by the cost of sampling. That is, we desire a very low probability of error, but our desire may be frustrated by lack of resources for monitoring.

We will test the hypothesis that the trend in success (or failure) rates over multiple years has a zero slope. Successes and failures can be modelled as independent Bernoulli trials which conform to a binomial distribution. The minimum sample size per year, *n*, needed for detecting a trend in independent binomial samples can be estimated with the equations (Nam 1987):

$$n = (n^*/4)\ [\ 1 + \sqrt{1 + 2/(Dn^*)}]^2 \qquad (1)$$

$$n^* = \left[z_{1-\alpha}\sqrt{k(k^2-1)\bar{\pi}(1-\bar{\pi})/12} \; + \; z_{1-\beta}\sqrt{\sum c_i^2\, \pi_i(1-\pi_i)} \right]^2 / \, D^2 \qquad (2)$$

$$\bar{\pi} = \frac{(\pi_0 + \pi_1 + \ldots + \pi_i)}{k} \qquad (3)$$

$$D = \sum c_i\, \pi_i \qquad (4)$$

$$c_i = i - 0.5(k-1) \qquad (5)$$

where k is the number of independent times at which measurements are taken, π_0, π_1, and π_i are failure rates (expressed as proportions) at times 0, 1, and i; α is the probability of Type I error; β is the probability of Type II error; z represents the standard normal distribution.

To determine sample size we need values for k, α, β, and $\pi_0, \pi_1, \ldots, \pi_i$.[6] Effect size is represented by and $\pi_i - \pi_0$. Relationships of sample size to k, α, β, and effect size are shown in Figure 6. The null hypothesis (H_0) is zero slope in the trend of success (or failure) rates over year 0 to year i. We assume that we want to see "results" in 3 to 4 years. We are more concerned with avoiding Type I error than avoiding Type II error. That is, we do not want to infer a change in success rate when in reality there was no change in success rate. Therefore, we choose $\alpha = 0.05$. According to Table 1, failure rates for culverts were about 0.30 to 0.35 in 2010. We want (and expect) the effect size to be negative, i.e. we expect to see improvement. To detect an effect size of -0.1, sample sizes range from 234 to 442 per year, depending on β and k (Table 5), which would require approximately 15 to 28 separate site visits per week during the field season. To detect an effect size of -0.2, sample sizes are greatly reduced: 56 to 103 samples per year, which would require approximately 3.5 to 6.5 site visits per week, on average. Fifteen site visits per week for the entire field season is beyond our current capacity. One the other hand, an effect size of 0.2 corresponds to a change in failure rates from 0.3 to 0.1, which seems rather unlikely. Therefore, we need the ability to detect smaller changes on the order of 0.15. To detect an effect size of -0.15, sample sizes range from 102 to 190, depending on β and k, which would require approximately 6.5 to 95 separate site visits per week. A β of 0.3 means a 30% chance of not detecting a change in success rate when a change actually occurred. That error rate seems too high. A β of 0.2 is more reasonable, and therefore, our minimum sample size is between 128 and 144 sites per year, which corresponds to about 8.5 site visits per week, on average.

The minimum sample size of 136 culverts per year does not take into account the realities of the HPA implementation monitoring. The actual number of culverts that can be monitored each year is limited by the number of culverts that are actually available and the number of field staff available to do the monitoring. We may not meet our minimum sample size because the number of available culverts in a given year is too small or because the number of field staff available for monitoring is too small. Problems with field season duration, notification of project completion, and access to private lands may reduce available culverts to a number less much less than 136.

In 2013, 48 culverts were visited for implementation monitoring. This was accomplished with 1 field crew leader and 1 field crew all working half-time (i.e., 20 hours/week) on culvert monitoring. If sample size is only 48 sites per year, then we can detect a change in failure rate a little less than -0.25 over 3 to 4

---

[6] Sample size calculations need an estimate of variance. In equation 2 variance estimates are incorporated by $\pi_x(1-\pi_x)$. The sample size determination depends on initial failure rate, π_0, because effect sizes for proportions are symmetrical about 0.5.

years. In other words, in order to detect a change that is statistically significant, HPA implementation failure rates must improve from the current failure rate of about 0.30 to 0.05.

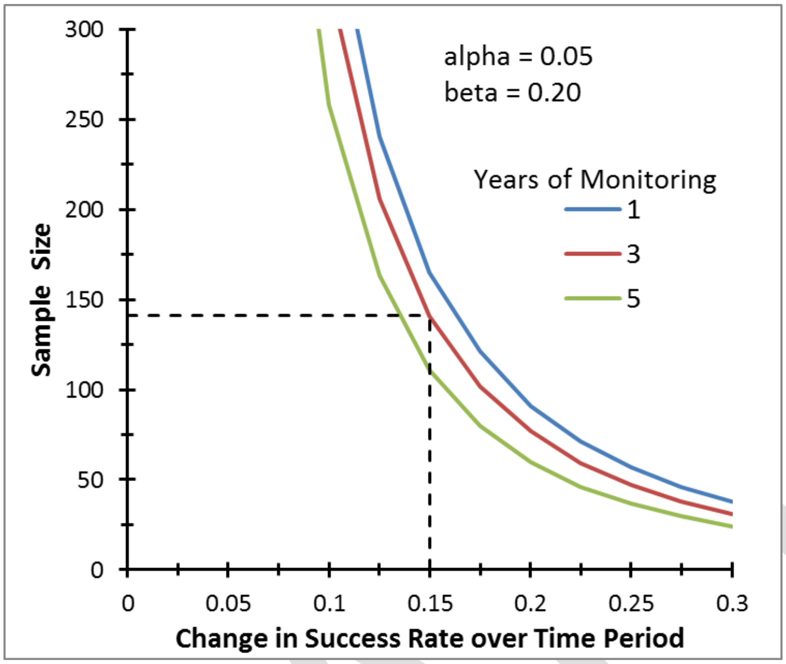


Figure 6. Examples of relationships used to determination of sample sizes in Table 5. Initial failure rate π_0 = 0.65 was used to determine sample size. Years of monitoring equals *k*-1.

Table 5. Sample Size for different values of β, the probability of Type II error, and effect size. For these values $\alpha = 0.05$ and $\pi_0 = 0.65$. The proposed minium sample size is between the values highlighted in gray.

Time Period (*k-1*)	Effect Size	**β**		
		0.3	**0.2**	**0.1**
3 years	0.25	39	48	63
	0.2	63	78	103
	0.15	114	144	190
	0.1	263	332	442
4 years	0.25	34	43	56
	0.2	56	80	92
	0.15	102	128	170
	0.1	234	297	395

Data Analysis. The primary hypothesis we are testing is that implementation of HPA permits for culverts is improving over time. Hence, we are attempting to detect a trend in the rate at which HPA permits and certain provisions in the permits are implemented correctly by both permittor and permitee. The following analyses are preliminary ideas which may change after new data provide information and insights on the best ways to assess implementation of HPA permits. Data collection procedures are in Appendix A.

1. Rates of overall Implementation Failure per Year: Data will show the estimated rates of implementation failure (i.e., percent of all culverts that fail) over the years of monitoring. Implementation failure rates will be calculated for both permittors and permittess. Failure rates can also be calculated for different elements of the hydraulic project: culvert width, culvert slope, etc. Trends will be estimated, but statistically significant trends will be determined through tests of null hypotheses. We assume that failure rates in different years are statistically independent. Because failure rates may either increase or decrease the tests are two-sided. The following is a subset of the statistics and hypothesis tests we can perform with the data to be collected:

   Test 1a: $F_t = 0$, where
   F_t is the estimated trend of implementation failure rates from year 0 to year t;

   Test 1b: $F_{tPermittor} = 0$ where
   $F_{tPermittor}$ is the estimated trend of permittor failure rates from year 0 to year t;

   Test 1c: $F_{1Permittee} = 0$, where
   $F_{tpermittee}$ is the estimated trend of permittee failure rates from year 0 to year t;

   Test 1d: $F_{tPermitWidth} = 0$, where
   $F_{tPermitWidth}$ is the estimated trend from year 1 to year t at which permits fail to have the correct culvert width;

   Test 1e: $F_{tCulvertWidth} = 0$, where
   $F_{tCulvertWidth}$ is the estimated trend from year 1 to year t at which culverts fail to be constructed with the same width specified on the permit.

2. Number of Permit Errors: Failure or success rates are based on binary assessments of the permitting and construction processes. A hydraulic project fails to be implemented properly with any number of defects greater than one. We can also report the mean number of defects per hydraulic project and test for statistically significant changes over time in the mean number of defects. For example:

   Test 2a: $D_1 = D_t$, where
   D_1 and D_t are the estimated mean number of defects per hydraulic project in year 1 and in subsequent years, respectively

   Test 2b: $D_{1Permittor} = D_{tPermittor}$, where
   $D_{1Permittor}$ and $D_{tPermittor}$ are the estimated mean number of defects per hydraulic project committed by permittors in year 1 and in subsequent years, respectively

   Test 2c: $D_{1Permittee} = D_{tPermittee}$, where
   $D_{1permittee}$ and $D_{tPermittee}$ are the estimated mean number of defects per hydraulic project committed by permittees in year 1 and in subsequent years, respectively

3. Magnitude of Permit Errors: For certain defects in hydraulic project permitting or construction we can report the magnitude of the defect. For instance, for hydraulic projects that fail because the culvert width is narrower than the width specified on the permit, we can estimate the mean difference between the permitted culvert width and the constructed culvert width. For example:

Test 3a: $M_{1CulvertWidth} = M_{tCulvertWidth}$, where

$M_{1WidthPermit}$ and $M_{tWidthCulvert}$ are the estimated mean differences between permitted culvert width and actual culvert width per hydraulic project in year 1 and in subsequent years, respectively

Test 3b: $M_{1SlopePermit} = M_{tSlopeCulvert}$, where

$M_{1SlopePermit}$ and $M_{tSlopeCulvert}$ are the estimated mean differences between permitted culvert slope and actual culvert slope per hydraulic project in year 1 and in subsequent years, respectively

Effectiveness Monitoring

The design of effectiveness monitoring must be mindful of the dynamic and unpredictable nature of stream channel morphology. We expect considerable variation in the measured changes of stream channel morphology over time. Some portion of this variability will be caused by the natural movement of water, sediment, and wood, and some variability will be caused by variation in culvert designs or dimensions, the remaining variability we can attribute to chance events. The product of different site conditions and different culvert dimensions could result in a high level of variation in stream channel responses. This variation could be reduced by stratifying sites according to hydrological and geomorphological variables and replicating culvert designs within strata. We could stratify post-hoc, but a small sample size may preclude that statistical tactic. We have no control over culvert design, and hence, replication within strata is impossible at this time. While this is not particularly problematic for implementation monitoring, it will diminish the strength of inference (i.e., robustness) of the results obtained through effectiveness monitoring.

The design for effectiveness monitoring is less developed than implementation monitoring because of the many unknowns about the responses of stream channels to current culvert designs, such as stream simulation, the degree of variation in those responses, and the best tools and techniques for accurately characterizing those responses.

Objectives. Reliably detecting failures and accurately quantifying trends in hydraulic project failure rates is a prerequisite for improvement of the HPA program. The main objectives are to detect changes in the culvert structure and detect changes in channel morphology near the culvert. Net loss of habitats caused by a properly implemented hydraulic structure is defined as an effectiveness failure, however, at present, we have no way of determining net loss of habitats. Therefore, inferences regarding the effectiveness of current culvert design standards and design guidelines will rely on detecting changes in stream channel morphology. To be clear, effectiveness monitoring is not intended to make strong inferences to the mechanisms for specific project failures (but correlative inferences will be possible). It is intended to describe central tendencies and temporal trends of failures.

Population / Sampling Frame. Culverts found to be properly implemented through implementation monitoring will form the culvert population for effectiveness monitoring. In other words, our sampling frame is all culverts, new or replacement, that have an HPA issued by WDFW, were constructed immediately prior or during our field season, and were properly implemented by the permittor and permittee. At present, we will limit our sampling frame to culverts in WDFW's Regions 4, 5, and 6 (i.e., western Washington). As more resources become available we will expand our sampling frame to eastern Washington.

