WASHINGTON STATE DEPARTMENT OF
Natural Resources

PETER GOLDMARK
Washington State Commissioner of Public Lands

MEMORANDUM

July 18, 2011

TO: Forest Practices Board

FROM: Jim Hotvedt, Adaptive Management Program Administrator

SUBJECT: Policy recommendations on actions resulting from completion of the roads sub-basin scale effectiveness monitoring report and amphibian genetics report

CMER recently completed two reports:

- Washington Road Sub-Basin Scale Effectiveness Monitoring First Sampling Event (2006-2008) Report and
- Type N Experimental Buffer Treatment Study: Baseline Measures of Genetic Diversity and Gene Flow of Three Stream-Associated Amphibians.

Upon completion of final reports, Forests and Fish Policy is expected to make a recommendation to the Forest Practices Board regarding formal petition for rulemaking or a non-rulemaking alternative action. Policy may also recommend that the Board take no action. Both reports are the first in a series of two or more sampling efforts to inform their underlying key questions. Policy recommends no action be taken at this time based on the results of the two studies.

Roads sub-basin study

The Road Sub-Basin Scale Effectiveness Monitoring (Road Sub-Basin) Study was designed to determine the degree to which forest roads under the jurisdiction of Washington forest practices rules are meeting the *road surface erosion* and *hydrologic connectivity* resource objectives, in other words, to determine

- If road characteristics that affect runoff and sediment delivery to streams are improving through time (emphasis on trends) and
- The extent to which roads meet the FFR performance targets.

The road sub-basin study report summarizes results of the first sample event planned for the Road Sub-Basin Scale Effectiveness Monitoring Program, and provides the first look at the current status of forest roads. As I reported to you at the February 2011 meeting, a high percentage of roads in the sample units across the state were reported by local managers to either have RMAP work complete, or already be up to current road rule standards, with over half of the sample units reported to have at least 85 percent of road length meeting standards. An average of 11 percent of the road network was hydrologically connected and assumed to deliver water and sediment to streams or wetlands. Across the state, 62 percent of the sample units met the hydrology performance target (miles of delivering road/miles of stream) and 88 percent of the

1111 WASHINGTON ST SE • PO BOX 47041 • OLYMPIA, WA 98504-7041
TEL: (360) 902-1250 • FAX: (360) 902-1780 • TTY: (360) 902-1125
Equal Opportunity / Affirmative Action Employer

units met the sediment performance target (tons of delivered sediment/year/miles of stream). A statistically significant decreasing relationship was found between sediment delivery in the sample unit and percent of roads in the unit reported to be up to standards. These findings suggest that many roads show decreasing sediment inputs as they are brought up to standards.

Trends in road performance cannot be evaluated directly until after future re-visits to the sample units. Sample units are planned to be re-surveyed two more times through the RMAP performance period to determine if road conditions are either meeting targets or improving with respect to surface erosion and hydrologic connectivity. The roads sub-basin report contains a recommendation that the monitoring project continue with the next planned monitoring period (planned interval of 5 years) to evaluate the expectation that, in general, road conditions improve and that they meet the performance targets as they are brought up to forest practices standards.

The Forests and Fish Policy committee recommends that no action be taken at this time based on the results of the study.

Amphibian genetics study

Data on amphibians from Type N systems suggest that stream-associated amphibians are ecologically important in headwater habitats, but how they respond to current forestry practices is essentially unknown. One of the major objectives of the Westside Type N Experimental Buffer Treatment (Hard Rock) Study is to investigate how different forest buffers surrounding non-fish-bearing streams may influence stream-associated amphibian populations. The amphibian genetics study will estimate potential changes in genetic structure associated with applications of the forests and fish HCP patch buffer compared to no riparian buffer, a complete riparian buffer, and an unharvested reference.

The recently completed report describes baseline, pre-treatment measures of genetic structure for three species of stream-associated amphibians: Coastal tailed frog (*Ascaphus truei*), Cope's giant salamander (*Dicamptodon copei*) and Coastal giant salamander (*Dicamptodon tenebrosus*). Four main objectives were addressed:

1) Classify giant salamander individuals to the correct species and identify any hybrids;
2) Estimate indices of genetic diversity for the three species at each sampled basin;
3) Test for significant differences by region, block and anticipated treatment for any genetic diversity parameter; and
4) Determine genetic clusters for each species.

Measures of genetic diversity and genetic differentiation within and among amphibian populations can allow for species identification, provide insight into trends in population size and identify the level of migration among sites. The genetics data will also help distinguish between giant salamander species – Cope's and Coastal giant salamanders – during density sampling to ensure proper determination of species composition and proper calculation of species richness and density metrics for each species. Genetic data will also allow identifying hybrid individuals, as well as their frequency of occurrence and distribution.

A summary of findings from this baseline report include:

- Giant salamander individuals were identified to species with high probabilities. Thirty-one (31) hybrids were identified out of 1504 total sampled *Dicamptodon* individuals. However, no hybrid individuals represented the current F_1 generation, suggesting hybridization is not currently ongoing.
- Levels of genetic diversity were high in Coastal tailed frogs and intermediate for Cope's giant salamander and Coastal giant salamander.
- Fourteen sites (out of 47 possible) had high levels of inbreeding and/or evidence of declines, but no consistent patterns were found among the three species.
- Effective population sizes (number of breeding individuals) varied greatly among species, with very large sizes for Coastal tailed frogs (3000-7000), intermediate sizes for Cope's giant salamander (150-3150) and low sizes for Coastal giant salamanders (50-750).
- Generally, no differences existed by region or block, with the notable exception of Cope's giant salamander, which had lower values of allelic richness, heterozygosity, inbreeding coefficient and effective population size in the Olympics.
- Finally, geographically large genetic clusters (at a regional scale) were detected for Coastal tailed frogs and Coastal giant salamanders, consistent with their broad species range and increased ability to disperse terrestrially. In contrast, Cope's giant salamander had geographically restricted genetic clusters, whereby individual sites often were genetically distinct from nearby sites. However, degree of divergence varied by region, with a larger cluster in the Olympics and the most restricted gene flow in the South Cascades.

This report does not present how amphibian genetics and its relationship to amphibian populations and viability are affected by implementation of the above mentioned harvest treatments. To inform this, post-harvest data will need to be collected 7-8 years after harvest, allowing approximately one generational turnover. Nevertheless, the study does provide the critical pre-harvest baseline that can be compared to post-harvest genetic sampling, which will produce a substantial gain in the understanding of how Type N buffers effect amphibian population sizes and genetic variability.

The Forests and Fish Policy committee recommends no action be taken at this time based on the results of the study.

Attachments (4)

Washington Road Sub-Basin Scale Effectiveness Monitoring First Sampling Event (2006-2008) Report

Kathy Dubé, Alice Shelly, Jenelle Black, Karen Kuzis

WASHINGTON STATE DEPARTMENT OF
Natural Resources
Peter Goldmark - Commissioner of Public Lands

Cooperative Monitoring
Evaluation & Research

September 2010

CMER 08-801

Washington State Forest Practices Adaptive Management Program
Cooperative Monitoring, Evaluation, and Research Committee (CMER)
Report

Washington Road Sub-Basin Scale Effectiveness Monitoring First Sampling Event (2006-2008) Report

Prepared by:

Kathy Dubé, Alice Shelly, Jenelle Black, Karen Kuzis

Prepared for the
Cooperative Monitoring, Evaluation, and Research Committee (CMER)
of the

Washington State Forest Practices Board
Adaptive Management Program
Washington State Department of Natural Resources
Olympia, Washington

May 2010 **CMER 08-801**

Forest Practices Adaptive Management Program

The Washington Forest Practices Board (FPB) adopted an adaptive management program in concurrence with the Forests & Fish Report (FFR) and subsequent legislation. The purpose of this program is to:

> Provide science-based recommendations and technical information to assist the board in determining if and when it is necessary or advisable to adjust rules and guidance for aquatic resources to achieve resource goals and objectives. (Forest Practices Rules, WAC 222-12-045)

To provide the science needed to support adaptive management, the FPB made the Cooperative Monitoring, Evaluation and Research Committee (CMER) a participant in the program. The FPB empowered CMER to conduct research, effectiveness monitoring, and validation monitoring in accordance with guidelines recommended in the FFR.

Disclaimer

This technical report contains scientific information from a research study designed to evaluate the effectiveness of the Forest Practices Rules in achieving one or more of the Forests & Fish performance goals, resource objectives, and/or performance targets. The document was prepared for CMER and was intended to inform and support the Forest and Fish Adaptive Management Program. The project is part of the Roads Rule Group Road Sub-Basin Scale Effectiveness Monitoring Program, and was conducted under the oversight of the Uplands Processes Scientific Advisory Group (UPSAG).

This document has been reviewed by CMER and has been assessed through the Adaptive Management Program's independent scientific peer review process. CMER has approved this document for distribution as an official CMER document. As a CMER document, CMER is in consensus on the scientific merit of the document. However, any conclusions, interpretations, or recommendations contained within this document are those of the authors and may not reflect the views of all CMER members.

Proprietary Statement

This work was developed with public funding including contracts: PSC 05-386 and PSC 08-73. As such it is within the public use domain. However, the concept of this work originated with the Washington Department of Natural Resources (WDNR) and the authors; permission must be obtained from the originators to use the results in the way developed herein. Use of results without permission of WDNR and the authors may be deemed a violation of federal statures under preview of the Office of Research Integrity. As a public resource document, this work should be given proper attribution and be properly cited.

Full Reference

Dubé, K., A. Shelly, J. Black, and K. Kuzis. 2010. Washington road sub-basin scale effectiveness monitoring first sampling event (2006-2008) report. Cooperative Monitoring, Evaluation and Research Report CMER 08-801. Washington Department of Natural Resources. Olympia, Washington.

This report was prepared under contact to the Washington Department of Natural Resources as part of Contacts No. PSC-05-386 and PSC-08-73.

Author Contact Information

Kathy Dubé
Watershed Professionals Network
kdube@watershednet.com

Alice Shelly
TerraStat Consulting Group
alice@blarg.net

Jenelle Black
Northwest Indian Fisheries Commission
jblack@nwifc.org

Karen Kuzis
ENVIRON International Corporation
kkuzis@Environcorp.com

Acknowledgements

Many thanks to the members of the UPSAG committee and all others who reviewed various work products throughout both phases of this project. Your comments and thoughts are appreciated. Reviewers included Julie Dieu, Mark Hicks, Doug Martin, Dick Miller, Bob Palmquist, Nancy Sturhan, Laura Vaugeois, Curt Veldhuisen, and the Independent Science Panel Review committee members. Jenelle Black performed the site selection process. Thanks also to the field crews who drove and walked all those road miles: Scott Hotchkiss, Pete Malinak, Matt Rourke, Jesse Saunders, Dan Thomas, and Brett Winterowd, and to Ted Hitzroth for GIS support.

ABSTRACT

New Forest Practice Rules for forest roads under Washington Department of Natural Resources (WDNR) jurisdiction were adopted in Washington State in 2001. Implementation of the road rules for existing roads is planned over a 15-year time horizon through the completion of Road Maintenance and Abandonment Plans (RMAP). The objectives of the Road Sub-Basin Scale Effectiveness Monitoring Program are:

- to determine if the road characteristics that affect runoff and sediment delivery to streams are improving through time as RMAP are implemented between 2001 and 2016; and
- to determine the extent to which roads on lands subject to WDNR forest practice rules meet the FFR performance targets.

Characteristics of forest roads in a total of 60 random four-square-mile sample units across the state were inventoried between 2006 and 2008, five to seven years after the forest practice rules were adopted. This is the first sample event planned for the Road Sub-Basin Scale Effectiveness Monitoring Program, and provides the first look at the current status of forest roads.

A high percentage of roads in the sample units across the state were reported to either have RMAP work complete, or already be up to current road rule standards, with over half of the sample units reported to have at least 85 percent of road length meeting standards. An average of 11 percent of the road network was hydrologically connected and assumed to deliver water and sediment to streams or wetlands. Across the state, 62 percent of the sample units met the Forests & Fish Report (FFR) hydrology performance target (miles of delivering road/miles of stream) and 88 percent of the units met the FFR sediment performance target (tons of delivered sediment/year/miles of stream).

Statewide, no relationship was found between length of road delivering in a sample unit and percent of roads in the unit reported to be up to standards; however, a statistically significant decreasing relationship was found between sediment delivery in the sample unit and percent of roads in the unit reported to be up to standards. These findings suggest that, while there may be locations where there is higher hydrologic connectivity or sediment input from roads, many roads show decreasing sediment inputs as they are brought up to standards.

Approximately ninety-five percent of the land sampled was owned by large industrial forest owners and state/local governments. The study was intended and designed to incorporate land owned by small forest landowners, but due to the sample site criteria and layout and their fragmented ownership patterns, little of that ownership type was able to be incorporated. Development of a companion study specifically designed to obtain access to and evaluate small forest ownerships could be pursued in order to characterize those lands.

It is recommended that the monitoring project continue with next planned monitoring period (planned interval of 5 years) to evaluate the expectation that, in general, roads conditions improve and they meet the performance metrics as they are brought up to FFR standards.