For effectiveness monitoring, culverts will be visited multiple times − at years 1 (immediately post construction), 2, 5, 10, and 20. The data collected will form a time series of changes in channel morphology.

Controls. Effectiveness monitoring for changes stream channel morphology (i.e., monitoring for habitat failure) consists of two sets of upstream and downstream cross-channel (Figure B1). Transects farthest from the culvert act as controls. The degree of change in the upstream/downstream control transects will be compared to the degree of change in the other upstream/downstream transects.

Cross-channel transects will be measured post-construction only. Using only post-treatment data can provide robust inferences, given sufficient planning and thought, by selecting attributes that identify failures rather than comparing changes at a project site with those of other locations. For example, development of an upstream mid-channel bar, development of a downstream scour pool, and development of a hydraulic jump at the inlet or outlet after one or more high water flow events are indicative of failure of the culvert to pass water, sediment, and likely fish, and thus failure of the culvert, regardless of conditions in other nearby stream reaches. The assumption that underlies this approach is that most conditions at the project site are expected to not substantially change through time.

While before-after sampling is often desired to detect changes, each sample requires resources and for culvert monitoring will likely provide little additional information because differences in the project site before and after culvert construction. Reference sites are also often useful, especially in retrospective comparisons where pre-project data are not available, but they are discouraged for much of this work because of the limited usefulness of most comparisons (e.g., we expect stream reference and treatment sites will differ and change at different rates due to their unique conditions), but they might be necessary for monitoring some effects of some project types. Further, temporal delays in the effects of the projects on habitat are expected, because changes are expected to be most likely and large as a result of high water flows. Sampling post-construction only will allow for sampling additional sites.

Replication. Effectiveness monitoring is not testing the effects of different treatments, therefore no replication is necessary. The Habitat Program's Science Division has no control over hydraulic project design, permitting, or construction processes, and consequently, replication is not possible.

Stratification. If sites are stratified, it will be done *post-hoc*. Stratification could be based on, for example, stream size (e.g., stream order or basin area), predominant stream bed substrate, or percent impervious surface covering the upstream drainage basin.

Site Selection. The culverts selected for effectiveness monitoring are a randomly selected subset of the culverts selected for implementation monitoring. Culverts used for effectiveness monitoring must pass all aspects of implementation monitoring. If a culvert has been randomly selected for effectiveness monitoring, then immediately after collecting implementation monitoring data, the culvert will be evaluated for implementation failure. If the culvert fails, then it will be dropped from the subsample of culverts selected for effectiveness monitoring and a new randomly selected culvert will be added to the subsample.

If multiple culverts are covered under an HPA permit, then only one of the culverts should be sampled to avoid potential sources of bias (e.g., spatial autocorrelation and pseudoreplication). Ideally, selected sites should be located in different, preferably unconnected watersheds. If that is not possible, then sites should be located at least 5 stream km apart and on different stream reaches, as defined by confluences.

Sample Size. Reliably detecting failures requires sample sizes that are sufficient to reduce the probability of Type II errors (i.e., failing to detect a difference when one is present) during each reporting interval (e.g., annually). Because staff resources are limited and sampling many locations is desired, survey crews will visit each culvert site only once in years 1, 2, 5, and 10.

Given adequate resources our optimal minimum sample size would be 50 sites per year. A minimum sample size of 50 is a rule-of-thumb using a binomial test (success or fail, yes or no). Power is basically optimized at about that sample size. However, this optimal size is impractical because all sites will be revisited in years 2, 5, and 10, and consequently, some years would require visits to 100 separate sites. Not all sites need be resampled in every year, but 50 is generally the magic number if we want to detect change. If we want to extrapolate to a larger population regarding rates, etc., then more than 50 sites may be needed.

In 2013, 22 culverts were visited for effectiveness monitoring (however 5 of those culverts were later determined to have failed implementation). This was accomplished with 1 field crew leader and 1 field crew all working half-time (i.e., 20 hours/week) on culvert monitoring. If sample size is only 22 sites per year, then our capability to detect change is dramatically reduces from our optimal sample size of 50. We will do more detailed sample size estimates for effectiveness monitoring after we can estimate the variance of key variables and demonstrate how many sites can be visited in a single field season.

Interspersion in Space and Time. Interspersion of sites avoids pseudoreplication. Random selection of HPA permits from all three west-side WDFW administrative regions should ensure spatial interspersion. Spreading site selection over multiple years is a form of temporal interspersion (Table 6). In years 2, 5, and 10, we will attempt to visit each culvert at approximately the same time of year when the culvert was visited in year 1, i.e., between approximately July 1 and mid-October.

Data Analysis. The primary hypotheses we are testing for effectiveness are: 1) culverts that meet design standards in WAC 220-110 or design criteria in WDFW's guidelines (e.g., Bates et al. 2003) immediately post-construction continue to meet those standards/criteria of over time (i.e., not structure failure), and 2) stream channel morphology in the vicinity of the culvert does not change over time (i.e., no habitat failure). For the first hypothesis we are attempting to detect changes in the culvert's original dimensions over time. Measurements for testing the first hypothesis repeat most of the measurements done for implementation monitoring (Appendix A). Data analysis for the first hypothesis will report the proportion of structure failures per year over time.

For the second hypothesis we are attempting to detect changes in stream channel morphology and stream bed substrates. The following analyses are preliminary ideas which may change after new data provides information and insights regarding the effects of culverts on streams. Data collection procedures are in Appendix B.

1. Backwater: Stream crossing structures that impede the flow of water can result in increased depth of water upstream of the structure during high stream flow events and associated high flow velocities can result in failure of the structure to pass fish or persist. Elevation of water directly upstream of the structure should never meet or exceed the elevation of the top of the flow structure (e.g., inlet to culvert). Unfortunately, surveyors cannot be present to witness such events and placing measuring devices is likely too expensive, so indicators of highest water level must be identified and measured.

   Test 1: $E_{bw} \geq E_{inlet}$, where
   E_{bw} is the measured elevation of the backwater, and
   E_{inlet} is the elevation of the top of the inlet.

2. Mid-channel Bar: Stream crossing structures that impede water flow often result in the formation of mid-channel bars or small islands, especially upstream of structures. Structures that constrain high flows can reduce upstream velocity and stream power, resulting in sediment deposition. Presence of mid-channel bars, increase of bar size, and reduced stream gradient indicate likely current or

imminent failure of the structure to pass fish and sediment. Herein, a mid-channel bar is defined as a streambed deposit of alluvial sediment with a maximum elevation that is 1) less than the bankfull channel elevation, 2) greater than the minimum observed elevation of the wetted channel, and 3) with the distance of the maximum elevation greater than one-quarter of the wetted channel width from both the right and left wetted edge of the channel.

Test 2a: $\overline{E}_{bar} \geq \overline{E}_{we}$, where

$\overline{E}_{bar}$ is the mean elevation of the mid-channel bar, where the bar is defined as the centermost three elevation measurements based on the distance from the right to the left wetted edge or bankfull edge if the stream is dry, and

$\overline{E}_{we}$ is the mean measured elevations of the wetted edges.

Test 2b: $\overline{E}_{bar} \geq 1.5\overline{E}_{wc}$, where

$\overline{E}_{bar}$ is the mean elevation of the mid-channel bar, where the bar is defined as the three highest, contiguous elevation measurements in the mid-half of the wetted channel or bankfull channel if the stream is dry, and 1.5 requires the bar height to be ≥ 50% higher than the remainder of the wetted channel, and

$\overline{E}_{wc}$ is the mean measured elevations of the wetted channel as measured from the right to left wetted edge, not including the elevations used to measure the bar.

3. Mid-channel Trough: Stream crossing structures that impede the flow of water can result in high flow velocity and stream power at the structure outlet that can result in failure of the structure to pass fish or persist. When the flow of sediment is also constrained by the structure a mid-channel trough is often formed, detrimentally effecting downstream habitat and likely fish passage.

Test 3: $\overline{E}_{trough} \leq 1.5\ \overline{E}_{we}$, where

$\overline{E}_{trough}$ is the mean elevation of the mid-channel trough, where the trough is defined as the three lowest, contiguous elevation measurements in the mid-half of the wetted channel or bankfull channel if the stream is dry, and 1.5 requires the trough depth to be ≥ 50% higher than the remainder of the wetted channel, and

$\overline{E}_{we}$ is the mean measured elevations of the wetted channel as measure from the right to left wetted edge, not including the elevations used to measure the trough.

4. Aggradation and Degradation: Although stream cross section and longitudinal profiles are expected to change due to natural stream dynamics, the relative cross section area upstream and downstream of the structure should remain relatively stable as water and sediment flow through the system. Stream crossing structures that fail to allow sufficient flow of water and sediment can result in aggradation of sediment upstream and degradation (incision) of the channel downstream of the structure that can develop into blockages and outlet drops, respectively.

Test 4a: $\overline{G}_{upstream,\ t1} < 1.2(\overline{G}_{upstream,\ t1+i})$, where

$\overline{G}_{upstream,\ t1}$ is the mean of all upstream cross section Gini coefficients at year one,

$\overline{G}_{upstream,\ t1+i}$ is the mean of all upstream cross section Gini coefficients at a subsequent year, and 1.2 is used to ensure a change of at least 20% of the mean Gini coefficient.

Test 4b: $T_{upstream,\ t1} > 1.2(T_{upstream,\ t1+i})$, where

$T_{upstream,\ t1}$ is the thalweg gradient at year one,

$T_{upstream,\ t1+i}$ is the mean thalweg gradient in a subsequent year, and 1.2 is used to ensure a decrease of at least 20% of the original gradient.

Test 4c: $\overline{G}_{downstream,\ t1} > 1.2(\overline{G}_{downstream,\ t1+i})$, where

$\overline{G}_{downstream,\ t1}$ is the mean of all downstream cross section Gini coefficients at year one,
$\overline{G}_{downstream,\ t1+i}$ is the mean of all upstream cross section Gini coefficients at a subsequent year, and 1.2 is used to ensure a change of at least 20% of the mean Gini coefficient.

Test 4d: $T_{downstream,\ t1} > 1.2(T_{downstream,\ t1+i)},$ where

$T_{downstream,\ t1}$ is the thalweg gradient at year one,
$T_{downstream,\ t1+i}$ is the mean thalweg gradient in a subsequent year, and 1.2 is used to ensure a decrease of at least 20% of the original gradient.

5. Pool Frequency: Pools are accepted as important, often limited habitat for many salmonids and aggradation (i.e., filling) of streams due to non-equilibrium conditions of the flow of sediment can reduce the frequency of presence of pools and their depth where they are present. For this work we will simply track the frequency of pools and their mean depth through time.

Test5a: $P_{upstream,\ t1} > P_{upstream,\ t1+i}$,

$P_{upstream,\ t1}$ is the of all upstream pools during year one, and
$P_{upstream,\ t1+i}$ is the number of all upstream pools during a subsequent year.

Test 5b: $P_{downstream,\ t1} > P_{downstream,\ t1+i}$),

$P_{downstream,\ t1}$ is the of all downstream pools during year one, and
$P_{downstream,\ t1+i}$ is the number of all downstream pools during a subsequent year.

Test 5c: $\overline{P}_{upstream,\ t1} > \overline{P}_{upstream,\ t1+i}$,

$\overline{P}_{upstream,\ t1}$ is the mean pool depth upstream of the project during year one, and
$\overline{P}_{upstream,\ t1+i}$ is mean pool depth upstream of the project during subsequent years.

Test 5d: $\overline{P}_{downstream,\ t1} > \overline{P}_{downstream,\ t1+i}$, where

$\overline{P}_{downstream,\ t1}$ is the mean pool depth downstream of the project during year one, and $\overline{P}_{downstream,\ t1+\ i}$ is the mean pool depth downstream of the project during subsequent years.