WASHINGTON ROAD SUB-BASIN SCALE EFFECTIVENESS MONITORING FIRST SAMPLING EVENT REPORT (2006-2008)

1 Introduction 1
1.1 Literature Summary 1
2 Monitoring Objectives 2
3 Methods 5
3.1 Project Coordination and Meetings 5
3.2 Field Protocol Development 6
3.3 Quality Assurance/Quality Control Program 7
3.4 Sample Units 7
3.4.1 Sampling Frame 7
3.4.2 Sample Size and Unit Area 8
3.4.3 Site Selection 10
3.5 Landowner Contacts 12
3.6 Field Sampling 12
3.7 Database Development 16
3.8 Calculation of Monitoring Measures 16
3.8.1 Statistical Analysis 19
4 Results and Discussion 20
4.1 Crew Variability Test 20
4.2 Sample Units 22
4.2.1 Sample Unit Selection and Rejection 24
4.2.2 Sample Unit Results 25
4.3 Monitoring Questions and Measures 30
4.3.1 Monitoring Question 1 33
4.3.2 Monitoring Question 2 45
4.3.3 Monitoring Question 3 48
4.3.4 Monitoring Question 4 52
4.3.5 Monitoring Question 5 54
4.3.6 Monitoring Question 6 54
4.4 New Roads Evaluation 59
5 Summary and Recommendations for Future Sample Events 60
6 References 62

LIST OF FIGURES

Figure 1(a) and 1(b). Change in the Mean Power and P-value for (a) RSED and (b) RLEN as the Area of Each Sample Unit from Phase I was Increased. 9
Figure 2. Sample Unit/Quarter Section Selection Process (from Raines et al. 2005) 10
Figure 3. Example Road Segments 13
Figure 4. Length Delivering and Sediment Production by Surveyor for 2006 and 2008 Tests. 21

Figure 5. Location of Road Monitoring Sample Units. 23
Figure 6. Un-weighted Road Length Delivering in Sample Unit vs. Road Density.....28
Figure 7. Estimated Sediment Delivered to Streams in Sample Unit vs. Road Density.....28
Figure 8. Un-weighted vs. Weighted Road Length Delivering for the Sample Units. 29
Figure 9. Length of Road by Sediment Yield Category in Each Sample Unit Showing Overall Variability Among Sites. 30
Figure 10. MQ1.1 Total Road Length Delivering to Streams per Unit Area. 33
Figure 11. MQ1.2 Percent of Road Network Draining to Streams.....34
Figure 12. MQ1.3 Percent of Delivering Road Network by Surfacing. 36
Figure 13. MQ1.4 Percent of Road Network by Traffic Level. 38
Figure 14. MQ1.5 Percent of Road Network by Cutslope Cover Category.....40
Figure 15. MQ1.6 Percent of Drainage Points by Connectivity Class.....42
Figure 16. MQ1.7 Percent of Delivering Road Length by Rutting Category. 44
Figure 17. H2b1 Road Surfacing Index.....45
Figure 18. H2b2 Road Traffic Index.....46
Figure 19. H2b3 Cutslope Cover Index.....47
Figure 20. H2b4 Miles of Delivering Road with Ruts Interfering with Drainage.....48
Figure 21. MQ3.1 Miles of Delivering Road per Mile of Stream (FFR Hydrology Performance Measure).....49
Figure 22. MQ3.1 Cumulative Percent of Sample Units, Miles of Delivering Road per Mile of Stream (RLEN).....50
Figure 23. MQ3.1 Sediment Delivered/Year per Mile of Stream (FFR Sediment Performance Measure).....51
Figure 24. MQ3.1 Cumulative Percent of Sample Units, Tons of Delivered Sediment per Mile of Stream (RSED).....52
Figure 25. MQ4.1 Miles of Delivering Road per Mile of Stream Divided by Regional Performance Target. 53
Figure 26. MQ4.2 Sediment Delivered/Year per Mile of Stream Divided by Regional Performance Target. 54
Figure 27. Percent of Sample Units with Percent of Roads Reported to be up to FFR Road Standard 55
Figure 28. Miles of Delivering Road/Miles of Stream (MQ3.1) Expressed as Percent of Regional Target vs. Percent of Road Miles in Sample Unit Reported to be up to Standards. ..56
Figure 29. Tons of Delivered Sediment/Miles of Stream (MQ3.2) Expressed as Percent of Regional Target vs. Percent of Road Miles in Sample Unit Reported to be up to Standards.....56
Figure 30. Sample Unit Maps Showing Difference in Delivery Between Units with Ridgetop Roads (left) and Midslope/Valley Bottom Roads (right).....58

LIST OF TABLES

Table 1. FFR Sediment Performance Targets for Roads 3
Table 2. Monitoring Questions/Hypotheses and Measures.....5
Table 3. Data Collected on Delivering Road Segments.....14
Table 4. Methods Used to Compute Monitoring Measures and Statistics 16

Table 5. Potential Study Unit Rejection Reasons and Rates 24
Table 6. Summary of Data Collected by Sample Unit 26
Table 7. Mean, Median and Standard Deviation of Monitoring Measures by Region and State. 31
Table 8. Summary of New Roads by Sample Unit 60

LIST OF APPENDICES

Appendix A. Meeting Minutes
Appendix B. QAPP (available electronically on CD)
Appendix C. QA/QC Report (available electronically on CD)
Management Visit Memos
Replicate Survey Memos
Crew Variability Assessment Report
Third Party Report
Appendix D. Field Protocol (available electronically on CD)

1 Introduction

The Washington Department of Natural Resources (WDNR) implemented new Forest Practice Rules for forest roads under WDNR jurisdiction in 2001. The 2001 road rules were developed in the wake of the Forests & Fish Report (FFR) and are designed to improve many aspects of the road network by reducing the occurrence of road landslides, culvert plugging, and road surface erosion, and by eliminating barriers to fish passage. The resource objectives for roads include sub-basin scale performance targets for surface sediment and water diversion to the stream network associated with forest roads (Raines et al. 2005). The resource objectives and performance targets for road hydrologic connectivity to the stream network and sediment input to streams are linked because runoff is required to transport sediment from roads to streams. By limiting the delivery of sediment and excess water to streams, the road rules will help to protect water quality and aquatic resources. Implementation of the road rules for existing roads is planned over a 15-year time horizon through implementation of Road Maintenance and Abandonment Plans (RMAP) that concludes in 2016. After that time, all roads will be required to be maintained to forest practice rule standards.

As part of the Cooperative Monitoring, Evaluation, and Research (CMER) Committee mission, a number of different monitoring programs have been developed to evaluate the Forest Practice Rules. Monitoring projects are grouped by 'Rule Group' (roads, riparian, etc.) and include effectiveness and validation monitoring at several different scales (FY 2010 CMER WORK PLAN p.12). The Road Sub-Basin Scale Effectiveness Monitoring Project (Raines et al. 2005) was designed to determine the degree to which forest roads under the jurisdiction of Washington Forest Practices Rules are meeting the road surface erosion and hydrologic connectivity resource objectives. This monitoring project has been set up to document current road characteristics related to runoff and surface erosion from roads within randomly-selected four-square-mile sample units of forest land. Sample units are planned to be re-surveyed three times through 2016 to determine if road conditions are either meeting targets or improving with respect to surface erosion and hydrologic connectivity. This is a voluntary study relying on the cooperation of landowners and does not evaluate landowner compliance with the Forest Practices Rules.

This report describes the first data collection effort, conducted from 2006 to 2008. It includes the field protocols, quality assurance plan, and analysis of data collected during the first sampling effort. Trends in road performance can not be evaluated until after future revisits to these (and/or other) sample units.

1.1 Literature Summary

There is a large network of unpaved forest roads in Washington State that provides access to private and state-owned timberlands. While these roads provide many positive benefits, they can also be sources of runoff and sediment that have the potential to affect aquatic habitat and water quality where the roads are hydrologically connected to streams or wetlands. Because the documentation for the Washington Road Surface Erosion Manual (Dubé et. al 2004, Appendix A) provides a detailed narrative of past research, the following discussion is limited to a brief summary.

The characteristics of a road and where it is constructed on the landscape control how much sediment is produced and whether or not this sediment is delivered to a waterbody. Road characteristics that have a major influence on erosion rates include:

- Road segment length – longer lengths have more erosion (Mills et al. 2003)
- Road gradient – steeper roads have more erosion (Luce and Black 1999)
- Surfacing type and durability – surfacing with durable gravel or vegetation reduces erosion (Meyers 2007, Coe 2006, Mills et al. 2003, Foltz 1996, Burroughs and King 1989, Kochenderfer and Helvey 1987, Swift 1984, Reid and Dunne 1984)
- Traffic use – higher use, particularly during precipitation events, produces more erosion (Mills et al. 2003, Foltz 1996, Reid and Dunne 1984, Sullivan and Duncan 1981)
- Grading and rut development – rutting produces more erosion; disturbance from grading temporarily increases erosion (Meyers 2007, Sugden and Woods 2007, Foltz and Burroughs 1990, Burroughs and King 1989)
- Interception of cutbank sub-surface flow by the road ditch – more interception results in ditch erosion (MacDonald et al. 2001)

Factors that influence the delivery of sediment to streams include:

- Distance between runoff point and waterbody – shorter distances result in more delivery (Brake et al. 1997, Megahan and Ketcheson 1996, Swift 1985, Trimble and Sartz 1957)
- Hillslope gradient and number of obstructions between road and waterbody – steeper gradient hillside and/or fewer obstructions results in more delivery (Brake et al. 1997, Megahan and Ketcheson 1996)
- Volume of flow and erosion from road segment – higher runoff and/or erosion volumes results in delivery at greater distances (Ketcheson and Megahan 1996)

The present monitoring project collected information on all the road conditions that have a major influence on erosion as well as indicators of hydrologic connectivity between the road and a stream or wetland.

2 Monitoring Objectives

The Road Sub-Basin Scale Effectiveness Monitoring Project was developed by members of CMER to determine if road characteristics that affect runoff and sediment delivery to streams are improving through time and the extent to which roads meet the FFR performance targets (Table 1). Details of the monitoring design are in Raines et al. (2005).

Table 1. FFR Sediment Performance Targets for Roads

Measure	Performance Target	
	New Roads	**Existing Roads**
Road sediment delivered to streams	Virtually none	
RLEN - Ratio of road length delivering to streams/total stream length (mile/mile)		Not to exceed: East of Crest 0.08-0.12 Coast (Spruce) 0.15-0.25 West of Crest 0.15-0.25
RSED - Ratio of road sediment production delivered to streams/total stream length (tons/yr/mile)		Not to exceed: East of Crest 1-3 Coast (Spruce) 6-10 West of Crest 2-6

(Source: Forests & Fish Report, Schedule L-1, June 2000)

The performance targets for road hydrologic connectivity to streams (RLEN) and surface sediment delivery (RSED) were developed in 2000 to supplement qualitative road standards identified in the Forests & Fish Report (Table 1). Target values were derived from sediment production estimates for forest road networks across Washington inventoried as part of Watershed Analyses done during the 1990s. Watershed analysts compiled road sediment delivery within sub-basins of similar scale to this study, each of which was given an aquatic hazard rating of Low, Moderate or High. The ranges bracketed by the RLEN and RSED targets correspond to sub-basins rated Low hazard and the lower values rated Moderate hazard. The group chose to document the targets by a 'range' rather than a single value because they felt the scientific information available to quantify aquatic sensitivity to sediment was insufficient to support a single threshold value. Separate targets were developed for the western and eastern sides of the Cascades and the coastal zone to reflect differing precipitation rates and channel densities. Although sub-basins where road metrics are within the local target range are expected to have acceptably low impact levels, the target development group recommended initiating further sediment-sensitivity studies to evaluate and possibly refine the target values. Readers should also be aware that although road evaluation methods for this monitoring study are similar to those used for Watershed Analysis, there have been numerous minor changes to methodologies since target development that have changed road sediment estimates.

The FFR performance target for new roads, virtually no sediment delivered to streams (Table 1), was not specifically addressed in the monitoring design for this study. For all the metrics developed as part of the monitoring design, existing and new roads were evaluated together and compared to the regional FFR targets for existing roads (Table 1). A small number of new roads were present within some of the sample units. The lengths of hydrologically connected road and estimated road sediment delivered to streams were compiled and are reported separately for the new roads in Section 4.4.

Six monitoring questions were developed to provide feedback on how well improvements in road characteristics translate into changes in the target measures for sediment and hydrology at the sub-basin scale:

- Monitoring Question 1: What is the condition of forest roads at each sample event, specifically those attributes management can change relative to sediment production and delivery?

- Monitoring Question 2: Have road attributes that affect sediment production and delivery improved over time?
    - Hypothesis 2a: No reduction in road drainage connectivity to streams has occurred since the previous sampling event(s).
    - Hypothesis 2b: No improvement in road attributes that affect sediment production and delivery has occurred since the previous sampling event(s).

- Monitoring Question 3: What is the status of road performance measures for drainage connectivity and sediment delivery to streams at each sample event?

- Monitoring Question 4: What is the status of road performance measures relative to their targets, by performance target region, at each sample event?

- Monitoring Question 5: Have measures of road sediment performance improved over time?
    - Hypothesis 5a: No reduction in the road drainage connectivity performance measure has occurred since the previous sampling event(s).
    - Hypothesis 5b: No reduction in the road sediment delivery performance measure has occurred since the previous sampling event(s).

- Monitoring Question 6: Will roads judged to meet FFR road standards meet the performance targets?
    - Hypothesis 6a: There is no direct relationship between the percentage of the road system that is judged to meet road standards and the reported road drainage connectivity performance measures.
    - Hypothesis 6b: There is no direct relationship between the percentage of the road system that is judged to meet road standards and the reported road sediment delivery performance measures.

In order to assess these monitoring questions and hypotheses, a series of numerical monitoring measures was developed that include the primary road characteristics influencing the delivery of sediment and water from a road system to aquatic environments (Table 2). These monitoring measures will be computed and compared for each sample event to determine the trend in road conditions through time.