Pools are defined as having 1) depth greater than the mean depth of the reach, 2) little or no discernible surface roughness, 3) lower velocity than the mean reach velocity, and 4) length of at least one wetted stream width. Pools are counted and measured between transects A through J. Pool depth is calculated as the difference in elevation of the deepest part of the pool and the highest part of the pool tail. Depth of water in each pool is measured at the deepest point.

6. Large Wood Debris: The abundance of LWD can have important implication for site stability, suitability of fish habitat and the continued function of many stream crossing structures. The abundance of LWD is expected to change at sites, but we expect the changes to be relatively similar up- and down-stream of the site, if initial values are similar and if the project can pass LWD.

Test 6a: $LWD_{upstream,\ t1} > LWD_{upstream,\ t1+i}$, where

$LWD_{upstream,\ t1}$ is the count of LWD between transects F and J at time 1, and
$LWD_{upstream,\ t1+i}$ is the count of LWD between transects F and J during subsequent years.

Test 6b: $LWD_{downstream,\ t1} > LWD_{downstream,\ t1+i}$, where

$LWD_{downstream,\ t1}$ is the count of LWD between transects A and E at time 1, and

$LWD_{downstream,\ t1+i}$ is the count of LWD between transects A and E during subsequent years.

Logistics of Culvert Monitoring

Logistics is the planning, coordination, and implementation of the details of an operation. Logistics addresses scheduling, communications, staffing, and equipment. In this plan we address only major logistical issues: 1) identifying culverts available for monitoring, 2) collecting permit and plans information, 3) scheduling site visits, and 4) staffing.

Identifying Culverts Available for Monitoring. Identifying culverts for implementation and effectiveness monitoring is challenging primarily because it depends on communication and coordination with people who are not invested in the success of the HPA program, namely, HPA permittees. Until very recently, there was no automatic permit provision to notify WDFW when culvert construction was completed. Even when a notification provision is included in an HPA permit, permittees do not always comply with it. Hence, at present, we must repeatedly contact permittees to learn when their culvert project is done. In November 2013 a standard notification was added to APPS. This should help immensely, but there is still a large two to four year backlog of permits that do not contain the new notification provision

The field season begins in June 1 when we attempt to identify culverts that will be constructed during the coming summer (Figure 7). The field crew leader searches HPMS (and in the near future APPS too) for HPA permits issued for culverts over the past 5 years. At present, we search only for permits issued in WDFW Regions 4, 5 and 6. A list of projects is compiled with permittee contact information. The field crew leader contacts permittees by e-mail or telephone and keeps a record (e.g., phone log) of his/her contacts. If the permittee is contacted and does not intend to install that culvert the coming summer, then that HPA is eliminated from the list of potential sites. If the permittee intends to install the culvert, then an approximate completion date is requested. Some permittees may need to be contacted repeatedly through the entire field season, however, after the first contact is made, other permittees will provide notification of project completion.

Collecting Permit and Plan Information. Recall that for implementation monitoring permittee failure occurs when the constructed hydraulic project fails to meet or exceed the provisions in the HPA or the plans attached to the HPA. Hence, the first step in data collection is locating and interpreting key structural dimensions on the HPA permit and the attached plans. This information is recorded on the data form in Appendix A. Because HPA applications have no standards, guidelines, or “templates” for engineering drawings and applications are not required to summarize of the project’s key structural dimensions, locating and interpreting information on HPA and plans is a tedious and laborious process. Consequently, a substantial amount of office time must be allocated for review of HPA permits and plans. For implementation monitoring of culverts, the recommended ratio of in-office HPA review time to in-field data collection is 2:3.

Scheduling Site Visits. Scheduling site visits depends upon receiving notification of project completion, obtaining landowner permission to enter private property, and coordinating multiple site visits to maximize efficiency.

The minimum sample size for implementation monitoring is 75 sites per year, and the preliminary minimum sample size for effectiveness monitoring is about 50 sites accumulated over 8 years (Table 7). However, due to staff limitations in 2013 we did implementation monitoring on 48 culverts and effectiveness monitoring on 22 culverts. Hence, roughly half of culverts selected for implementation monitoring were also selected for effectiveness monitoring. In future years, if staff resources are similar to 2013, then we expect to visit roughly the same number of sites (48 and 22) for implementation and effectiveness monitoring. Because sites are revisited at years 2, 5, and 10 for effectiveness monitoring, we estimate that about 16 sites per year, on average, must be visited for effectiveness monitoring (Table 7). In 2014, 17 culverts visited for effectiveness monitoring will be old sites and 5 will be new sites.

If travel to, between, and from sites is minimal, then one or two culverts can be visited by a field crew per day (Table 6). At the current staffing level, field staff are working half time (2.5 days/week) on hydraulic project monitoring. Allowing for ½ day per week for in-office work, four culverts can be visited per week. Hence, over a 12 week field season, 48 culverts can be visited for implementation monitoring. To minimize travel time, the field crew leader will cluster site visits in time according to travel distance between sites. When a sufficient number of randomly selected, completed hydraulic projects within a region become available, then those culverts should be scheduled for site visits. The field crew leader will schedule a multi-day trip.

Temporal stratification of sites will require some forethought on scheduling. Random selection of sites will occur every two weeks from the temporal stratum assembled during the previous two weeks. In other words, the sites visited from August 1 to August 15 will be those for which we received notification of project completion between July 16 and July 31. Sites will be randomly selected from the July 16 to July 31 sampling frame. We recognize that this simplistic scheme could result in inefficiencies. For example, sites within a single 2 week period could be widely distributed in space resulting in excessive travel time. A smarter approach may be to take random samples within 2 week periods but defer site visits until a number of sites can be visited more efficiently in single trip.

Staffing. Staffing is challenging because the number of culverts completed per week varies substantially over time; for instance, the number of culverts completed in September is much greater than in July (Figure 7).

The temporal stratification of our monitoring design calls for proportional allocation of sampling effort across the field season. In other words, after a two week period during which relatively more culverts were completed, the following two weeks would have relatively more site visits. This sampling scheme requires uneven levels of staff time throughout the field season. Proportional allocation across temporal strata requires close attention to logistics, in particular, scheduling of site visits and staff resources.

Given current funding, we will visit about 48 culverts per year for implementation monitoring and 16 sites, on average, for effectiveness monitoring. That requires about 4 site visits per week, on average, over a 12 week field, but the actual number of site visits could range from 2 (early in the season) to 8. Hence, our current level of staffing may need to be doubled during the busiest part of the field season.

Table 6. Estimates of time in hours to complete implementation and effectiveness monitoring measurements at one culvert.

	Persons in Field Crew	
Type of Monitoring	**1**	**2**
Implementation	¾	⅔
Effectiveness	4½	3
Implementation & Effectiveness	5¼	3⅔

Table 7. Approximate annual sample size and site visit schedule for effectiveness monitoring. Gray rows − years 1, 2, 5, and 10 – denote years when sites are visited for effectiveness monitoring. Sites may be dropped from monitoring because of catastrophic structural failure, terminated permission to access to private lands, or reduction in funding. Larger number of sites may be added because of increases in funding. At current level of funding, the number of sites visited for implementation monitoring will remain relatively constant at about 48 sites per year. Effectiveness monitoring sites are a random subset of those 48 sites.

Sample Year	**Calendar Year**																					
	1		2		3		4		5		6		7		8		9		10		11	
	2013		**2014**		**2015**		**2016**		**2017**		**2018**		**2019**		**2020**		**2021**		**2022**		**2023**	
	visit	no visit	visit	no visit	visit	no visit	visit	no visit	visit	no visit	visit	no visit	visit	no visit	visit	no visit	visit	no visit	visit	no visit	visit	no visit
1	22		5		10		5		0		10		0		5							
2			17		5		10		5		0		10		0		5					
3						17		5		10		5		0		10		0		5		
4								17		5		10		5		0		10		0		5
5									17		5		10		5		0		10		0	
6												17		5		10		5		0		10
7														17		5		10		5		0
8																17		5		10		5
9																		17		5		10
10																			17		5	
11																						17
sub totals	22	0	22	0	15	17	15	22	22	15	15	32	20	27	10	42	5	47	27	25	5	47
total sites	17*		22		32		37		37		47		47		52		52		52		52	

* Five of 22 culverts were later determined through implementation monitoring to have failed implementation, and therefore, they were removed from further effectiveness monitoring

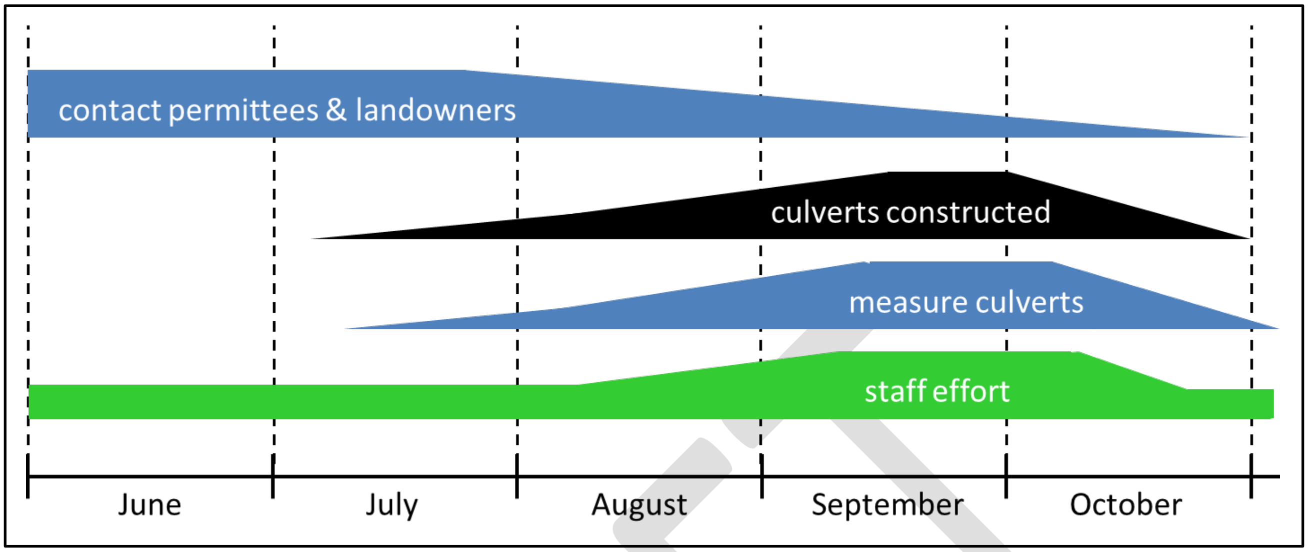


Figure 7. Approximate annual activity time line of field efforts for implementation and effectiveness monitoring of culverts.

Monitoring Design for Marine Shoreline Armoring

Implementation Monitoring

Implementation monitoring measures the results of the hydraulic project permitting and construction processes. Implementation monitoring is essentially descriptive (i.e., observational) studies. That is, the main purpose of implementation monitoring is to describe over time failure (or success) rates for implementation of hydraulic projects. Because there are relatively few structural guidelines for shoreline armoring relative to stream crossing projects, implementation monitoring for shoreline armoring will focus on the location of the structure relative to MHHW or OHW, the length of the structure, the building materials used, and any additional measures stated in the plans or the permit.

Objectives. The main objective of implementation monitoring is to detect failures to properly implement the hydraulic code rules or HPA permits. Reliably detecting failures is a prerequisite for identifying projects suitable for effectiveness monitoring.

Population / Sampling Frame. The sampling frame is recently permitted projects for the construction of new marine shoreline armoring. The sampling frame may include replacement armoring if few new armoring projects are identified, and resources permit additional sampling. Current sampling frame is limited to coastal zones of San Juan and Kitsap Counties. We will expand the sampling frame to include other parts of Puget Sound, Region 4 and Region 6,as more resources become available.

Controls. Implementation monitoring is mainly descriptive. That is, the main purpose of implementation monitoring is to report success and failure rates for implementation of hydraulic projects. Nevertheless, we are testing a hypothesis – success rates for hydraulic project implementation are improving. The null hypothesis (H_0) is no change in success (or failure) rate between year 1 and subsequent years. Therefore, in effect, our control or reference is the first year of monitoring.