Table 2. Monitoring Questions/Hypotheses and Measures

Monitoring Questions or Hypothesis	Reported Monitoring Measures
Monitoring Question 1	1. Total road length draining to streams (road miles/sq mi) 2. Percent of road network draining to streams 3. Percent of road in each surface category 4. Percent of road in each traffic category 5. Percent of road in each cutslope cover category 6. Percent of drainage points by connectivity class 7. Percent of road in each road rutting category
Hypothesis 2a	1. Total road length draining to streams (road miles/sq mi) 2. Percent of road network draining to streams
Hypothesis 2b	1. Road surfacing index 2. Road traffic index 3. Cutslope cover index 4. Miles of delivering road with ruts interfering with drainage
Monitoring Question 3	1. Miles of forest road delivering to streams per miles of stream (road hydrology performance measure) 2. WARSEM modeled tons of road sediment delivered to streams per miles of stream per year (sediment performance measure)
Monitoring Question 4	1. Miles of forest road delivering to streams per miles of stream (road hydrology performance measure) divided by the performance target by target region 2.WARSEM modeled tons of road sediment delivered to streams per miles of stream per year (sediment performance measure) divided by the performance target by target region
Hypothesis 5a	Miles of forest road delivering to streams per miles of stream (road hydrology performance measure)
Hypothesis 5b	WARSEM modeled tons of road sediment delivered to streams per miles of stream per year (sediment performance measure)
Hypothesis 6a	Miles of forest road delivering to streams per miles of stream by percent of road length meeting FFR road standards
Hypothesis 6b	WARSEM modeled tons of road sediment delivered to streams per miles of stream per year by the percent of road length meeting FFR road standards

(Source: Raines et al. 2005)

3 Methods

The first sample event included two phases of work. The first phase (Phase I) included development of a detailed field protocol and QA/QC plan, training of the inventory crew, and sampling 14 blocks of land in 2006. The results of the Phase I samples were used to conduct a statistical power analysis to estimate the total sample size required for the first sample event, which resulted in a decision to use a sample size of 60. The second phase of the sample event was conducted in 2008 and included updating the field protocol, re-training, and sampling the remaining 46 sample units.

3.1 Project Coordination and Meetings

Several meetings between the contractor and CMER personnel were scheduled as part of the Phase I and Phase II work. Meeting minutes are included in Appendix A.

The Phase I Start-Up Meeting was held on December 14, 2005, to review the monitoring objectives and methods. The Phase I Second Unit Meeting was held on March 31, 2006, with

the field crew to discuss any changes needed to the field protocol after sampling the first two sample sites. Attendance by project principal investigator Kathy Dubé at the October 24, 2006 CMER meeting satisfied the Phase I completion meeting and is documented in the CMER meeting minutes.

A kickoff meeting for Phase II was held on March 21, 2008, to review the Phase II objectives and methodology with UPSAG committee members. The Phase II Second Unit Meeting was held on May 23, 2008, following sampling of two Phase II units. The purpose of the Second Unit Meeting was to review data collection procedures and discuss any issues that came up with the survey crews. The Phase II Survey Completion Meeting was held on January 29 and continued on February 5, 2009 to discuss the sampling results, any issues that arose during the project, suggestions from the field crew for future sample events, and data analysis required to meet project goals.

3.2 Field Protocol Development

A standardized field protocol manual was developed to provide consistent methods for collecting road condition data (Appendix D). The Field Protocol was developed based on guidelines in the WARSEM manual (Dubé et al. 2004) and the information required in the monitoring design (Raines et al. 2005).

Several revisions to the 2006 field protocol document were made prior to the 2008 sampling based on suggestions by the field crews and to incorporate UPSAG requests. The following key changes were made and implemented during the 2008 sampling:

- Tread configuration (the percent of the road tread that drained to a particular drainage point being inventoried) was changed from three categories in 2006 [full (100%), half (50%), none (0%)] to five categories in 2008 (0%, 25%, 50%, 75%, 100%) at the suggestion of the field crews to better reflect conditions actually seen on the ground and to improve consistency between surveyors who were having a difficult time determining tread configuration in some cases.

- Delivery of ditch water across long, vegetated fillslopes was clarified to help improve consistency. If there was less than 50 feet of vegetated fillslope between the ditch outlet and the stream, 100% delivery is assumed. If there is more than 50 feet of vegetated fillslope between the ditch outlet and the stream, the delivery flowchart is used to determine delivery.

- Addition of a fifth delivery category to document road segments that closely paralleled streams or obvious and mapped wetlands but did not fall within another delivery category. This category was called “Parallel within 20 feet of a stream.”

- Addition of a data field to document the length of road that is parallel within 20 feet of a stream. This includes cases where a segment may deliver directly, but the road closely parallels a stream and may be constructed in or occupying the floodplain.

- Updating the flowchart that provides guidance for delivery and breaking roads into segments to reflect changes to delivery categories.

The final 2008 field protocol is included on the data CD.

3.3 Quality Assurance/Quality Control Program

A Quality Assurance Project Plan (QAPP) was developed and implemented during the initial 2006 field season, and updated during the Phase II field season in 2008. The main elements of the QAPP include:

1. Standard field protocols
2. Data management procedures
3. Field crew training program
4. Quality controls during field data collection
5. Monthly field visits by management personnel
6. Duplicate crew surveys
7. Quantification of crew variability
8. Third party QA analysis
9. QA/QC reporting

Field inventory personnel were trained in the field protocol during three-day sessions prior to both the 2006 and 2008 field seasons; these included PowerPoint presentations and field training. In addition, the Principal Investigator and contractor project management personnel worked in the field with crew members during the course of the project to answer questions, review protocols, and discuss potential revisions and improvements to the protocol ("Field Assistance Visits"). Crew members also worked together for several days during the sampling season to discuss how they were making determinations in the field and improve the consistency of data collected ("Duplicate Crew Surveys").

An assessment of crew variability was conducted during both the Phase I and Phase II field seasons to: 1) understand where field assistance and additional training were needed to minimize differences between observers; and 2) help to understand, if possible, how much of the total variance in monitoring results may be due to observer error. Three road sections were selected from a West of Crest sample unit in 2006 (Phase I) and three different road sections were measured in an East of Crest unit in 2008 (Phase II). The test lengths were chosen to represent the range of road use and maintenance conditions within each sample unit. The variability tests were conducted after initial training and again near the end of data collection. Each of the road sections was evaluated by each of the crew members using standard project data collection instructions and techniques.

A final version of the 2008 QAPP includes details of each of these elements and is included in Appendix B of this report. Appendix C is the QA/QC report for the project.

3.4 Sample Units

3.4.1 Sampling Frame

Sample units were selected in a stratified random manner from all forest practices rules-regulated (FFR) timberlands within the state. The number of units in each geographical performance target region was based on percent of FFR land area within each region:

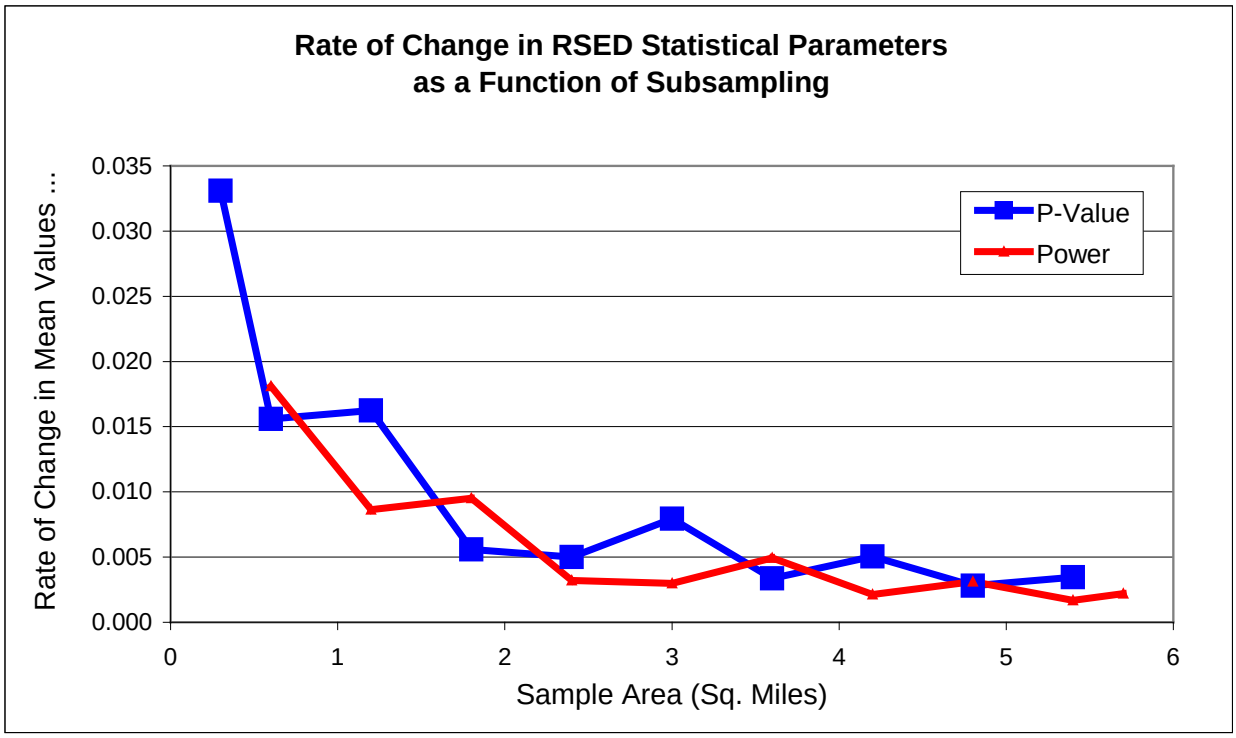
- Coastal Spruce = 11%
- East of Crest = 51%
- West of Crest = 38%

FFR land was estimated from the "CMERlands" GIS coverage developed by the Department of Natural Resources Forest Practices Division (Washington Department of Natural Resources 2005). The CMERlands coverage was developed from USGS "forested" polygon coverage (assessed from LandSAT imagery), Habitat Conservation Plan (HCP) areas, public land ownership coverages, and Native American land coverages. FFR lands were selected as those forestlands that were neither Federal, Tribal, nor covered by an HCP *plus* DNR HCP lands that were not within the habitat range of spotted owls or marbled murrelets and so are managed under normal forest practices rules (these occurred in eastern Washington). The "forested" designation was inclusive and often included land that was actually scrub-shrub steppe or rangeland as well as some agricultural land. These non-forest areas were further filtered out from individual sites during the site screening process. Property from small forest landowners was specifically included in this study.

3.4.2 Sample Size and Unit Area

This study is designed to evaluate trend, which entails evaluating differences in results from sample events spaced 5 years apart. The study is designed to resample the same units in each sampling event and to use paired t-test analyses to assess change (Raines et al. 2005). In order to estimate the number of samples required to obtain good statistical power for a paired t-test, a measure of variability among differences (e.g., future year metric minus current year metric for each sample unit) would be required. Since such an estimate will not be available until multiple years of data are collected, the power for a simple two-sample t-test was estimated instead. When measurements are correlated (i.e., same sites through time), the paired t-test has higher power than the two-sample t-test, so the estimated sample size should be more than adequate. Based on variability among parameters found during earlier watershed analyses, the study plan originally estimated a sample size of 60 units. With 60 units at each sampling time, the power was greater than 80% for detecting a change of 30% in RLEN and a change of 50% in RSED using a two-sample t-test. Thus, with a paired t-test for a paired sample, the power is expected to be even greater. Data from the 14 units sampled during Phase I were analyzed to refine this original estimate.

During Phase I, six-square-mile units (24 quarter sections) were used for sampling based on the average subbasin area in Watershed Analysis data (average area = 6.26 mi^2, n=60, SD = 3.96 mi^2; Raines et al. 2005). However, obtaining sites of this size proved to be very difficult and resulted in an unacceptably high site rejection rate (see next section), especially in those areas on the fringes of large industrial forestland blocks. Sampling such large areas was also very expensive due not only to the increased land area but to the fact that accessing many landowners and road systems greatly multiplied the sampling effort and costs. As part of an effort to reduce the sample unit rejection rate (by requiring fewer approved quarter sections) and to reduce costs, CMER analyzed the 2006 field data for the two main target metrics from the 14 Phase I sites using fixed area re-sampling (Figure 1).



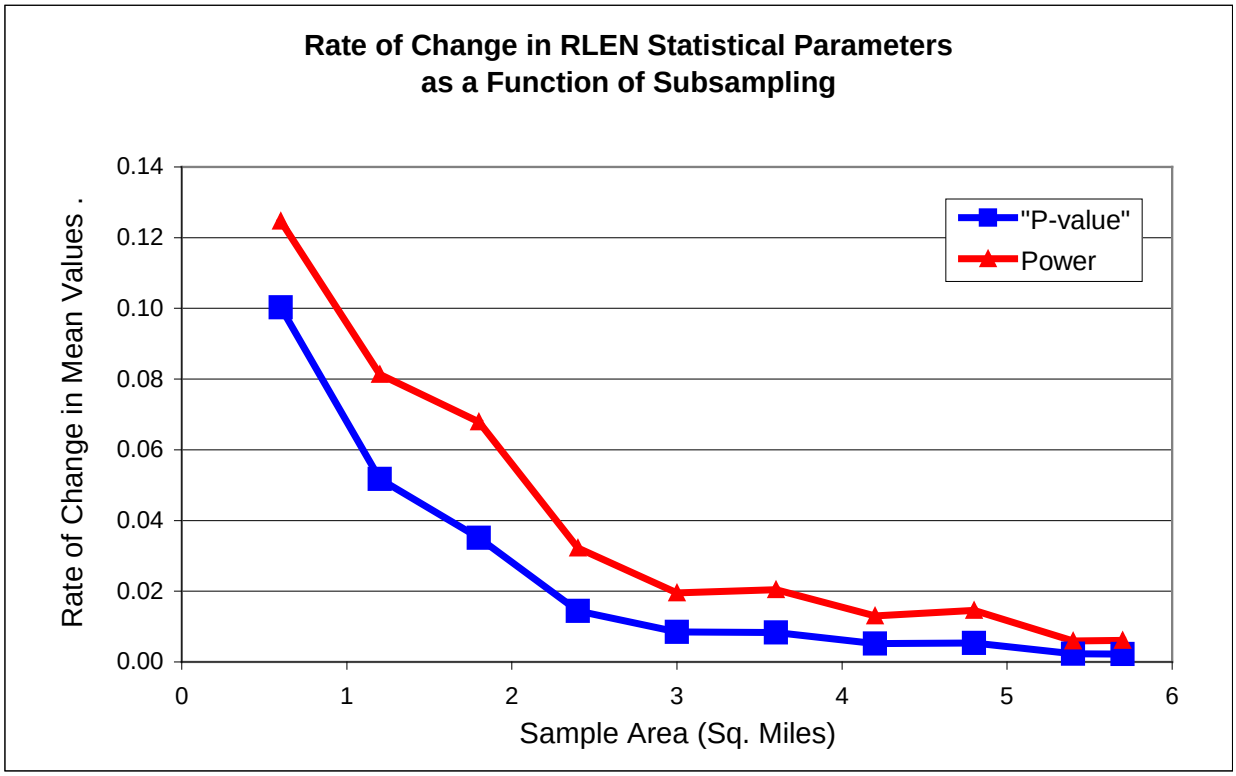


Figure 1(a) and 1(b). Change in the Mean Power and P-value for (a) RSED and (b) RLEN as the Area of Each Sample Unit from Phase I was Increased.