Replication. Implementation monitoring is mainly descriptive. It is not testing the effects of different treatments, therefore no replication is necessary. The Habitat Program's Science Division has no control over hydraulic project design, permitting, or construction processes, and consequently, replication is not possible.

Stratification. Stratification divides a population in strata (or subpopulations) that are thought to be more homogenous. Since will measure nearly every HPA for new shoreline armoring projects, stratification will be post hoc. If stratification is deemed appropriate, we may divide projects based on armoring type (soft, hard), permittee, or permittor.

Site Selection. We will aim to survey the whole population of new marine shoreline protection projects approved and completed in the Puget Sound.

Sample Size. We will aim to survey whole population of new marine shoreline protection projects approved and completed in the Puget Sound. Based on conversations with regional habitat biologists that work in the Puget Sound region, we expect up to about 5 new shoreline armoring projects per county to be approved annually (about 60 total). If few new armoring projects are identified, and resources permit additional sampling, implementation monitoring may be expanded to include replacement armoring.

Data Analysis. Analysis of implementation will be a comparison of the plans and conditions approved in the HPA and the observations and measures taken at the project site. Thresholds will be used to determine whether each of the selected elements, such as the location of the structure relative to MHHW or OWH, the structure length, construction material, etc., either pass (within the acceptable bounds of the threshold) or fail (beyond the bounds of the threshold). For example, approved project plans may stipulate that armoring will extend for 150 feet; if we use a threshold of three feet, then if the structure is measured to be 153 feet or greater, then project would fail for this provision, but if it measures 153 feet or less, it would pass. The magnitude of such discrepancies between permitted plans and actual measure will be recorded. We will summarize the measures and observations recorded for each site and the pass/ fail assessment for each measure. We will also summarize pass/fail rates across projects for common measures taken at multiple sites and pass/ fail rates will be compared across years.

Effectiveness Monitoring

Effectiveness monitoring of marine shoreline armoring projects is designed to detect changes in nearshore fish and shellfish habitats that result from the armoring project. Detecting such changes will be difficult to detect because indications of failure may be spread beyond the area of the project site, may take years to manifest themselves, and may be difficult to differentiate from the effects of natural processes or the effects of other structure on the beach. Effectiveness monitoring is intended to describe central tendencies and temporal trends of failures. Therefore, inferences regarding the effectiveness of shoreline armoring practices will rely on detecting changes to nearshore habitat that result from the construction of shoreline armoring.

Objectives. The main objective is to detect changes to fish habitat caused by marine shoreline armoring. Negative change to habitat caused by a properly implemented hydraulic structure is defined as an effectiveness failure.

Population / Sampling Frame. The sampling frame is recently permitted projects for the construction of new marine shoreline armoring. Current sampling frame is limited to coastal zones of San Juan and Kitsap Counties. We will expand the sampling frame to include other parts of Puget Sound, Region 4 and Region 6, as more resources become available.

Controls. Control or reference sites are often useful, especially in retrospective comparisons where pre-project data are not available. Because beaches are dynamic systems that change, we will identify reference sites for comparison to project sites. All references site will be located within the same drift cell as the project site to minimize natural variability. Each reference site will be an unarmored stretch of beach that is surveyed for the same measures during the same time frame as the project site.

Replication. Effectiveness monitoring is not testing the effects of different treatments, therefore no replication is necessary. The Habitat Program's Science Division has no control over hydraulic project design, permitting, or construction processes, and consequently, replication is not possible.

Stratification. Post-hoc stratification will be applied if warranted.

Site Selection. Site selection for effectiveness monitoring will be conditioned upon first being deemed suitable by passing minimum implementation thresholds for critical permit provisions. Additional factors influencing site selection will include our ability to identify a suitable reference beach for the project, and the willingness of the property owner of the project beach as well as the reference beach to allow long term monitoring of their beaches. Additional logistical considerations may also play a role in site selection.

Sample Size. During the first few years of sampling we will attempt to sample every new marine shoreline armoring project with the goal of sampling 50 sites. Resampling of a subset of sites to look for changes or trends, may commence after the first year of sampling. Not all sites will be resampled in every year, and the number of sites sampled from the first year and subsequent years may be reduced to accommodate sampling of new sites.

Data Analysis. Effectiveness monitoring is testing two null hypotheses: 1) no difference in physical characteristics of beach over time, and 2) no difference in physical characteristics of beaches at project site and at undisturbed reference site. Analysis of effectiveness will be a comparison of measures of habitat characteristics of a project site either with an undisturbed reference site, and/or with measures taken before or shortly after construction, and well after construction at the project site. The analysis will look for differences or changes in measures that may be impacted by the project, such as the location of ordinary high water or beach width, amount of accumulated large wood material, and contribution of local marine riparian wood to the beach. We will also compare measures across projects to look for differences or trends in groups of undisturbed/disturbed sites, or before/after measures.

Logistics of Marine Shoreline Armoring Monitoring

The logistics of marine shoreline armoring monitoring will be similar to those of culvert monitoring. Seasonal differences in storm frequency, tide elevation, and occurrence of daytime low tides will dictate the timing of many of the armoring projects as well as the timing of surveys.

Additional logistical challenges are presented by the use of a reference site. While most HPA sites will be accessible by boat without the need to cross neighboring properties, armoring often extends the entire length of a property's water front, so reference sites will likely need to be located on neighboring properties within the drift cell. Since these properties are not associated with the HPA, obtaining permission to survey these properties as a reference site may pose a challenge. Also, some drift cells may be so impacted by armoring that it may not be possible to locate a suitable reference site within the same drift cell. As a result, we expect it may take multiple seasons to achieve the desired sample size.

Data Management

Recording and Storage

At present, all data collected in the office and in the field are recorded on paper data forms. We may develop custom electronic data forms for an I-pad based data logger. We plan to develop a custom database in Microsoft Access for implementation and effectiveness monitoring that is linked to WDFW's Aquatic Protection Permitting System (APPS) and WDFW's GIS database. The transition from the old HPA data system, HPMS, to the new HPA data system, APPS, may require manual transfer of data from HPMS to APPS.

New field data will be transcribed to the database on a weekly basis. All entered data will be double checked for accuracy against the field forms and copies of all data sheets will be archived. Data recording and storage processes include:

- A system for naming files uniquely, with date and project,
- Office and field staff and the monitoring lead will keep a copy of data as back up,
- All electronic data are stored on the WDFW network and backed up daily.

On potential source of data error is measurement unit systems. All HPA permits use the English system (inches, feet), but scientists tend to use the metric system (meters). At present, our data forms use metric, and hence, dimensions on plans attached to HPA permits must be converted from English units to metric units.

Quality Assurance and Quality Control

Data assurance is in large part addressed by consistent execution of standard procedures; ours are documented in Appendices A, B, and C. The experience and training of project staff are critical to achieving data quality goals. To address measurement accuracy and bias, we will ensure that:

- all instruments are inspected, tested, and calibrated according to manufacturer instructions ,
- all equipment are checked for damage before use in the field,
- the same standardized procedures are followed at every site
- separate field crews are issued identical measurement instruments (e.g., GPS, laser level)
- field staff are well-trained by experienced biologists.
- whenever practical, two-person field crews will perform measurements and record data,

When conducting data analysis, we will:

- double check the measurement unit system (i.e., English or metric)
- double check conversion factors between measurement unit systems
- check suspicious outliers for data entry error.

References

Barnard, R.J., J. Johnson, P. Brooks, K.M. Bates, B. Heiner, J.P. Klavas, P.C. Ponder, P.D. Smith, and P.D. Powers. 2013. Water Crossing Design Guidelines. Habitat Program, Washington Department of Fish and Wildlife, Olympia, WA.

Barnard, R.J., S. Yonkers, A. Nagygyor, and T. Quinn. 2013. An evaluation of the stream simulation culvert design in Washington State. unpublished report. Habitat Program, Washington Department of Fish and Wildlife, Olympia, WA.

Bates, K., B. Barnard, B. Heiner, J.P. Klavas, and P.D. Powers. 2003. Design of Road culverts for Fish Passage. Habitat Program, Washington Department of Fish and Wildlife, Olympia, WA.

Blumm, M.C., and J,G. Steadman. Indian treaty fishing rights and habitat protection: the Martinez decision supplies a resounding judicial reaffirmation. Natural Resources Journal 49: 653-706.

Bormann, B.T., P.G. Cunningham, M.H. Brookes, V.W. manning, and M.W. Collopy. 1994. Adaptive ecosystem, management in the Pacific Northwest. PNW-GTR-341. Pacific Northwest Research Station, USDA Forest Service, Portland, OR.

Cramer, M., K. Bates, D. Miller, K. Boyd, L. Fotherby, P. Skidmore, and T. Hoitsma. 2003. Integrated Streambank Protection Guidelines. Washington Departments of Fish & Wildlife, Ecology, and Transportation, Olympia, WA.

Dunn, T., and L.B. Leopold. Water in Environmental Planning. 1978. W.H. Freeman and Company, San Francisco, CA.

Habitat Program. 2011. Hydraulic Project Approval Compliance Monitoring, Progress Report. Unpublished report. Protection Division, Habitat Program, Washington Dept. of Fish and Wildlife, Olympia, WA.

Habitat Program. 2012. Hydraulic Project Approval Compliance Monitoring, Progress Report. Unpublished report. Protection Division, Habitat Program, Washington Dept. of Fish and Wildlife, Olympia, WA.

Johannessen, J., A. MacLennan, A. Blue, J. Waggoner, S. Williams, W. Gerstel, R. J. Barnard, R. Carman, and H. Shipman. 2013. Draft Marine Shoreline Design Guidelines. Washington Department of Fish & Wildlife, Olympia, WA.

Lovaas, J. 2013. 'Culvert case' could be costly issue for state. The News Tribune, Tacoma, WA.

MacDonald, L.H., A.W. Smart, and R.C. Wissmar. 1991. Monitoring Guidelines to Evaluate Effects of Forestry Activities on Streams in the Pacific Northwest and Alaska. EPA/910/9-91-001. USEPA, Region 10, Seattle, Washington. 166pp.

Morisset, M.D., and C.A. Summers. 2009. Clear passage: the Culvert Case decision as a foundation for habitat protection and preservation. Seattle Journal of Environmental Law and Policy 1:29-63.

Mulvihill, C.I., B.P. Baldigo, S.J. Miller, D. DeKoskie, and J. DuBois. 2009. Bankfull discharge and channel characteristics of streams in New York State. USGS Scientific Investigations Report 2009-5144. U.S. Geological Survey, Reston, VA.

Nam, J. 1987. A simple approximation for calculating sample sizes for detecting linear trends in proportions. Biometrics 43:701-705.

Price, D.M., T. Quinn, and R.J. Barnard. 2010. Fish passage effectiveness of recently constructed road crossing culverts in the Puget Sound Region of Washington State. North American Journal of Fisheries Management 30:1110–1125, 2010

Quinn, T. S. Kalinowski, R. Bicknell, C. Olds, M. Schirato, D. Price, C. Byrnes, D. Kloempkin, and R. Barnard. 2006. A Pilot Study of Hydraulic Permit Compliance, Implementation, and Effectiveness in Region 6. Unpublished report. Habitat Program, Washington Dept. of Fish and Wildlife, Olympia, WA.

Rosgen, D.L. 1994. A classification of natural rivers. Catena 22(3):169-199.

USDA and USDI. 1994. Standards and Guidelines for Management of Habitat for Late-Successional and Old-Growth Forest Related Species within the Range of the Spotted Owl. USDA Forest Service, Portland, Oregon.

Sokal, R.R., and F.J. Rohlf. 1995. Biometry. W.H. Freeman and Company, New York, NY.

Wilhere, G.F. 2002. Adaptive management in habitat conservation plans. Conservation Biology 16:20-29.

WDFW 2009. Fish Passage Barrier and Surface Water Division Screening Assessment and Prioritization Manual . Habitat Program, Washington Department of Fish and Wildlife, Olympia, WA.