Based on this analysis, CMER determined that the sample size should remain at the original 60 but that the sample unit areas could be reduced to as little as three square miles and still provide

the required certainty to assess trends (Figure 1). Three square mile units would still be within the standard deviation of the average Watershed Analysis subbasin area used as the basis in the study design. Sampling four-square-mile blocks would allow for some area attrition within blocks (due to ownership or land use changes incompatible with the study) in future years without having to remove the entire block from the sample. This is a very important point given the current trend of land development, the desire to gain information about small forest land management as well as large industrial land management, and the effort required to obtain and permit sample units that meet the study criteria. Therefore, UPSAG recommended and CMER and the Forest Practices Policy Committee approved the decision to retain the original sample size of 60 but reduce the sample unit areas to four-square miles. The results of the CMER analysis suggested that reducing the sample area from six- to four-square miles would reduce power by less than 0.5% for RSED and 2% for RLEN (the two primary metrics being tracked) while reducing the costs by approximately $175,000. For the final analysis combining data from Phases I and II, only the first 16 selected quarter sections from the Phase I units were used (see following section for a description of sample unit quarter section numbering) so that all units analyzed had the same area. The 240 square miles sampled represents 1.7 percent of the estimated 9.1 million acres of Forest Practice Rule land. If 60 of the original 6 square-mile units were inventoried, it would have been 2.5 percent of the total land area.

3.4.3 Site Selection

Initial unit selection followed the plan described in Raines et al. 2005 and began by randomly selecting section corners within the FFR land sample area. The potential sample unit consisted of a sixteen-square-mile block surrounding that section corner (Figure 2). Units were screened in the numerical order in which the corner was selected. As a unit was rejected, the next unit in the initial selection was evaluated until the total sample number was reached. Extra units were initially selected and screened to allow for further attrition as more refined site screening and permitting proceeded.

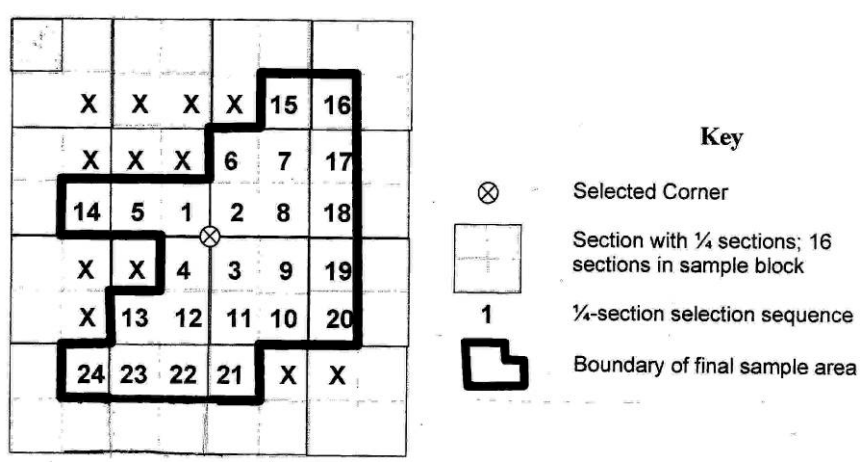


Figure 2. Sample Unit/Quarter Section Selection Process (from Raines et al. 2005).

The intent was to sample blocks of land that contained coherent networks of roads that were managed as forest roads under the forest practices rules. The first step in screening units entailed assessing the amount of FFR land within the block. If there was less than four square miles of FFR land, as designated from the GIS layers, the unit was discarded without further assessment. If there was enough FFR land, the unit was inspected to see if the FFR areas were laid out such that there was a contiguous block of FFR land of suitable size, or two large contiguous blocks that together would make up the requisite area. The objective of assessing blocks of contiguous road was the key factor. For example, in a few cases the unit spanned a large river, and large contiguous blocks on each side could be formed and were used to make up the unit. In other cases, large blocks could not be formed and the unit was rejected.

After the initial screening using the CMERlands FFR GIS layer, quarter sections within units were evaluated using aerial photos, land parcel ownership, personal knowledge, and field inspection. Typically, quarter sections in the unit that could be readily identified as non-usable were rejected in the table for that unit. Per the study design (Raines et al. 2005, Appendix A), quarter sections were discarded if they contained any unacceptable non-forest use (residence, quarry, fire station, commercial, etc.). Roads, power lines, and railroad rights of way were considered acceptable non-forest uses and quarter sections containing those were retained. UPSAG determined that properties containing small (e.g., one-room) hunting cabins would be retained in the sample since such a structure's use did not fundamentally change the use of that forestland, nor did it particularly indicate the intention to develop the land further. On the other hand, quarter sections that consisted of land parcels less than 10 acres in size were assumed to be under development or likely to be developed by the next sample event and were rejected. Agricultural, scrub-steppe range land, and large water body inclusions were assessed on a case-by-case basis. If the non-forested area made up over half of the quarter section, that quarter was discarded.

Following the initial identification and recording of ineligible quarter sections, each potential quarter section surrounding the central section corner was evaluated in more detail in a clockwise spiraling pattern out from the selected section corner (Figure 2). If the quarter section was eligible, it was added to the potential sample unit in order (quarter section 1, 2, 3, 4, etc. in Figure 2) and numbered. Weeding out the clearly ineligible quarters first simplified the detailed assessment by allowing working unit maps to be created in the GIS and quarters numbered. Using those working maps, county assessor records of land parcel boundaries and ownership provided further information on land use and ownership for sites and quarter sections that passed the photo screening. Parcel information was obtained in several ways. County assessors were contacted and most provided their parcel data to us in a database, and, where they had it, a GIS layer. When it became available, the Statewide Parcel Database (RTI 2007) was used. Internet searches were used to verify parcel database information and to obtain information for those counties that did not provide information directly. Where graphical parcel data were not available, property legal descriptions were used to identify properties within quarter sections. Since this level of data screening was very time intensive, it was left until last and only applied to those quarter sections that passed previous screening. Parcel data were inspected to ensure there were no incompatible improvements on the properties and to obtain landowner contact information.

Landowners were then contacted to determine if they were interested in participating in the monitoring project. Contact was made via letter, by telephone, and by personal contact if somebody whom we knew also knew the landowner. If the owner was interested in participating, the quarter section was retained; if any owner within a quarter denied access, that quarter section was dropped and the next quarter section in the sequence was used. Additional criteria were applied to ensure that units were reasonably blocked up and did not have short segments of unrelated roads distributed across them (see Raines et al. 2005, Appendix A for more details).

3.5 Landowner Contacts

Landowners within each sample unit were contacted to provide permission to sample the roads, and to acquire the following information about roads within the unit:

- Landowner category (large or small by WDNR standards)
- Road name/numbering
- Traffic levels over past 1 year - landowner either provided average number of loads/day or selected traffic category from those provided in the Field Protocol for each road within the sample unit on their ownership
- Percentage of roads within the sample unit that the landowner considered up to current forest practice standards (either up to standards or RMAP has been implemented)
- Recent road maintenance activities (within last 5 years)
- Age of roads

This information was used for statistical analysis and in the WARSEM analysis.

3.6 Field Sampling

Roads in the selected sample units were inventoried from March – August, 2006 during Phase I (14 units) and from May – December, 2008 during Phase II (46 units). Field crews during both phases included five crew members; four of these crew members participated in both field seasons. One Phase I crew member was not available during Phase II and was replaced by another person. Details of the field sampling methods are included in the field protocol (Watershed Professionals Network 2006 and 2008).

During the field inventory, all forest roads in the sample unit were either walked or driven to collect information on the road condition. Forest road for the purpose of this study are as defined by the Washington Administrative Code (WAC) Chapter 222-16-010. On that basis, the following were not included in the inventory:

- Foot trails
- Skid roads
- Highways/county roads
- Officially orphaned or abandoned roads [defined in WAC 222-24-053(3) and (4)]
- Vegetated roads: non-drivable vegetated roads with over 90% cover on the tread were inspected for at least 200 meters to ascertain there was no evidence of erosion. Vegetated roadways with no traffic result in little surface erosion (Swift 1984). If no

evidence of erosion was found, the remainder of the road was not surveyed. Non-drivable roads with less that 90% vegetative cover were inventoried.

Roads were divided into segments by determining the hydrologic connectivity (drainage to streams or typed wetlands) for each portion of the road network. Delivery to streams/typed wetlands was classified as follows, based on studies of road sediment delivery and sediment trap efficiency as described in Appendices A and C of the WARSEM manual (Dubé et al. 2004), Ketcheson and Megahan (1996), and NCASI (2000):

0 – No delivery – the road or drainage structure outfall drains to the forest floor, with no evidence of sediment plumes reaching a stream or wetland

1 – Direct delivery – the road, ditch, or drainage structure outfall drains directly into a stream (e.g., at a stream crossing) or wetland

2 – 35% delivery – a sediment plume from the road drainage reaches a stream or wetland, or there is a constructed sediment trap that traps some sediment (Ketcheson and Megahan 1996, NCASI 2000)

3 – 10% delivery – a sediment plume from the road drainage ends between 1-20 feet from a stream or wetland (Ketcheson and Megahan 1996)

4 – Direct via gully – drainage from the road is delivered directly to a stream or wetland through a gully or landslide scar

5 – Road parallel to and within 20 feet of a water body – a road that does not fall into another delivery category (1-4) but parallels a stream or wetland (edge of tread is within 20 feet of stream/wetland/floodplain). This delivery category was added for Phase II field work; it was not delineated for Phase I roads.

The road network was broken into delivering and non-delivering segments that were generally delineated by drainage structures (culverts, bridges, ditchouts) and grade breaks. Figure 3 shows an example of a road system broken into segments; delivering segments that either drain directly to streams (e.g., segments 1, 2, 3, 5) or are in the 35% or 10% delivery categories (e.g., segments 4 and 6), or are closely parallel to a stream (e.g., segment 7) are numbered in the diagram.

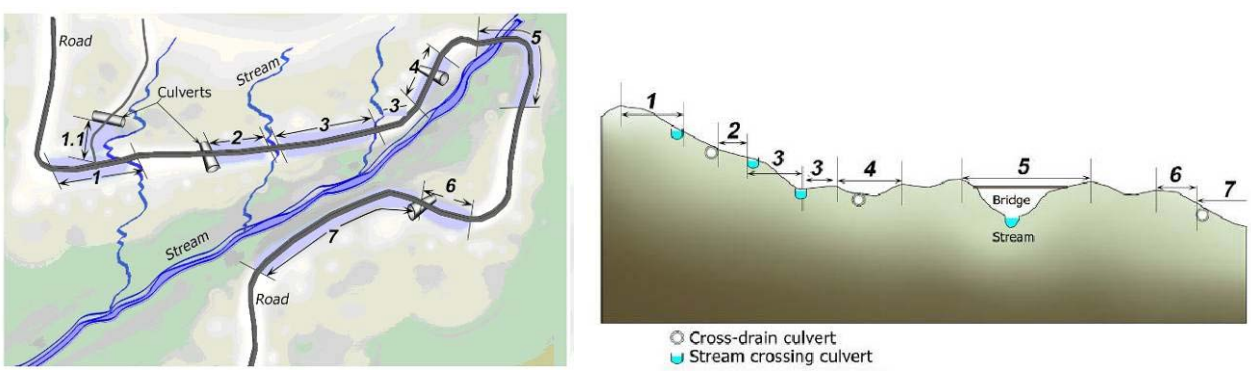


Figure 3. Example Road Segments
Map view (on left) and road profile (on right) show how roads were divided into delivering segments (numbered) and non-delivering segments (un-numbered).

For each delivering road segment, a GPS location was collected, and information on the characteristics of the road tread, cutslope, and ditch were collected and entered into the GPS (Table 3). A JAMAR RAC Plus I odometer was used in each vehicle to track stationing in feet (a string box was used for roads inventoried on foot). The location of each non-delivering culvert (e.g., ditch relief pipes) was also collected. No information on road characteristics was collected for non-delivering road segments since these segments do not contribute sediment to surface waters (Raines et al. 2005).