Appendix A. Implementation Monitoring for Culverts

Implementation monitoring consists of monitoring for permittor failure and permittee failure. There are 3 types of permittor failure: a) a permit does not comply with WAC 220-110; b) a permit approves a culvert that is inappropriate for the project site, i.e., the approved culvert does not follow WDFW's design guidance; and c) the inaccuracy of key measurements necessary for project design and submitted with the plans result in a culvert that is inappropriate for the project site. Permittee failure occurs when a completed culvert fails to meet or exceed design specifications stipulated in an HPA permit. To evaluate permittee failure, the constructed culvert will be compared to the provisions in the HPA and the plans attached to the HPA.

The first steps of implementation monitoring involve checking the permit and construction plans. This serves two purposes: 1) preparing for collection of field data, and 2) determining whether the permit complies with hydraulic code rules and/or WDFW's design guidelines.

Office Data Collection

A. <u>Basic Information</u>
   See data form.

B. <u>Evaluate Permit Provisions</u>
   1. Do the permit or plans show the following key site measurements:
      a. channel slope?
      b. bank full width (BFW)?
   2. Determine type of culvert: no-slope, stream simulation, bottomless, or hydraulic design
      Does permit state the type of culvert design?
      If no, then determine the type of culvert design from the provisions and plans.
      (Note: If culvert specifies hydraulic design, then discontinue the in-office permit review. The culvert will still be evaluated for permittee failure.)
   3. If the permit specifies a no-slope culvert, then do the permit or plans specify or show:
      a. channel gradient less than 3%
      b. level, flat, or 0% culvert slope?
      c. culvert width at the stream bed equal to or greater than average BFW
      d. minimum 20% counter sinking of the culvert at the outlet
      e. maximum 40% counter sinking of the culvert at the inlet
   4. If the permit specifies a stream simulation culvert, then do the permit or plans specify or show:
      a. culvert slope less than or equal to 1.25*upstream channel slope?
      b. culvert width at the stream bed equal to or greater than 1.2*average BFW + 2 ft
      c. 30% to 50% counter sinking of the culvert at the outlet
      d. size of bed material to be placed in culvert
      e. placement of rock bands in culvert.
   5. If the permit specifies a bottomless culvert, then do the permit or plans specify or show:
      a. culvert width at the stream bed equal to or greater than 1.2*average BFW + 2 ft
      b. size of bed material to be placed in culvert

C. <u>Evaluate Appropriateness of Culvert Design</u>
   1. Given the information recorded in Section B, select one of the following:
      a. the biologist permitted an appropriate culvert design for the site
      b. the biologist lacked key site measurements needed to evaluate the appropriateness of the proposed design

c. the biologist lacked culvert dimensions needed to evaluate the appropriateness of the proposed design
d. the biologist permitted an inappropriate culvert design for the site
Explain why:

Field Data Collection

A. Basic Information
See data form.

B. Measure Key Culvert Design Parameters
Measure the habitat attributes used to design the culvert:
1. Measure BFW
See Appendix C in Barnard et al. (2013)
2. Measure channel slope
3. If the HPA permit includes provisions specifying the required size of stream bed sediments, then measure size of stream bed material. See methods for measuring bed material in Appendix B, step I.

C. Collect other data needed to complete the form Freshwater Culverts, Implementation Monitoring.

D. Complete a Level A barrier assessment.
See WDFW 2009 for procedures and methods.

E. If barrier status cannot be determined with Level A assessment, then complete Level B barrier assessment (see WDFW 2009 for procedures and methods).

Table A1. Tolerances for key measurements necessary for project design. The tolerance to apply is the more lenient of either percent error or absolute error.

	Measurement	**Allowable Percent Error**	**Allowable Absolute Error**
channel characteristics	bankfull width	± 10%	± 0.5 ft
	channel slope	± 5%	± 1%
	substrate size		
no-slope culvert design	culvert slope	--	± 2%
	culvert width at streambed	- 5%*	- 0.5ft
	culvert length	+ 5%*	
	% countersink at inlet		-5%
	% countersink at outlet		-5%, +20%
stream simulation culvert design	culvert slope	--	2%
	culvert width at streambed	- 5%	-0.5 ft
	culvert length	+ 5%	
	% countersink at inlet		-5%
	% countersink at outlet		-5%, +20%
	substrate size		

* If + or – rather than ±, then tolerance is one-sided. That is, if "+", for instance, then there is a limit on dimension being bigger and no limit on dimension being smaller.

Permitting Biologist ______________________________ Biologist's telephone_________________________-

Completion Date ________________________

Latitude _______________________________ Longitude __________________________

Landowner/Applicant Name __________________________________

Landowner/Applicant Phone ___________________________

Freshwater Culverts
Implementation Monitoring

<u>Pre-site Information from HPA Permit and Construction Plans</u>

Water Crossing Structure Design

☐ no-slope ☐ stream-simulation ☐ hydraulic ☐ unknown

☐ other ________________________

Culvert Shape

☐ round ☐ elliptical ☐ box ☐ squash

☐ bottomless ☐ unknown ☐ other ________________________

Culvert Material

☐ PCC ☐ CPC ☐ CST ☐ SST ☐ CAL ☐ SPS ☐ SPA ☐ PVC ☐ TMB

☐ MRY ☐ UNK ☐ OTH ________________________

Configuration

culvert span _________ culvert rise _________ culvert length _________ culvert slope _________

Culvert Bed

streambed slope (within culvert @ thalweg) _________

countersunk depth at outlet[1] _________ countersunk depth at inlet[2] _________

outlet invert elevation _________ inlet invert elevation _________

Number of Coarse Bands __________ Baffles __________

Culvert width at streambed: downstream _________ upstream _________

BFW (if provided) _________ Streambed slope (if provided) ____________________________________

Other features described in permit or plans (e.g., armoring, grade controls, fishways, LWD, etc.) AND any comments:

____

____

____

[1] Downstream end of culvert must be countersunk minimum of 20% of culvert rise. N/A for bottomless.

[2] For stream sim: upstream end of culvert must be countersunk 30-50% of culvert rise. For all other designs: upstream end of culvert should be countersunk 20-40% of culvert rise. N/A for bottomless.

On-site Information

Reviewer Name(s) ___________________________
Date of field review __________________________
Lat/Long ___________________________________

Water Crossing Structure Design

☐ no-slope ☐ stream-simulation ☐ hydraulic ☐ unknown

☐ other ___________________________

Culvert Shape

☐ round ☐ elliptical ☐ box ☐ squash

☐ bottomless ☐ unknown ☐ other ___________________________

Culvert Material

☐ PCC ☐ CPC ☐ CST ☐ SST ☐ CAL ☐ SPS ☐ SPA ☐ PVC ☐ TMB

☐ MRY ☐ UNK ☐ OTH ___________________________

Configuration

culvert span __________ culvert rise __________ culvert length __________ culvert slope __________

Culvert Bed

streambed slope (within culvert @ thalweg) __________

countersunk depth at outlet __________ countersunk depth at inlet ___________

benchmark located ☐ Yes ☐ No

their benchmark elevation __________ your benchmark elevation __________

outlet invert elevation __________ inlet invert elevation __________

Number of Coarse Bands ___________ Baffles ___________

Culvert width at streambed: downstream __________ upstream __________

BFW __________ WDIC __________ Backwatered ☐ Yes ☐ No Apron ☐ Yes ☐ No Road Fill __________

Other features (e.g., armoring, grade controls, fishways, LWD, etc.)

____
____
____

Comments (e.g., tidally influenced, pond or wetland, near confluence, etc.):

____
____
____
____

Appendix B. Effectiveness Monitoring for Culverts

Data Collection

A. Basic Info. Record the HPA Number, date, location information, and crew members on the survey form.

B. Benchmark. Identify a sound location for placing an elevation benchmark (six-foot-long metal fence post); a location that is unlikely to be disturbed by natural or anthropogenic processes for the next 50 years. It should be outside of the current or anticipated bankfull channel, not adjacent to the road, near an easily identified landmark and within clear sight of the project. Place the fence post until only about one foot remains above ground. Collect and record the GPS coordinates of the benchmark (BM). Take a photograph of the BM that includes the landmark to facilitate easy relocation in subsequent years.

C. Structure Integrity. Complete field data form for implementation monitoring of culverts (Appendix A).

   This is also part of implementation monitoring in year 1. Continuing to meet design standards or WDFW's design criteria in subsequent years (years 2, 5, 10, and 20) is a measure of project effectiveness; that is, whether it continues to be passable through time based on our current design standards is a measure of effectiveness.

D. Culvert Slope.
   a. Measure elevation of the top and bottom (e.g., invert, bed) of the inlet and outlet using a laser level. (0.00 m)
   b. Measure and record culvert length. (0.00 m)
   c. At the inlet of the structure, record whether a mid-channel bar (i.e., island with or without vegetation) is present. Take a photograph.
   d. At the outlet of the structure, record whether a mid-channel trough (i.e., entrenchment within the active channel) is present. Take a photograph.

E. Layout Cross-sections.
   a. At 4 locations, each of which are about 10 m apart, measure bankfull width and estimate mean bankfull width ($\bar{W}_{bf}$).
   b. Transects E and F are to be positioned at the culvert's outlet and inlet, respectively
   c. Transects D and G are to be positioned two $\bar{W}_{bf}$ down- and up-stream of the culverts outlet and inlet, respectively. All other transects are to be positioned 2 $\bar{W}_{bf}$ apart and evenly spaced, *perpendicular* to stream flow (see Figure B1). (0.0 m)
   d. Cross-sections are all measured from left to right, facing upstream, from left bankfull edge to right bankfull edge.
   e. Cross-sections are lettered A through J from downstream to upstream and should be measured in that order to allow for estimate of gradient-based statistics.

F. Site Measurement – follow standard stream elevation survey methods. All measurements must be collected using metric instruments (e.g., metric tape measures, metric survey rods, etc.). Metric measurements allow the survey team to quickly and accurately calculate changes in elevation and distance while surveying, allowing for quality assurance checks in the field and simple calculation of statistics and metrics for monitoring.

a. *Benchmark* All elevations are referenced to the benchmark (BM) elevation, 100.000 m. Note that the BM might best be placed within accurate measurement distance of Transect A (downstream-most) and the project, if possible, to minimize the need for turn points.
b. *Instrument height* (HI) is measured by taking a backsight (BS) to the BM. HI can be negative or positive, depending on the relative locations of the BM and instrument (e.g., survey autolevel). (0.000 m)
c. *Backwater* While upstream of the structure, the elevation of the highest observable water line on or within one BFW of the structure should be measured to detect backwater events at high flows. (0.000 m)
d. *Order* Transects should be measured in order, from A through J, to allow for simple and direct mapping of the stream and estimate of stream statistics (e.g., stream gradient, mean depth, mean cross-sectional area, mean width-to-depth ratio).
e. *Transect Measurements* On each transect, measurements of distance (tape measure; 0.00 m), elevation (0.000 m), water depth (0.000 m), and substrate size (see below) should be taken at a minimum of twenty relatively evenly spaced locations and at the locations of any large changes in the cross-section profile. The first elevation at each cross-section should be at Left Bankfull, where depth will usually be zero. Between left bankfull and left wetted edge at least one measure of the bank elevation should be collected to allow for estimation of changes in bank angle. Within the wetted channel (or dry stream bed) these same measurements should be taken wherever an abrupt change in channel shape occurs, including the edges and top of mid-channel bars (e.g., islands), or at approximately equal distances, when the channel shape is regular. Note that we are collecting data that describe channel shape, not fish habitat available, so information that describes the elevation of mid-channel bars and dry stream bed (i.e., depth = 0) are very important.
f. *Sediment Sizes* At each location where a bank or bed elevation is measured, estimate the size of a single particle of the surface sediment that is located at the left-forward corner of the survey rod. Size classes are described in Table B1.

Table B1. Size classes for streambed sediments.