Table 3. Data Collected on Delivering Road Segments

Attribute	Possible Values	How Measured or Determined
Site ID	Text	Input site ID number.
Segment ID	Unique number; integer	The GPS unit assigns this number automatically.
Road Name	Text	Recorded on first/last segment of road, rest blank.
Group ID	Integer	Leave field blank unless this road segment drains to the same point as another segment; then assign both (or all) segments draining to the same point the same Group ID number (unique to those segments).
Surveyor	Name	Input surveyor name.
Weather	Clear Rainy Cloudy Snow	Determine average conditions for day.
Reference Point ID	Text	Record on first/last segment of road.
Segment Start Station	Number	Record JAMAR reading at start of segment.
Drainage Point Station	Number	Record JAMAR reading at drainage point.
Segment End Station	Number	Record station at end of segment (JAMAR reading or string box).
Drainage Point Type	Culvert Bridge Ditchout Drivable dip Sag point Other segment Other/unknown	Record type of drainage structure. Other segment – drains to other segment that delivers.
Structure Purpose	Stream crossing Relief, no gully Relief, gully No structure	Stream – defined bed and banks up <u>and</u> downstream of culvert Relief – no defined bed/banks Gully – defined bed/banks downstream <u>but not</u> upstream of culvert No structure – no structure observed
Delivery	0-none 1-direct or direct w/full trap 2-35% sediment 3-10% sediment 4-direct via gully or landslide scar 5-parallel within 20 feet	Determine delivery of ditch, drainage outfall, or road segment if outsloped based on flowchart "Guidance for delivery and breaking road segments" included in the field protocol.
Stream Adjacent Length	Length in feet	Measure length of road parallel and within 20 feet of a stream/wetland.
Culvert drop >9 inches?	Y/N	Is the vertical measurement between the water flowing out of the culvert and the water surface of the plunge pool more than nine inches? If stream is dry, measure from bottom of pipe to bottom of plunge pool.
Culvert downstream end countersunk?	Y/N	Is the downstream end of the culvert below the surrounding substrate?
Culvert upstream end countersunk?	Y/N	Is the upstream end of the culvert below the surrounding substrate?
Substrate extends through culvert?	Y/N	Look through culvert to determine if substrate extends the entire length of culvert.

Attribute	Possible Values	How Measured or Determined
Average Tread Width	Width in feet	Measure the full width of tread surface that could be driven on at 3-4 locations to nearest foot. Record average value (nearest foot).
Average Ditch Width	Width in feet	Measure width of ditch at 3-4 locations. Record average value (nearest foot).
Ditch Condition	P - recently pulled or graded E - scoured, eroding A - aggrading R - armored S – stable N/A	P – ditch has been recently pulled or graded E – ditch is eroding/incising. A – ditch is aggrading (full) R – ditch has been rocked/armored or is vegetated S – ditch appears stable (not eroding currently)
Maintenance Category	1-none recent 2-grading and/or ditch pulling only 3-other BMPs	Visually inspect the road to determine if maintenance has been completed recently (within last year).
Road Slope	<5% 5-10% >10%	Measure and record average gradient of tread with clinometer or estimate within slope class: <5% - flat or gently sloping road 5-10% - moderately sloped road segment >10% - steep road Average the gradient over entire segment. If the segment is a V-shaped stream crossing, estimate gradient on each side of crossing and average.
Surfacing	A-asphalt G-good gravel P-pitrun or worn gravel N-native	Determine surfacing on road tread. Use the following guidelines: Good gravel – a good gravel surface; little dust or fines on surface Pitrun or worn gravel – poor quality gravel surface; lots of fines or dust Native – dirt surface
Rutted?	No Yes/interfering Yes/not interfering	Are there ruts in road (over 2” deep)? If so, are they interfering with tread drainage?
Grassed tread?	Y/N	Is grass/vegetation covering more than 50% of tread surface?
Road Configuration (percent of tread width that drains to structure)	0% 25% 50% 75% 100%	Look at configuration of road prism. Evaluate the drainage path of water on the tread. Record the percent of the road tread drainage that delivers to the drainage structure to the nearest 25%.
Cutslope Average Height	25 ft 10 ft 5 ft 2.5 ft no cutslope	Average height of cutslope (slope length).
Cutslope Cover Density	90-100% 70-90% 50-70% 30-50% 10-30% 0-10% N/A	Determine the average percent of the cutslope area that is covered with vegetation, rock, leaf litter, or other non-erodible material.
Erosion Rating	High Low Default	Record rating only if obviously high (major cutslope raveling due to soil conditions) or low (bedrock). Otherwise record default.
Secondary segment?	Y/N	Record Y if this is a secondary segment – see field protocol for details.
Difficult segment?	Y/N	Indicates segment that is difficult or has complicated drainage.
Comments	Text	Note any unusual situations in the comment field.
Photo numbers	Text	Enter photo number from digital camera.
Date and time	yy/mm/dd hh:mm:ss	Automatically recorded by GPS unit.

3.7 Database Development

Concurrent with the development of the field protocol, an Access® database was prepared to store data collected during the road inventory. Data tables within the database include:

- Sample Unit Table – contains information about each sample unit.
- Landowner Table – contains contact information for each landowner.
- Road Segment Table – contains the information collected in the field, as well as modeled sediment delivery (from WARSEM) for each road segment that was determined to deliver to a stream, lake, or wetland.
- Non-Delivering Segment Table – contains information collected in the field for non-delivering drainage structures.

The database structure is included as an appendix to the Final Field Protocol.

3.8 Calculation of Monitoring Measures

Data for each sample unit and road segment was compiled for computation of monitoring measures and graphing. The monitoring measures for each of the monitoring questions or hypotheses listed above in Table 3 were computed for each sample unit based on the formulas shown in Table 4 below. Each monitoring measure was also compiled for the three geographical performance target regions (East of Crest, West of Crest, Coastal/Spruce) as well as for the entire state-wide sample. Statistical analyses are described in more detail in Section 3.8.1.

Table 4. Methods Used to Compute Monitoring Measures and Statistics

Monitoring Measure	**Computation Method**
MQ1.1, H2a.1 Total road length draining to streams (road miles/sq mi)	a. Un-weighted delivering road length[2] (mi)/Area of sample unit (sq mi)
	b. Weighted delivering road length[3] (mi)/Area of sample unit (sq mi)
MQ1.2 Percent of road network draining to streams	a. Un-weighted delivering road length[2] /Total surveyed road length[1]
	b. Weighted delivering road length[3] /Total surveyed road length[1]
MQ1.3 Percent of road in each surface category	Surveyed road length in each surface category[1] /Total surveyed road length[1]
MQ1.4 Percent of road in each traffic category	Surveyed road length in each traffic category[1] /Total surveyed road length[1]
MQ1.5 Percent of road in each cutslope cover category	Surveyed road length in each cutslope cover category[1] / Total surveyed road length[1]
MQ1.6 Percent of drainage points by connectivity class	Number of delivery points in each connectivity category / Total number of surveyed delivery points
MQ1.7 Percent of road in each road rutting category	Surveyed road length in each rutting category[1] /Total surveyed road length[1]
H2b.1 Road surfacing index	Road attribute index for surfacing (see [B] below)
H2b.2 Road traffic index	Road attribute index for traffic (see [B] below)
H2b.3 Cutslope cover index	Road attribute index for cutslope cover (see [B] below)
H2b.4 Miles of delivering road with ruts interfering with drainage	Surveyed length of roads[1] with ruts that were classified as interfering with drainage (mi)

Monitoring Measure	**Computation Method**
MQ3.1 Miles of forest road delivering to streams per miles of stream (road hydrology performance measure)	a. Un-weighted delivering road length[2] (mi) / Adjusted length of streams within sample unit[C] (mi)
	b. Weighted delivering road length[3] (mi) / Adjusted length of streams within sample unit[C] (mi)
MQ3.2 WARSEM modeled tons of road sediment delivered to streams per miles of stream per year (sediment performance measure)	WARSEM modeled tons of road sediment delivered to streams (tons/yr) / Adjusted length of streams within sample unit[C] (mi)
MQ4.1, H5a Miles of forest road delivering to streams per miles of stream (road hydrology performance measure) divided by the performance target by target region	a. Un-weighted delivering road length[2] (mi) / Adjusted length of streams within sample unit[C] (mi) divided by upper performance target: East of Crest = 0.12 West of Crest = 0.25 Coastal (Spruce) = 0.25
	b. Weighted delivering road length[3] (mi) / Adjusted length of streams within sample unit[C] (mi) divided by upper performance target
MQ4.2, H5b WARSEM modeled tons of road sediment delivered to streams per miles of stream per year (sediment performance measure) divided by the performance target by target region	WARSEM modeled tons of road sediment delivered to streams (tons/yr) / Adjusted length of streams within sample unit[C] (mi) divided by upper performance target: East of Crest = 3 West of Crest = 6 Coastal (Spruce) = 10
H6a Miles of forest road delivering to streams per miles of stream by percent of road length meeting performance standards	Data is displayed graphically
H6b WARSEM modeled tons of road sediment delivered to streams per miles of stream per year by the percent of road length meeting performance standards	Data is displayed graphically

A. Note that there are three different ways the length of road was used in the formulas in Table 4. Superscripted numbers in the right column apply to the following definitions:
   1. Total surveyed road length in the unit includes length of all roads surveyed (delivering and not delivering).
   2. Total delivering road length = the total (un-weighted) length of road that was in any delivering class (direct, direct via gully, 35% delivery, 10% delivery). Does not include non-delivering or stream-parallel length.
   3. Weighted delivering road length = total delivering road length (#2 above) for each segment multiplied by the delivery percent (100%, 35%, 10%) multiplied by the percent of the tread width delivering (100%, 75%, 50%, 25%, 0%).

B. Road Attribute Indexes for categorical values are computed using the following formula:

$$\frac{\sum \text{Surveyed road length in each attribute category}^{1} \text{ x WARSEM factor for that attribute category}}{\text{Total surveyed road length}^{1} \text{ x WARSEM factor associated with least amount of erosion for that attribute}}$$

C. Adjusted stream length = length of streams within each sample unit taken from WDNR GIS coverage x [number of road-stream crossings found in field / number of road-stream crossings in GIS]. This adjustment takes into consideration the fact that the GIS stream coverage may miss or have extra streams in many locations.

Note that the length of streams within each sample unit is included as part of the calculation of several of the monitoring metrics. The length of stream was based on the WDNR GIS stream layer, which is often not a true representation of actual streams on the ground. The road inventory crew checked each road/stream crossing within the sample unit to assess whether mapped and un-mapped streams were or were not present. The miles of streams in the GIS layer were then proportionally adjusted based on the actual number of stream crossings found within the unit versus the number of stream crossings in the GIS layer (see note C below Table 4). While this may not result in a completely accurate portrayal of stream miles, it was the most cost-effective method of adjusting stream miles within the scope of this project.

The length of road delivering to steams was measured in the field and included length in each of the delivering categories (direct, direct via gully, 35% delivery and 10% delivery). This length was used for comparison with the FFR length performance measure (RLEN, Table 1 above). The total delivering length metric does not take into account some strategies that landowners may use to reduce road connectivity such as crowning or outsloping a road, installing cross drains (if they are still within proximity to a stream), or sediment traps. For this reason, a weighted road length was computed for comparison with future monitoring events. The weighted delivering road length was computed based on the total delivering road length, delivery category, and percent of tread width draining to the delivery point (Table 4).

The multi-year, state-wide scale of the road effectiveness monitoring study precluded direct measurement of sediment production. For this reason, a modeling approach was chosen to estimate sediment production for the FFR sediment metric (MQ3.2, MQ4.2, H5b, and H6b). WARSEM was chosen to estimate sediment delivery for the following reasons (Raines et al. 2005):

1. WARSEM is closely related to the Road Surface Erosion Module calculations in the Washington Watershed Analysis Manual (Washington Forest Practices 1997) that was used to develop the FFR performance targets for sediment delivery;

2. WARSEM input variables include the important road attributes controlling road sediment production with categories that provide clear interpretation of improving trends;

3. Input data from WARSEM can be used in alternate models if the need arises; and

4. WARSEM is easy to use and is well suited for estimating sediment delivery for large datasets.

It is recognized that output from any road surface erosion model is not an accurate measure of sediment production or delivery at the scales of individual delivery points or individual road segments; however, models are useful for comparing trends in sediment production through time in response to changes in road conditions (Dubé et al. in press). The monitoring measures and statistics evaluated here include 15 different measures/indices; only two of these use WARSEM-generated numbers.

The Washington State Road Surface Erosion Model (WARSEM) was used to estimate the sediment production and delivery from each delivering road segment. The characteristic of each road segment (length, tread and ditch width, tread configuration, surfacing, road gradient,

cutslope height and cover, traffic use, delivery, and Township/Range/Section location) were entered into WARSEM.

The erosion and delivery calculations in WARSEM are based on a set of empirical relationships that were developed from research on road erosion (Dubé et al. 2004). The model uses the following formulas to calculate road surface erosion:

> Total Sediment Delivered to a Stream from each Road Segment (in tons/year) = (Tread & Ditch Sediment + Cutslope Sediment) x Road Age Factor
>
> Tread & Ditch = Geologic Erosion Factor x Tread Surfacing Factor x Traffic Factor x Segment Length x Road (Tread + Ditch) Width x Road Gradient Factor x Rainfall Factor x Delivery Factor
>
> Cutslope = Geologic Erosion Factor x Cutslope Cover Factor x Segment Length x Cutslope Height x Rainfall Factor x Delivery Factor

The value for each factor in the equations is selected by WARSEM based on the road characteristics collected in the field and entered for each individual road segment. The delivery categories assigned in the field (100%, 35%, and 10%) correspond directly to WARSEM delivery factors.

3.8.1 Statistical Analysis

The statistical analysis relevant to monitoring questions 1, 3, and 4, and hypothesis 6 followed the methods discussed in the Monitoring Design (Raines et al., 2005). The following are clarifications and modifications to the statistical analysis methods described in that document. Arithmetic means, medians, and standard deviations of the unit results are provided for each region. The statewide median is simply the median across all 60 units. The statewide mean and standard deviation are weighted means of the regional results, weighted by sample size. Some discussion as to the appropriateness of the arithmetic mean as an estimate of statewide metrics is necessary

Sample units of roughly equal land area were selected randomly in three geographic regions, proportional to the estimated land area subject to FFR rules. However, the sampling design within each region can be looked at as a cluster sample, where each unit is a cluster of roads (or streams, depending on the metric being estimated). Some units have many roads or streams while others have few roads or streams, so the clusters differ in size with regard to the amount of road or stream. Typically, when calculating a landscape average, larger clusters (i.e., units with more roads) should have greater weight than smaller clusters if one wants to infer results to additional stream- or road-miles. However, Washington state forestlands are tracked, assessed, and managed on an area-basis rather than on a road- or stream-mile basis. We need to be able to infer results to additional land area, not road- or stream-mile. Therefore it was decided that the unit-by-unit averages (i.e., not weighted by length of road or stream) were more relevant as management summaries of results. These averages are biased estimates of state-wide means of road characteristics with less precision than if samples were weighted by and inferred to road mile.

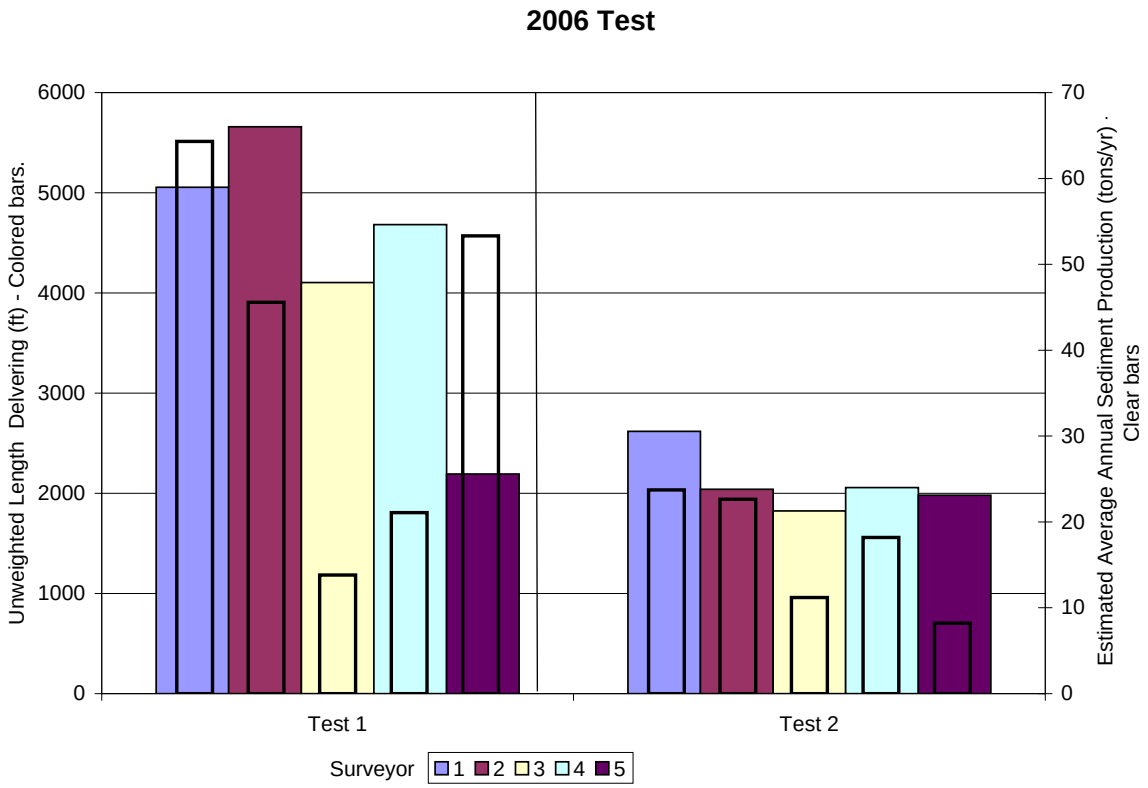
For hypothesis 6, the Monitoring Design states that linear regression will be used to determine if there is a relationship between the percentage of the road system that is judged to meet road standards and two reported metrics. Because the metrics are not normally distributed or transformable to normal distribution, a nonparametric (Kendall's tau; Conover 1980) regression/correlation test was used instead. Using the nonparametric test also loosens the assumption of linear regression that the x-variable is measured without error. For the nonparametric test, it is only necessary that the x-variable is ranked properly, which seems to be a reasonable assumption.

4 Results and Discussion

4.1 Crew Variability Test

The crew variability test originally had two goals: 1) to target areas for training to improve consistency; and 2) to provide a quantitative estimate of crew variability for delivering length and sediment yield.

The test was optimized to provide feedback to and guide training for field crews for quality assurance purposes. The test results were very helpful to target areas for further training. From the results of the first test in 2006, project management personnel targeted field assistance and additional discussion with field crew members on assessing delivery (using the flowchart developed for this purpose), using a clinometer to measure road gradient, determining configuration (full/half/none) with care, and developing a common understanding of surfacing between field workers in each unit. During the 2008 season, these measurements continued to be stressed, along with protocols for determining delivery over fillslopes and deciding if a drainage swale was considered a stream or not. Road variables influencing the sediment production that varied most among observers included delivery, gradient, portion of tread delivering, and surfacing. Several of the roads in the 2008 test unit were outsloped, which means that the road drains over the fillslope to the stream. These types of road drainage systems are prevalent in Eastern Washington roads. Determination of delivery in these cases is particularly difficult, and was one reason for the high variability between observers during the 2008 testing. Field crew members who were used to surveying Westside road systems with pronounced ditch systems were looking at delivery from the poorly developed/short ditch systems on the Eastside (2008 test unit) roads. Field crew members who normally surveyed Eastside road systems were considering delivery over the fillslopes. The post-test discussions led to a common way to look at outsloped Eastside road systems (delivery over fillslopes) and should have further reduced variability for this metric after the test. In general, there is less variation in results during the second test each field season than in the first test, suggesting that results became more consistent with continued training and working together during the duplicate crew surveys (Figure 4).



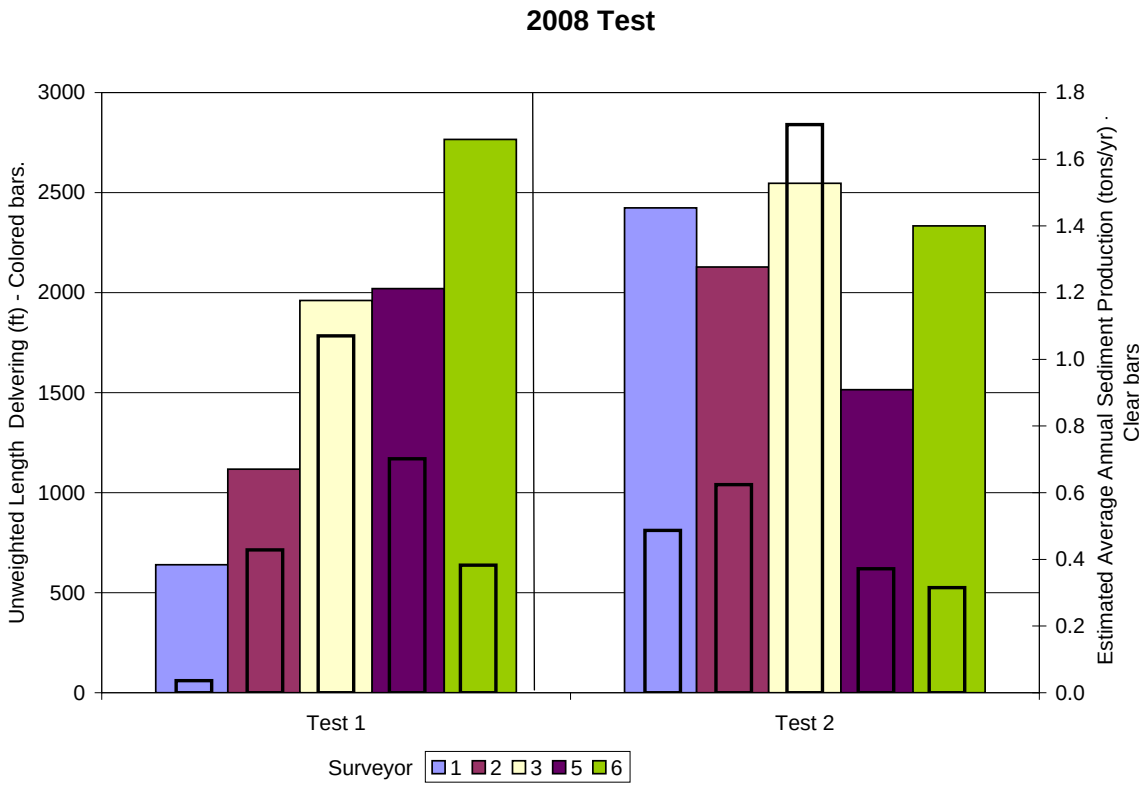


Figure 4. Length Delivering and Sediment Production by Surveyor for 2006 and 2008 Tests.

Because the crew variability tests were optimized for quality assurance and training purposes, the results are not useful for estimating measurement error for the entire project. While it is possible to calculate the amount of variance among surveyors based on the 2006 and 2008 test results, it is not appropriate to apply the results to the data collected in an entire sample unit for several reasons:

- The variability in delivering length and calculated total sediment production among observers on a single segment or section of road is quite large relative to differences among road segments. However, the variability among sums of road segments, which are being used for the larger project will be lower because they are sums (or averages), which have lower theoretical variance (i.e., the variance of a mean is reduced proportionally to the sample size). Since the crew variability study measured only a small length of road (1.4-2.4 miles) in relation to the total road miles in a sample unit (average 23 miles/unit), it was not appropriate to extrapolate the results to entire units.

- Crew variability differed from the first to the second test (Figure 4) and from 2006 to 2008. In addition, crew members sampled different amounts of roads in different regions across the state. This would complicate efforts to apply appropriate variance estimates to all the sample units across the state and through time.

- During subsequent sampling events, the Monitoring Design specifies that the same study sites will be re-sampled. Since trends between sampling events will be based upon change of only those variables that can be modified by management practices within each sample unit, there will be fewer variables with the potential for observer error.

For the study results, these crew variability results imply that differences observed through time (particularly for only two time periods) could be purely due to observation error. Therefore, strong interpretation of data results should only be made after consistent changes through time (after repeated measurements) have been observed. Minimizing observer variability is important to this and subsequent monitoring periods. Consideration of ways to minimize variability should be stressed during both the planning and implementation phases of future monitoring events.

4.2 Sample Units

This first sampling event provides information on the current (2006-2008) status of forest roads on lands subject to Washington Forest Practice Rule across the state. A total of 60 four-square-mile sample units were inventoried (Figure 5). The units were distributed by performance target geographical region based on the relative proportion of forest practice rule lands within each region (East of Cascade Crest = 30 sample units, Coastal/Spruce = 7 sample units, and West of Cascade Crest = 23 sample units).

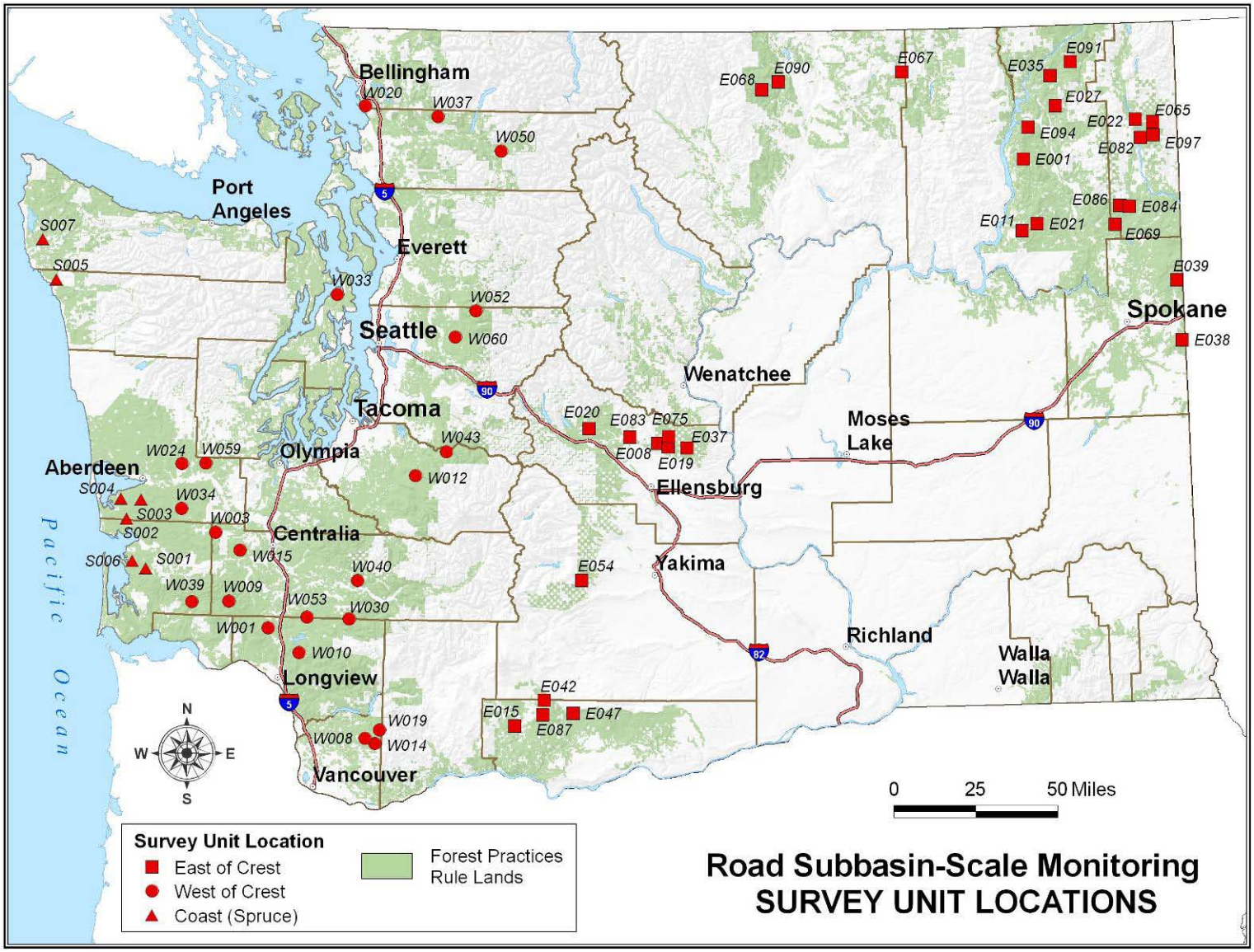


Figure 5. Location of Road Monitoring Sample Units.