Size Class	Code	Size Range (mm)	Description
bedrock; smooth	RS	> 4,000	larger than a car
bedrock; rough	RR	> 4,000	larger than a car
boulder	BL	250 - 4,000	basketball to car
cobble	CB	64 - 250	tennisball to basketball
coarse gravel	GC	16 - 64	marble to tennisball
fine gravel	GF	2 - 16	ladybug to marble
sand	SA	0.06 - 2	gritty to ladybug
silt/clay/muck	FN	< 0.06	loose fines, not gritty
hardpan	HP	na	firm consolidated fine
vegetation / organic soil	VO	na	could occur at top of bank
wood	WD	na	log or branch, any size
other	OT	na	

G. Between transect measurements.

a. Pool Depth Between each transect, if a pool is evident, measure its maximum depth and the depth of the pool tail (0.00 m).

b. LWD Because of the importance of large wood debris (LWD) to the transport of sediment in stream channels, LWD will be counted between each transect. Pieces of LWD that are within the bankfull channel and that are greater than 2-m long and have a diameter greater than 30-cm should be counted. Count each piece only once.

H. Streambed at Culvert Inlet and Outlet.

a. At the inlet of the structure, record whether a mid-channel bar (i.e., island with or without vegetation) is present. Take a photograph.

b. At the outlet of the structure, record whether a mid-channel trough (i.e., entrenchment within the active channel) is present. Take a photograph.

I. Streambed in Culvert. (stream simulation culverts only)

a. Thalweg profile – Measure bed surface elevation and water depth of the thalweg at 10 approximately equally spaced locations within the structure. Note that elevation measurements should be in reference to the benchmark elevation.

b. Sediment sizes – At every location where thalweg depth is measured collect three substrate particle sizes (roughly, one left, center and right of the center of the channel) use the same size classes as in section F, subsection f, above.

c. Measure the length of the structure (straight line maximum length, 0.00 m).

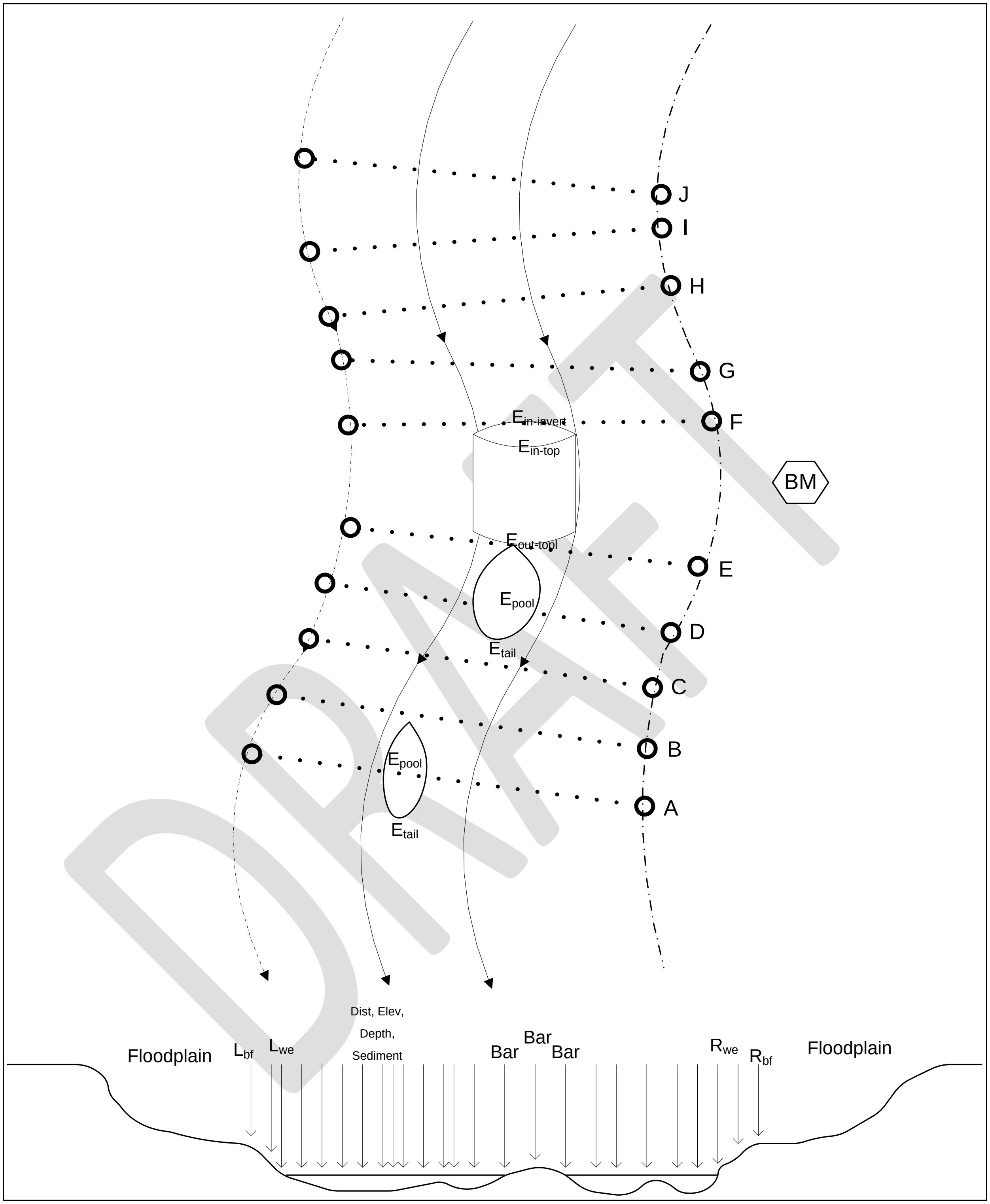


Figure B1. Basic layout of stream survey cross-sections for effectiveness monitoring of stream crossings.

Appendix C. Implementation and Effectiveness Monitoring for Marine Shoreline Bank Protection

Both implementation and effectiveness monitoring will be conducted at nearly all sites where it is practicable, hence, the methods for both are presented in a single appendix.

Field Procedure Summary

1. Pre-project evaluation:
    a. For sites where the structure has not yet been constructed, locate and document a local bench mark such as a large boulder, tree, or corner of a foundation that is not likely to move or be affected by the construction activity. Mark, photograph, and record the GPS location and elevation of the bench mark using the RTK GPS.
    b. Using the GPS, collect data points at the upland toe, MHHW, and ordinary high water within or seaward of the proposed footprint of the shoreline armoring structure.
    c. Proceed with the remainder of the survey tasks listed below in sections 3 through 7.
2. Structure Dimensions:
    a. If a local bench mark was established during a previous survey, return to the bench mark and record the location of the bench mark on the GPS.
    b. Records the construction material used for armoring and using the GPS, record the elevation of the toe of the armoring at the most seaward point of the structure, and the greatest height of the structure measured from the toe of the armoring.
    c. Using a tape measure, measure the straight length of the shoreline armoring, and the maximum perpendicular distance from the upland toe to the toe of the structure. (Depending on the size, shape, and type of structure, these measurements may be taken with the GPS by calculating the distance between two points).
    d. Record any additional site specific provision observations or measurements identified during the permit review process.
3. Beach Profile:
    a. Run a tape measure from toe of armor or toe of bluff to MLW or MLLW as determined from predicted tide times and levels.
    b. Using the tape measure as a guide, collect a data point with the RTK GPS every 2 to 4 meters, and gather a data point at OHW, the toe of bluff or toe of armor, center of wrack line(s), lower edge of log line, center of potential forage fish spawning substrate, and obvious changes in beach profile (slope) or grain sizes (e.g., cobble-sand transition that often occurs near MLLW).
4. Sediment grain sizes:
    a. Characterize the sediment by percent cover of each grain size category at several key elevations using the 0.25 m^2 quadrat. These elevations should include the wrack line and/or toe of armor, MHW, and MLW. For each sample estimate the percent cover of the surface layer of sediment, then scrape off the surface layer (~10 cm) and estimate the percent cover for the subsurface layer.
    b. In addition to percent cover samples, collect a bulk sample of sediment to the depth of the largest mobile grain size observed (generally less than 64 mm). The sample size should be equal to about 100x the weight of the largest grain as measured along the y-axis (Church et al. method). Bulk samples will be collected in one to three 2 gallon buckets depending on the grain size and wave dominant wave energy on the beach. Quantitative grain size analysis in the laboratory involves sieving dry sediments through a stack of progressively finer sieves using a mechanical sieve shaker for 12 minutes, and weighing the amount retained in each sieve.
5. Wood and wrack:

   a. Log surveys: Run a 50-m transect line near the upper edge of each beach. Measure the width of the log line (from bluff seaward) at 5 random points along the transect and count the number of logs intersecting the width measurement line. Record the number of large (> 2 m length) and small (< 2 m length) logs, and number that are in contact with the sediment. Count the number of trees fallen from the bluff and record the general orientation of logs relative to shore (parallel, perpendicular, skewed).
   b. Wrack surveys: Using the same transect line used for the log survey; taking measurements from the center of the wrack line whenever possible. Measure the width and the depth of the wrack line at 5 locations. Using a 0.25 m^2 quadrat, record the percent cover of the quadrat occupied by wrack at each location.
   c. Wrack samples: At the same 5 points, collect a 15 cm diameter core of wrack and the 1 – 2 cm of sediment beneath it and place this sample in a sealed bag for analysis in the lab. At the lab, separate the wrack from the sediment, dry and weigh the wrack; and remove and count associated amphipods.

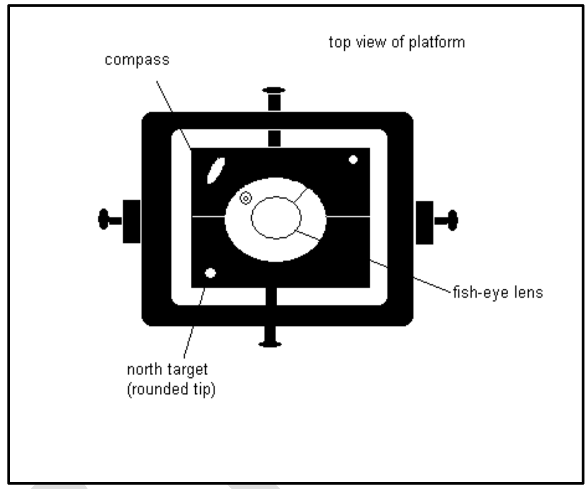
6. Forage Fish spawn: (*reference Moulton and Penttila (2001) for details*)
   a. At the each site identify the band of substrate near the likely center of spawning activity (Penttila 2011). Following the methods describer in Moulton and Penttila (2001), collect several scoops of the top several cm of sediment at four points along this band at intervals of about 10 m. The volume of sediment collected at each point should be similar, and the final volume of the sample should fill an 8” x 24” plastic bag ½ to 2/3 of the way full. Label each sample with its location and date. Store samples in a cool place for up to 48 hours until they are either processed or preserved in Stockard’s solution.
   b. Upon returning to the lab, sieve and winnow the samples as described by Moulton and Penttila (2001).
   c. Identify, count, and record numbers of live and dead eggs observed using Moulton and Penttila (2001) methods.
7. Shade

Hemispherical Photography Equipment and Protocol (adapted from Mark Hunter)

Equipment:

Camera (in hard case)
Memory card
2 Batteries & recharging units
Tripod
Platform with fish-eye lens

Objective: The hemispherical photography protocol will allow us to estimate the amount of shade a point on the beach receives during any period of time during the year (assuming a clear sky). The photo is interpreted by a computer software program that predicts the sun’s path based on longitude, latitude and date. Because we understand and can predict the sun’s path, the hemispherical photo can be taken at any time of day and on any day provided you follow the protocol below. Because we are trying to estimate the amount of shade provided by topography (e.g., cliffs and bluffs) and vegetation, we need to wait until full leaf out (i.e., summer months) to take these photos. The idea is to take a single photo at each of the sample sites at the center of the upper sampling transect (further defined below).



Protocol:

1. Set up tripod on stable surface near the center of the transect line, and attach the pan/tilt handle, and vertical tilt handle (these should be in the tripod bag).