Sample units beginning with the letter "E" are east of the Cascade crest, "W" are west of the Cascade crest, and "S" are in the Coastal Spruce zone.

4.2.1 Sample Unit Selection and Rejection

As described in Section 3.4.3, sample units initially selected for study were screened to ensure that each unit included four square miles of relatively contiguous blocks of FFR-regulated lands with acceptable uses, and that land owners granted access to their lands for the study. Of 164 sample units originally selected, 37% were rejected in the initial assessment due to lack of FFR land, and 14% were rejected because the FFR land was not contiguous enough to create blocks large enough to meet study unit criteria (Table 5). Units rejected at this first stage of screening were not assessed in enough detail to specify particular reasons for the lack of FFR land. In later stages of screening, where individual quarter sections were assessed, 3% of the units were rejected because of too much agricultural or other non-forest area; 6% due to development at various levels; 3% due to landowner denial; and one unit due to inaccessibility as the result of a large storm in southwest Washington.

Table 5. Potential Study Unit Rejection Reasons and Rates

Reason for Rejection	Number of Units	Percentage of Total Sites Assessed (n=164)
Not enough FFR	61	37%
Not enough contiguous FFR	23	14%
Development	10	6%
Agricultural land, non-forest	4	3%
Landowner denial	5	3%
Inaccessibility	1	1%
Total Rejected	104	63%

No sample units selected for use in the Southeastern part of the state around the Blue Mountains met the sample unit criteria and had enough of those landowners willing to grant access. Several units were initially identified in that region, but they were all rejected. Private forestland in the Blue Mountain area forms a narrow band between the agricultural and rangeland and the National Forest. Moreover, the forested areas tend to be in the valleys that weave among the agricultural land on plateaus and hillsides. The roads in these areas are agricultural roads and are not designed or maintained as forest roads under the forest practices rules and so they were not accepted into study units. These factors made it very difficult to compile a block of FFR land of the necessary size in that region.

Roads sampled include those owned by twenty-one large landowners, thirty small landowners, three state agencies, and four municipalities. The sampled area is dominated by large landowner and State forest land properties, which constitute approximately ninety-five percent of the sampled area. Although property from small forest landowners was specifically included in this study, small forest ownerships are underrepresented (by area) in the final sample. Across the state it is estimated that 40 percent of FFR land is owned by small landowners (landowners having an average harvest of less than 2 million board feet annually; Rogers and Cooke 2009). However, most of the small forest landowner properties occur on the margins of developed areas and include houses or agricultural lands that were incompatible with study site criteria. This intermingling with qualifying forest lands was the largest cause for rejection of individual quarter sections. Lack of landowner cooperation also contributed to the low acceptance rates of

quarter sections containing lands belonging to small forest landowners. CMER studies rely on voluntary participation by landowners. Although many small forest landowners were willing to cooperate, any one uncooperative landowner resulted in rejection of the entire quarter section for that unit. The end result is that, despite the relatively high number of participating small forest landowners, small forest parcels make up less than five percent of the sampled area. All large forest landowners and government entities approached participated in this study.

Because small non-industrial forest ownerships (by land area) were under-represented, this study likely does not adequately reflect road conditions on small forest ownerships and should not be used to infer road conditions on those ownerships. The high rejection rate for these has convinced us that either a change in the sampling design or a companion study specifically designed to obtain access to and evaluate small forest ownerships would be required to provide an accurate assessment of the status of roads on those lands.

4.2.2 Sample Unit Results

Table 6 summarizes the number of miles of road surveyed and miles of stream in each sample unit, the un-weighted and weighted length of road delivering to streams/wetlands, and the estimated sediment delivering to streams and the FFR sediment performance metrics.

Note that one East of Crest survey units had no streams. As a result, there were no delivery points in this unit, and no data on road conditions were collected. This unit has 0 miles of road delivering and appears as a blank on the bar charts in Section 4.3.

Table 6. Summary of Data Collected by Sample Unit

Area	Miles of Road Surveyed	Surveyed Road Density (mi/sq mi)	Adjusted Stream Length (mi)	Stream Density (mi/sq mi)	Un-weighted Road Length Delivering (mi)	Weighted Road Length Delivering (mi)	WARSEM Calculated Delivered Sediment (tons/yr)	RLEN - Un-weighted Miles of Road Delivering/ Stream Length (mi/mi)	Weighted Miles of Road Delivering/ Stream Length (mi/mi)	RSED - Sediment Delivered/ Stream Length (tons/yr/mi)
Sample units located East of Crest	27.3	7.12	5.2	1.35	1.51	0.94	1.4	0.3	0.18	0.28
	25.0	6.26	4.0	1.01	0.66	0.30	0.4	0.2	0.07	0.10
	26.0	6.50	5.6	1.40	0.83	0.52	0.7	0.1	0.09	0.13
	24.3	6.14	8.3	2.09	5.32	3.58	13.4	0.6	0.43	1.62
	20.5	5.15	3.9	0.98	1.83	1.24	2.2	0.5	0.32	0.57
	18.1	4.53	7.5	1.89	1.49	0.70	6.9	0.2	0.09	0.92
	29.8	7.16	3.0	0.72	1.25	0.68	1.0	0.4	0.23	0.34
	13.4	3.39	4.0	1.01	0.08	0.01	0.2	0.0	0.00	0.04
	14.1	3.51	0.0	0.00	0.00	0.00	0.0	0.0	0.00	0.00
	5.7	1.46	7.9	2.04	0.06	0.01	0.0	0.0	0.00	0.00
	13.9	3.47	1.9	0.48	0.10	0.01	0.0	0.1	0.01	0.01
	22.3	5.42	8.2	2.00	2.66	1.77	1.9	0.3	0.22	0.23
	2.7	0.69	7.6	1.90	0.08	0.01	0.0	0.0	0.00	0.00
	26.2	6.71	17.0	4.37	2.33	1.99	10.7	0.1	0.12	0.63
	17.0	4.10	7.4	1.78	0.70	0.33	0.8	0.1	0.05	0.11
	19.8	4.93	14.5	3.62	4.33	3.46	14.6	0.3	0.24	1.01
	20.1	5.05	3.9	0.98	1.19	0.32	3.6	0.3	0.08	0.92
	22.2	5.63	8.9	2.26	0.66	0.07	0.6	0.1	0.01	0.06
	4.1	1.05	10.9	2.79	0.27	0.10	0.1	0.0	0.01	0.01
	10.7	2.67	23.2	5.79	1.64	1.64	3.5	0.1	0.07	0.15
	25.3	6.31	4.3	1.07	1.50	0.59	1.2	0.4	0.14	0.28
	16.6	4.10	5.5	1.37	0.32	0.03	0.1	0.1	0.00	0.02
	21.9	5.49	8.6	2.14	1.67	1.01	4.6	0.2	0.12	0.53
	10.9	2.67	13.9	3.40	0.38	0.22	0.2	0.0	0.02	0.02
	31.1	7.46	13.6	3.26	1.58	1.03	7.6	0.1	0.08	0.56
	17.2	4.32	9.0	2.27	1.08	0.92	1.9	0.1	0.10	0.21
	15.5	3.83	5.3	1.31	0.08	0.07	0.1	0.0	0.01	0.01
	20.4	5.48	5.4	1.45	0.23	0.05	0.3	0.0	0.01	0.06
	20.9	5.37	9.9	2.54	0.41	0.28	0.5	0.0	0.03	0.05
	14.9	3.71	6.0	1.50	0.44	0.04	0.2	0.1	0.01	0.04
Units in Coastal/Spruce	25.9	6.50	24.2	6.06	4.09	1.93	117.9	0.2	0.08	4.88
	15.3	4.28	59.9	16.75	0.98	0.55	29.3	0.0	0.01	0.49
	25.2	6.32	11.0	2.75	0.71	0.23	173.9	0.1	0.02	15.83
	25.2	6.39	6.7	1.69	0.67	0.30	1.8	0.1	0.05	0.27
	27.8	6.95	48.9	12.23	6.21	2.86	224.9	0.1	0.06	4.60
	29.0	7.19	24.1	5.96	3.69	1.77	30.9	0.2	0.07	1.28
	21.6	5.60	27.2	7.05	4.55	2.34	102.4	0.2	0.09	3.77

Table 6. Summary of Data Collected by Sample Unit (continued)

Area	Miles of Road Surveyed	Surveyed Road Density (mi/sq mi)	Adjusted Stream Length (mi)	Stream Density (mi/sq mi)	Un-weighted Road Length Delivering (mi)	Weighted Road Length Deliverin g (mi)	WARSEM Calculated Delivered Sediment (tons/yr)	Un-weighted Length of Road Delivering/ Stream Length (mi/mi)	Weighted Length of Road Deliverin g/ Stream Length (mi/mi)	Sediment Delivered/ Stream Length (tons/yr/ mi)
Sample units located West of Crest	28.8	7.27	17.4	4.41	7.08	4.42	217.3	0.4	0.25	12.45
	29.9	7.63	26.8	6.85	4.78	2.36	407.6	0.2	0.09	15.19
	26.4	6.56	13.0	3.23	4.89	2.79	60.6	0.4	0.21	4.65
	33.0	7.63	15.5	3.60	7.35	3.76	137.0	0.5	0.24	8.82
	29.2	7.35	21.1	5.31	7.93	2.05	30.2	0.4	0.10	1.43
	22.9	5.83	22.7	5.78	3.13	1.20	26.1	0.1	0.05	1.15
	17.1	4.33	23.5	5.94	3.35	1.12	73.9	0.1	0.05	3.15
	27.7	6.95	4.6	1.16	1.72	0.60	7.6	0.4	0.13	1.66
	27.2	6.74	14.5	3.58	6.64	2.91	93.5	0.5	0.20	6.47
	14.6	3.79	15.8	4.11	1.86	1.00	9.7	0.1	0.06	0.62
	26.4	6.79	13.7	3.53	0.73	0.22	7.5	0.1	0.02	0.55
	26.0	6.51	18.7	4.69	5.96	2.02	47.6	0.3	0.11	2.54
	22.7	5.71	4.6	1.17	0.49	0.17	0.4	0.1	0.04	0.09
	25.4	6.41	22.6	5.71	3.25	0.93	24.3	0.1	0.04	1.07
	8.9	2.44	59.2	16.22	4.37	2.35	205.6	0.1	0.04	3.48
	20.6	5.64	14.3	3.91	5.69	2.35	163.7	0.4	0.16	11.47
	24.2	6.26	4.3	1.12	0.63	0.24	1.5	0.1	0.05	0.35
	24.7	6.31	19.2	4.90	4.74	2.07	266.0	0.2	0.11	13.88
	23.4	5.86	10.0	2.51	1.75	0.90	28.5	0.2	0.09	2.85
	24.8	6.09	23.9	5.86	3.21	0.79	39.9	0.1	0.03	1.67
	21.3	5.32	11.1	2.78	4.42	2.04	13.5	0.4	0.18	1.22
	25.4	6.43	32.6	8.25	0.49	0.26	7.6	0.0	0.01	0.23
	24.3	6.16	24.5	6.20	6.53	3.61	68.9	0.3	0.15	2.82

Road density (miles of road/square mile of area), is commonly cited as a metric for the impacts of roads on streams. Although data from sample units suggest a generally positive correlation between road density and the length of road hydrologically connected to streams (Figure 6), and the predicted sediment input to streams (Figure 7), there is not always a direct correlation due to other contributing factors. The location of the road network in relation to the stream network, as well as road condition and traffic levels are important factors for determining the relative effect of roads on streams, aquatic habitat, and water quality.