2. Attach platform onto tripod by seating the quick shoe at the base of the platform into the pan head of the tripod and close the locking lever to hold it in place (you must open the lever, and push the release button to remove the platform).

3. Attach camera: The camera lens screws into the fish-eye lens from the bottom of the platform. Be careful when attaching the camera to the lens as the threads can easily be stripped. Make sure that the camera battery is positioned opposite of (away from) the platform compass; otherwise the battery will interfere with the compass reading and you will not be able to orient the platform (see #4).

4. Using the platform compass, orient the tripod so that the rounded plastic tip is facing **magnetic north**.

5. Level the platform: Using the bubble indicator, adjust the tripod's pan/tilt and vertical positions, and the adjustable bar weights until the platform is level. You may have to get creative here by adding small pebbles to the outside of the platform until it is completely level, or by hanging the plastic lens cover from one edge of the platform. Check the view through the camera to make sure the view isn't obstructed.

6. Turn camera on. The camera battery will hold a charge for 2 days. Daily recharging of the battery is advised. There is 1 extra battery and 2 extra memory cards for the camera in the hard case. There are 2 charging units for the batteries.

7. Before taking the photo, make sure that everyone near the tripod ducks below the platform, or moves out of range, otherwise they will be captured in the photo, and will block the canopy shot. Make sure you record the number of the photo with the correct beach site number.

8. For best quality photos, take photos before or after the mid-day glare (before or after 10AM to 2PM), when the sun is directly overhead. The glare (sunlight reflecting off of surfaces like leaves) can appear as open sky and thus be difficult to interpret in the photo resulting in inaccurate assessments of canopy cover. Moderate to heavy overcast provides the best lighting for photos. Earlier morning or late afternoon times are best.

Sampling Procedures

We will measure the ecological and geomorphic characteristics at sites with recent shoreline armoring as well as sites with little to no recent armoring. Ideally unmodified sites will be located within the same drift cell and with similar geologic, oceanographic, and geomorphic settings. We will visit beaches during daytime low tides.

The sampling unit is a 50 meter transect running parallel to the shoreline. At each site a 50m long shore transect will be established for data collection. Replicates will be collected along the transect at randomly selected points using a random a number generator. Monitoring activities along the transect are split into several tasks.

Compliance assessment, and site and armoring characterization

Prior to the compliance assessment we will compile a check list of measurable and observable permit provisions for each site. This list will be used to identify and note whether observable provisions, such as armoring material are either in or out of compliance. For measurable provisions, such as structure length, measurements will be taken and recorded. In addition to the provision check list, the following data will be collected at each site. The type, length, height, width, building material, and orientation of armoring (if present) will be recorded. The backshore characteristics such as vegetation, bank height, and shore type will be noted, and we will photograph the armoring, and take a hemispherical photograph at the center of each site for use in analysis of shading.

Beach geomorphology and sediments

A tape measure will be laid out from the toe of armor or the upper extent of the beach to Mean Low Water (MLW) or Mean Lower Low Water (MLLW) as determined from predicted tide times and levels. This cross shore transect is the line down which beach elevation/slope will be measured and sediment samples to characterize the site will be collected. A GPS unit will be used to record points along the length of the cross shore transect. The cross shore transects will be within and perpendicular to the 50m long shore transect along which biological samples are collected.

An RTK GPS will be used for elevation profiles unless the signal strength is too weak to operate the GPS; in this case a laser level will be used to measure elevation along the transect relative to a local bench mark. We will gather elevation data along the long shore transect line at every two to four meters as well as at Ordinary High Water (OHW), the toe of the shore armoring or backshore, obvious changes in beach profile (slope) or grain sizes (e.g., cobble-sand transition), MLW, and MLLW if possible. Lengths along the beach profile will be measured to the nearest tenth of a meter, and elevation will be measured to the nearest centimeter. Data recorded in the GPS unit will be downloaded at the office and used to make two dimensional plots of the beach profile at each site.

We will sample beach sediment at several key elevations, including the wrack line, toe of armor or backshore, Mean High Water (MHW), and MLW. Sediment will be sampled by estimating the percent cover of sediment in each of five grain size categories using a 0.25 m^2 quadrat divided into 25 equal squares. An estimate of percent cover will be made for the surface layer of sediment and the subsurface layer by scrapping away the top 10cm of sediment and repeating the procedure. An additional sediment sample will also be collected at MHW using a frequency-by-weight bulk sampling method. In this method, a volume of material is excavated from the beach and sieved into half-phi size classes. Each size class is then weighed and cumulative frequency distributions developed. The frequency-by-weight sampling method requires a volume of material based on the largest mobile particle on the surface. The sampling volume must be large enough that the largest particle is less than 1%, by weight, of the sample to obtain significant results (Church et al.,1987). The surface area covered by the sample is approximates and depends on the type of sediment. The coarser the sediment, the larger the sample size: approximately 5 x 10 cm for sand, 10 x 15 cm for pebble, 20 x 20 cm for cobble. This method best represents the low % of coarse sediments in the overall sample. Quantitative grain size analysis in the laboratory involves sieving dry sediments through progressively finer sieves and weighing the amount retained in each sieve.

Forage fish egg presence and health
The Forage fish sample will evaluate presence, relative abundance, and condition of eggs and embryos. Field work will consist primarily of sampling the surface layer of beach substrate for surf smelt and sand lance eggs based on existing methods (e.g., Moulton and Penttila 2001) to document presence/absence and quantify proportions of embryo condition (live/dead). Timing for sample collection will be based on known spawning seasons for surf smelt near a given sample site. The samples will be labeled with collection site and date, and stored in a cool place for no more than 48 hours before they are processed or preserved in either Stockard's Solution or ethanol in 16oz. sample jars.

Laboratory analysis will measure egg/embryo counts, proportion live (at time of collection), and hatched eggs. Portions of each sample will be dispensed into clear petri dishes and examined under a dissecting microscope until the entire sample has been processed. Embryos will be considered live if translucent and dead if opaque. In cases of extremely high abundances, subsamples will be taken by mixing the sample and taking a fixed volume for processing under the microscope.

Log and wrack detritus
Log surveys include drift wood and trees recruited from the bluff. For trees recruited from the bluff, trees that have clearly fallen from the land (with roots) are counted; trees that have clearly been cut and thrown over the edge of the cliff are not counted. The general orientation of the fallen logs is recorded (parallel, perpendicular, skewed). This is usually 90 degrees (perpendicular) to the beach. For logs recruited from the sea, the width of the log line (from bluff seaward) is measured. At five random points along the transect tape, the number of logs intersecting the tape are recorded, along with their diameter: large (> 2 m length) or small (< 2 m length).

Using the same transect line as the log survey, we will measure the width and percent cover of wrack material at 5 random points using a 0.25 m^2 quadrat. The percent cover will include all things other than substrate under the quadrat and human debris.. At these 5 random survey points, we will also collect a 15 cm diameter core of the wrack and 1-2 cm of sediment immediately beneath the wrack and place in a plastic bag for analysis in the lab. The bag will contain a label with the site name and sample number on it. The samples will be frozen or preserved with ethanol for lab analysis. Lab analysis will consist of drying and weighing the sample, and sorting the sample to count the associated amphipods.

All field data will be recorded on data sheets adapted from other studies that have used similar methods. In addition to recording parameters measured in the field, field notes will also record the field personnel and their rolls, label information for collected samples, and a notes or comments section to record any unusual circumstances that may affect interpretation of results. All samples collected in the field will be labeled with a sample ID which will allow us to track the date, location, and replicate of the sample. Sample labels with either be attached directly to the sample container or placed within the sample container with the sample.

All measurement tools and monitoring equipment will be maintained in good working order. Forage fish sampling gear will be rinsed between sites to minimize the probability of cross site contamination.

All of our samples will be collected within the same water body (Puget Sound) helping to minimize the threat of contaminating areas with invasive species that were not already present. To further reduce the risk of invasive species transport and contamination we will follow the standard operating procedures outlined in the Washington State Department of Ecology document EAP070 (http://www.ecy.wa.gov/programs/eap/quality.html).

Measurements

Measurements made in the field or under controlled conditions, for which accuracy applies, are: weights, counts, lengths, shade, and elevation/location. Aside from user error, the accuracy of these parameters is inherent in the tool used to measure. The biological and sedimentological data we will be collecting are subject to so much natural variation that our goal is to capture a 'summary' condition through field replicates that we can then compare to other sites. Sampling stratified by parameters that we believe are primarily responsible for variation will result in patterns that can be translated into correlations. Table C1 lists the various field and lab methods, specifies applicable measurement accuracy and precision, and specifies the number of replicates we will collect to summarize site conditions and evaluate measurement precision.

Table C1: Summary of measurement methods for implementation and effectiveness monitoring of marine shoreline bank protection.

Parameter	Sample Description	Collection/Storage	Measurement Methods	QC Samples	Measurement Quality Objectives	Reference
Pre-project	Use GPS to record local bench mark and location of upland toe, MHW, and OHW	In situ measurements recorded on the inernal memory of the GPS	RTK GPS used to record 3D location of points of interest so that they may be reference after the shoreline armoring is constructed	2 replicates per beach	Accuracy: ±0.1m	Appendix A Section 1
Structure Dimensions	Use measuring tape and stadia rod, or GPS to measure length and height of the the structure	In situ measurements	Measure height, and length of shoreline armoring structure	na	Accuracy: ±0.1m	Appendix A Section 2
Beach profile	Use RTK GPS to measure x,y,z (location and elevation) of data points a long a transect perpendicular to the water	In situ measurements recorded on the inernal memory of the GPS	RTK GPS used to record 3D location of points spaced 2-4 m apart a long a cross beach transect line	2 replicates per beach	Accuracy: x,y,z= ±10cm	Appendix A Section 3
Bulk sediment samples	Collect 100x the weight of the largest cobble in view from the upper 10 cm of sediment.	Buckets, no preservation, dry samples	Sieve on standardized sieves, at modified Wentworth intervals, in a sieve shaker for 12 minutes. Weigh fractions on a digital scale.	Homogenize and split 1 sample per season	Accuracy: fraction wt. = ±10% Weigh fractions to the nearest 0.05 g	Appendix A Section 4b; Church et al. 1987
Sediment samples	0.25 m^2 quadrat divided into 25 equal squares used to estimate % cover of grain sizes within 5%	In situ measurements	Visual estimate of the percent of the surface sediment in each of 5 size classes for the surface sediment and sediment 5cm below the surface	2 replicates per elevation	This sample is redundant with sediment samples collected at MHW for lab sieving	Appendix A Section 4a
Detritus-logs	Measure width of the log line to the nearest 1 cm	In situ measurements	5 random samples along a 50 m transect where log count, orientation, and size class are recorded	5 replicates per beach	na	Appendix A Section 5a
Detritus-wrack sample	Measure width of the wrack line to the nearest 1 cm; estimates % cover within wrack line with 0.25 m2 quadrat divided into 25 equal squares used to estimate % cover within 5%	In situ measurements	5 random samples along a 50 m transect where wrack % cover are recorded using a 0.25 m^2 quadrat	5 replicates per beach	na	Appendix A Section 5b
Detritus-sediment core	15 cm diameter core of the wrack and 1 inch of sediment beneath the wrack	Sealed jar or bag containing label w/ site name & sample number	In the lab, sort to count associated amphipods, and weigh wrack material	5 replicates per beach	Accuracy: Sample wt. = ±1 g Precision: Dupe count = ± 10%	Appendix A Section 5c
Shade	Use fish eye lens to take hemispherical photo of sky	Digital photograph	Use software to anaylize photo and estimate shading	2 replicates per beach	Precision: shaded area = ± 10%	Appendix A Section 7
Forage fish	Scoops of top 3-5cm of sediment collected form 4 locations along transect to fill a 8"x24" plastic bag 1/2 to 2/3 full	Stored in a cool place and preserved in 16oz jar w/in 48 hours	Forage fish sample method in Moulton & Penttila (2001) http://wdfw.wa.gov/publications/01209/wdfw01209.pdf	2 replicates per site	Precision: Dupe count = ± 10%	Appendix A Section 6; Moulton & Penttila 2001

Appendix D.
Landowner/Permittee Contact Protocols for Hydraulic Project Implementation and Effectiveness Monitoring
(1/22/14 Draft)

This landowner/permittee contact protocol applies to WDFW employees engaged in hydraulic project implementation and effectiveness monitoring through the Habitat Program's Science Division. This protocol is different from those described in Policy - 5212: *Monitoring Compliance with State Hydraulic Code (Chapter 77.55 RCW)* and Procedure-5212 that pertain to Habitat Biologists and Wildlife Officers.

Purpose and Scope of Implementation and Effectiveness Monitoring

Implementation and effectiveness monitoring by the Habitat Science Division is part of an adaptive management process applied to the Habitat Program's hydraulic project approval (HPA) authority. The intent of adaptive management is to better understand and thus improve HPA implementation (process of issuing permits and the compliance with those permits) and effectiveness (how projects meet the goal of protecting fish life). The monitoring is designed to collect scientifically valid information that can be used by others in the Habitat Program to improve HPA outcomes. This monitoring is **not** a formal part of the process to identify and investigate hydraulic code violations. (see Policy and Procedure – 5212 for those activities).

Gaining Access to Hydraulic Projects for Monitoring Surveys

Notwithstanding an HPA permittee's signature on the Joint Aquatic Resource Permit Application that consents to entry onto property by permitting agencies, biologists engaged in implementation and effectiveness monitoring (hereafter Monitoring Biologists) must attempt to contact, by telephone or email, landowners to secure permission prior to conducting monitoring activities on their property. This request must include the Monitoring Biologist's name; contact information; the purpose, date, and time of the site visit(s); and the location of the site identified in the signed application for an HPA. In addition, the Monitoring Biologist must document the time and date and form (i.e., telephone or email) of the request.

If the landowner does not respond to a telephone or email request, then the monitoring activity can proceed provided the landowner is notified by telephone or email that the Monitoring Biologist will visit the project site. This notification must include the monitoring biologist's name; contact information; the purpose, date, and time of the site visit(s); and the location of the site identified in the signed application for an HPA. The Monitoring Biologist must document the time and date and form (telephone or email) of this notification.

If the landowner does respond and denies access to their private property, then the monitoring activity cannot proceed at that site. The Monitoring Biologist will inform the appropriate Habitat Biologist that access was denied.

In addition, if a Monitoring Biologist is visiting a site without explicit spoken or written approval from the landowner, and the landowner's residence is located on the site and readily accessible by foot, then the Monitoring Biologist must seek permission by attempting to contact the landowner at his or her residence. That is, the Monitoring Biologist must visit the landowner's residence to seek permission and provide the landowner with the following information: your name, contact information, the purpose, date and time of the monitoring visit(s). The Monitoring Biologist must document the time and date of their personal communication with the landowner. If the landowner is present and denies access or if the monitoring biologist feels threatened in any way, they should immediately exit the property.

Observing Potential Violations of a HPA Permit during Routine Monitoring Work

If a Monitoring Biologist observes a probable violation of an HPA permit during a site visit, then he/she will complete all monitoring procedures as they would with compliant HPAs, and then inform the appropriate Habitat Biologist by phone or email *as soon as possible*. Document the date you observed the probable violation(s), the nature of the violations and date you contacted the Habitat Biologist. Do not under any circumstances contact or confront the landowner about the potential violations. This is strictly the purview of the Habitat Biologist and Enforcement Officers.

If you observed probable HPA violations on projects other the ones you selected to monitor and have followed the protocols for **Gaining Access to HPA Projects during Monitoring Surveys** (above), contact the appropriate Habitat Biologist at your earliest convenience. Document the date you observed the probable violation(s), the nature of the violations and date you contacted the Habitat Biologist. Do not under any circumstances contact or confront the landowner about the potential violations. Again this is strictly the purview of Habitat Biologists and Enforcement Officers.

Appendix E. Minimum Sample Size Estimate for Detecting a Difference Between Two Proportions

Selecting sample size entails subjective judgment and compromise. We must make judgments regarding acceptable levels of Type I and Type II errors, but “acceptable levels” may be compromised by the cost of sampling. That is, we desire a very low probability of error, but our desire may be frustrated by lack of resources for monitoring.

We will test the hypothesis that success (or failure) rates in year 1 are different than success rates in subsequent years. Successes and failures can be modelled as independent Bernoulli trials which conform to a binomial distribution. The minimum sample size per year, *n*, needed for comparing proportions estimated from two independent binomial samples can be estimated with the equations (Sokal and Rohlf 1995, p. 768):

$$n = \frac{A\left[1 + \sqrt{1 + 4|\pi_0 - \pi_1|/A)}\,\right]^2}{4(\pi_0 - \pi_1)^2} \tag{E1}$$

$$A = \left[t_{\alpha,\infty}\sqrt{2\bar{\pi}(1-\bar{\pi})} + t_{2\beta,\infty}\sqrt{\pi_0(1-\pi_0) + \pi_1(1-\pi_1)}\right]^2 \tag{E2}$$

$$\bar{\pi} = \frac{(\pi_0 + \pi_1)}{2} \tag{E3}$$

where π_0 and π_1 are failure rates (expressed as proportions) at time 0 and time 1, α is the probability of Type I error, β is the probability of Type II error, *t* represents Student’s t-distribution.

To determine sample size we need values for α, β, π_0 (failure rate at time 0), and effect size[7]. Effect size equals π_0 - π_1. Relationships of sample size to α, β, π_0, and effect size are shown in Figure 6. The null hypothesis (H_0) is no change in success (or failure) rate between year 1 and subsequent years. We are more concerned with avoiding Type I error than avoiding Type II error. That is, we do not want to infer a change in success rate when in reality there was no change in success rate. Therefore, we choose α = 0.05. According to Table 1, failure rates for culverts were about 0.3 in 2010. We want (and expect) the effect size to be negative. To detect an effect size of -0.1, sample sizes range from 251 to 412 per year, depending on β (Table E1), which would require approximately 16 to 26 separate site visits per week during the field season. To detect an effect size of -0.2, sample sizes are greatly reduced: 59 to 92 samples per year, which would require approximately 4 to 6 site visits per week, on average. Sixteen site visits per week for the entire field season is beyond our current capacity. One the other hand, an effect size of 0.2 corresponds to a change in failure rates from 0.3 to 0.1, which seems rather unlikely. Therefore, we need the ability to detect smaller changes on the order of 0.15. To detect an effect size of -0.15, sample sizes range from 108 to 174, depending on β, which would require approximately 7 to 11 separate site visits per week. A β of 0.3 means a 30% chance of not detecting a change in success rate when a change actually occurred. That error rate seems too high. A β of 0.2 is more reasonable, and therefore, our minimum sample size is 134 sites per year, which corresponds to about 8.5 site visits per week, on average.

7 Sample size calculations need an estimate of variance. In equation E2 variance estimates are incorporated by $\pi_x(1-\pi_x)$. The sample size determination depends on initial failure rate, π_0, because effect sizes for proportions are symmetrical about 0.5.

The minimum sample size of 134 culverts per year does not take into account the realities of the HPA implementation monitoring. The actual number of culverts that can be monitored each year is limited by the number of culverts that are actually available and the number of field staff available to do the monitoring. We may not meet our minimum sample size because the number of available culverts in a given year is too small or because the number of field staff available for monitoring is too small. Problems with field season duration, notification of project completion, and access to private lands may reduce available culverts to a number less much less than 134.

In 2013, 48 culverts were visited for implementation monitoring. This was accomplished with 1 field crew leader and 1 field crew all working half-time (i.e., 20 hours/week) on culvert monitoring. If sample size is only 48 sites per year, then we can detect a change in failure rate of only -0.24. In other words, in order to detect a change that is statistically significant, HPA implementation failure rates must improve from the current failure rate of about 0.30 to 0.06.

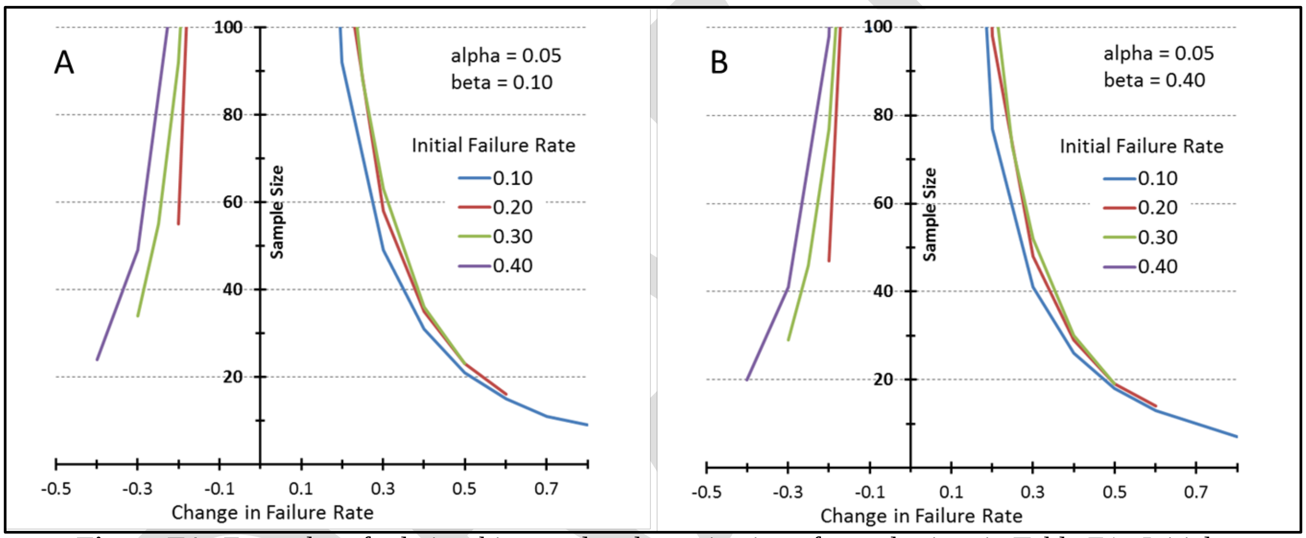


Figure E1. Examples of relationships used to determination of sample sizes in Table E1. Initial failure rate $\pi_0 = 0.3$ was used to determine sample size.

Table E1. Sample Size for different values of β, the probability of Type II error, and effect size. For these values $\alpha = 0.05$ and $\pi_0 = 0.3$. The proposed minium sample size is highlighted in gray.

Effect Size	β		
	0.3	**0.2**	**0.1**
-0.25	36	43	55
-0.2	59	72	92
-0.15	108	134	174
-0.1	251	313	412

Hypothesis tests for proportions could be applied to estimated rates of implementation failure (i.e., percent or proportion of all culverts that fail) over the years of monitoring. Implementation failure rates will be calculated for both permittors and permittess. We assume that failure rates in different years are statistically independent. Because failure rates may either increase or decrease the tests are two-sided. The following is a subset of the statistics and hypothesis tests we can perform with the data to be collected:

Test 1: $F_1 = F_t$, where
F_1 and F_t are estimated rates of implementation failure in year 1 and in subsequent years, respectively

Test 2: $F_{1Permittor} = F_{tPermittor}$, where
$F_{1Permittor}$ and $F_{tPermittor}$ are estimated rates of permittor failure in year 1 and in subsequent years, respectively

Test 3: $F_{1Permittee} = F_{tPermittee}$, where
$F_{1permittee}$ and $F_{tPermittee}$ are estimated rates of permittee failure in year 1 and in subsequent years, respectively

Test 4: $F_{1PermitWidth} = F_{tPermitWidth}$, where
$F_{1PermitWidth}$ and $F_{tPermitWidth}$ are estimated rate at which permits fail to have the correct culvert width in year 1 and in subsequent years, respectively

Test 5: $F_{1CulvertWidth} = F_{tCulvertWidth}$, where
$F_{1CulvertWidth}$ and $F_{tCulvertWidth}$ are estimated rates at which culverts fail to be constructed with the same width specified on the permit in year 1 and in subsequent years, respectively