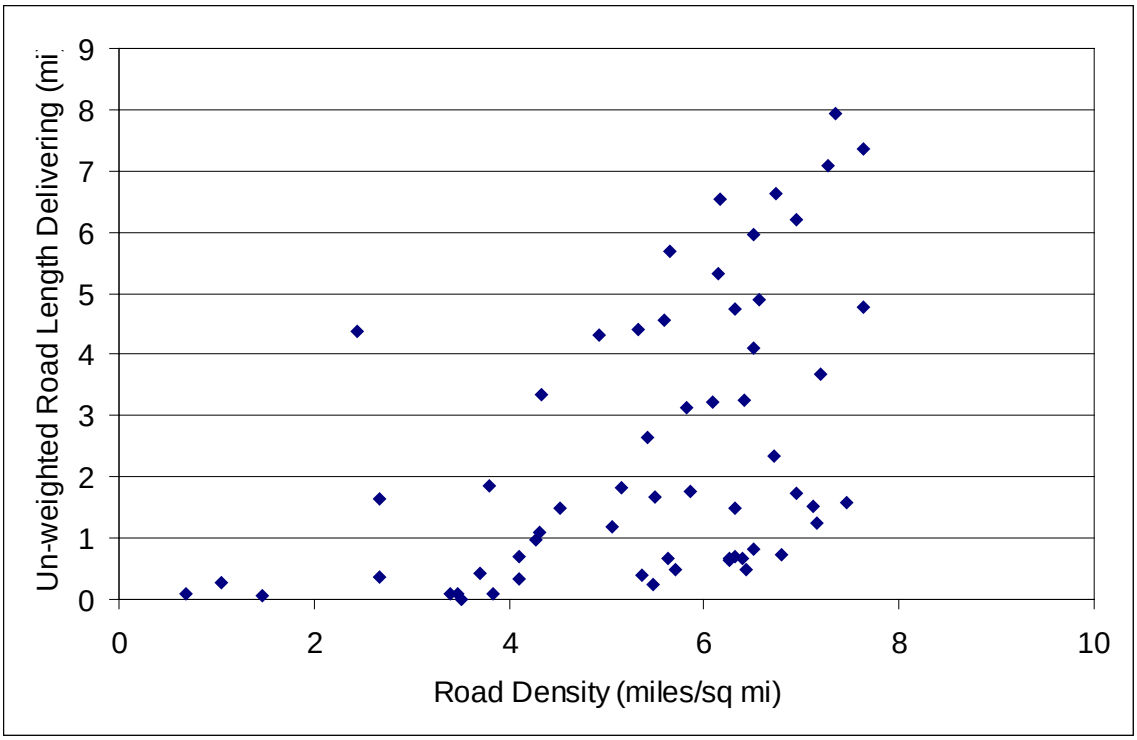


Figure 6. Un-weighted Road Length Delivering in Sample Unit vs. Road Density.
Includes all sample units across state

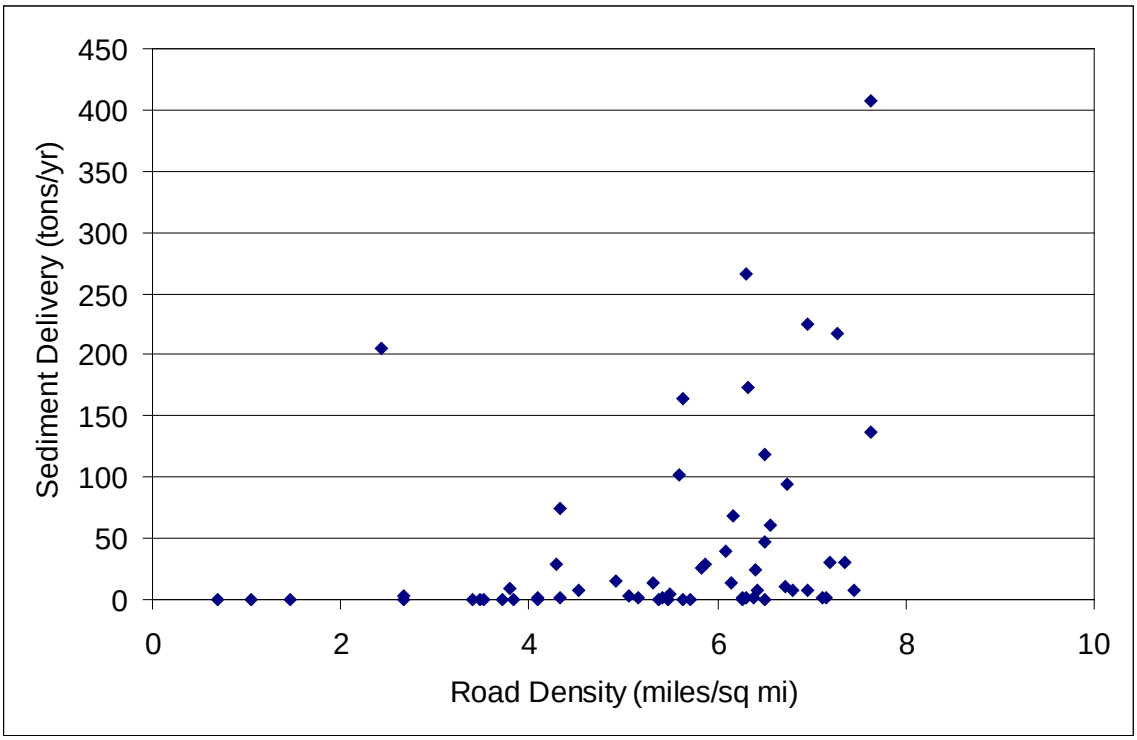


Figure 7. Estimated Sediment Delivered to Streams in Sample Unit vs. Road Density.
Includes all sample units across state

The total length of road delivering to streams was measured in the field and included both direct and indirect (35%, 10%) delivery categories, and so includes road segments where some of the

strategies land managers use to reduce the amount of road runoff and sediment draining to streams – such as crowning or outsloping a road, installing a sediment trap, or installing a cross-drain culvert with partial delivery – have been applied. In order to understand the magnitude of the difference such measures may make, a weighted delivering length was also calculated that takes into consideration delivery percentage and percent of road tread delivering to each drainage point. Figure 8 shows the comparison of un-weighted (total delivering length) and weighted road length. Across the state, the weighted road length delivering was an average of 45% of the un-weighted road length delivering.

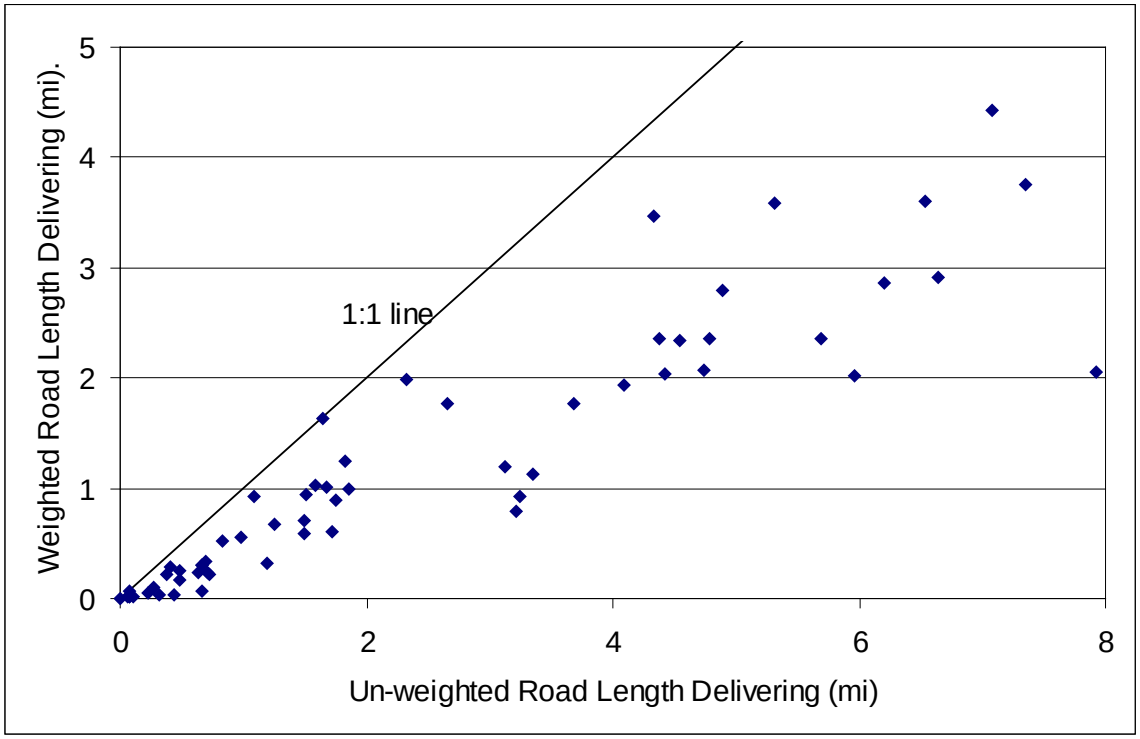


Figure 8. Un-weighted vs. Weighted Road Length Delivering for the Sample Units.
Includes all sample units across state

Road managers are also interested in which parts of their road system are delivering the most sediment. Figure 9 shows the length of road in four sediment yield categories (average annual delivered sediment in tons/yr at a drainage point) in each of the sample units. The majority of the road system in the units had either no sediment delivery, or less than one ton/yr of delivery to streams. A small portion of road drainage points in some of the Coast and West of Crest units had over 3 tons/yr of sediment delivery. Targeting segments with higher sediment yields for road maintenance and improvements may be a cost-effective method to reduce total delivery of road surface erosion to streams.

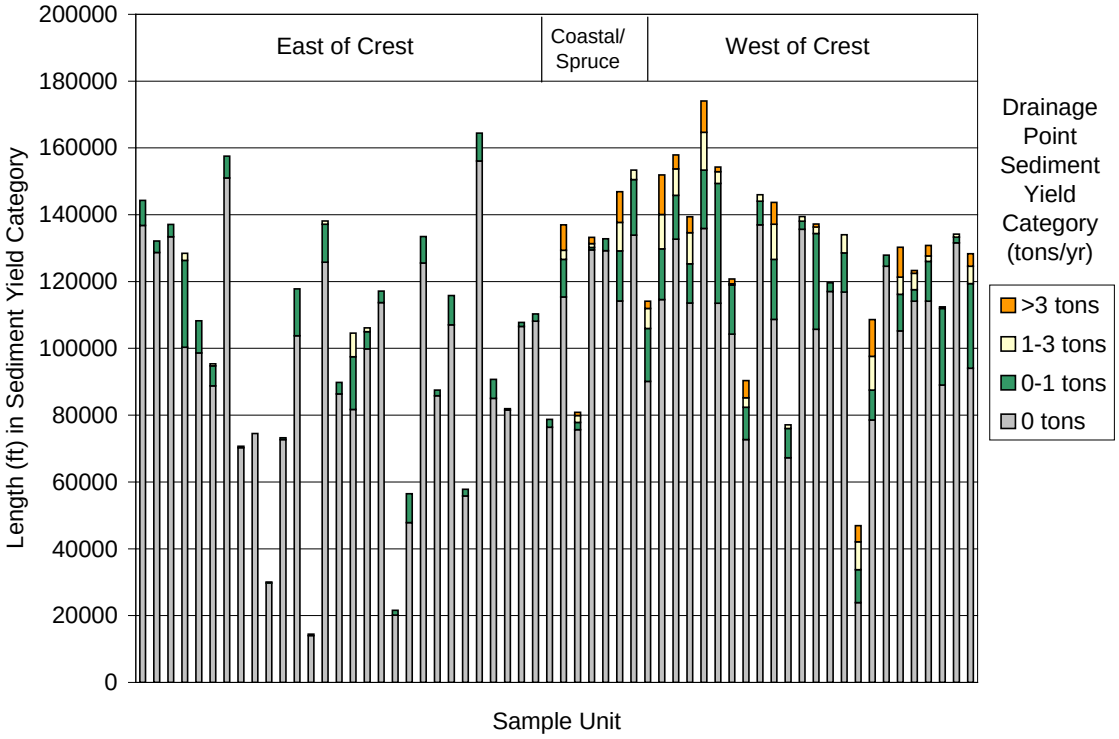


Figure 9. Length of Road by Sediment Yield Category in Each Sample Unit Showing Overall Variability Among Sites.

4.3 Monitoring Questions and Measures

Six monitoring questions/hypotheses, each with one or more metrics, were developed in the study plan to determine the status of forest road characteristics that relate to water and sediment delivery to streams. The mean median, and standard deviation of the monitoring measures listed in Table 2 (above) were computed for each geographic target region as well as state-wide (Table 7). The following sections describe the results for each of the monitoring questions/hypotheses.

Table 7. Mean, Median and Standard Deviation of Monitoring Measures by Region and State.

Monitoring Question or Hypothesis			East of Crest n=30			Coast/Spruce n=7			West of Crest n=23			Statewide n=60		
			Mean	Median	Stand. Dev.	Mean	Median	Stand. Dev.	Mean	Median	Stand. Dev.	Mean	Median	Stand. Dev.
MQ1, H2a	1a. Total un-weighted road length delivering per unit area (road miles/sq mi)		0.29	0.19	0.31	0.76	0.91	0.55	1.00	1.11	0.59	0.62	0.40	0.47
	1b. Total weighted road length delivering per unit area (road miles/sq mi)		0.18	0.08	0.24	0.36	0.44	0.26	0.44	0.51	0.30	0.30	0.23	0.27
	2a. Percent of road network delivering (un-weighted length/length inventoried)		6%	4%	6%	12%	13%	8%	17%	19%	11%	11%	7%	8%
	2b. Percent of road network draining to streams (weighted length/length inventoried)		4%	2%	5%	6%	6%	4%	8%	7%	6%	5%	4%	5%
	3. Percent of road in each surface category (based on total length delivering)	Asphalt	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	2%
		Good gravel	8%	0%	22%	23%	20%	13%	15%	15%	13%	12%	2%	18%
		Pitrun/worn gravel	14%	8%	18%	67%	73%	16%	77%	77%	20%	44%	48%	19%
		Native	74%	89%	32%	10%	9%	10%	8%	3%	14%	42%	25%	24%
	4. Percent of road in each traffic category (based on total length delivering)	Very heavy	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		Heavy	0%	0%	0%	9%	0%	22%	0%	0%	0%	1%	0%	8%
		Mod Heavy	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		Moderate	0%	0%	1%	7%	0%	12%	7%	0%	14%	4%	0%	10%
		Light	8%	0%	21%	3%	0%	4%	8%	0%	12%	7%	0%	17%
		Occasional	84%	100%	32%	81%	84%	22%	85%	92%	18%	84%	100%	26%
		None	5%	0%	20%	0%	0%	0%	0%	0%	0%	2%	0%	14%
	5. Percent of road in each cutslope cover category (based on total length delivering)	90-100 %	41%	36%	34%	45%	53%	15%	61%	65%	20%	49%	56%	27%
		70-89 %	16%	15%	15%	23%	21%	8%	14%	11%	11%	16%	15%	13%
		50-69 %	12%	5%	16%	11%	8%	9%	7%	4%	8%	10%	5%	13%
		30-49 %	4%	0%	9%	4%	4%	4%	3%	1%	6%	4%	0%	8%
		10-29 %	3%	0%	8%	2%	0%	4%	1%	0%	1%	2%	0%	6%
		0-9 %	0%	0%	1%	2%	0%	4%	0%	0%	0%	0%	0%	1%
		N/A	20%	16%	18%	13%	13%	14%	14%	13%	12%	17%	13%	15%
	6. Percent of drainage points by connectivity class	Direct	47%	49%	34%	55%	52%	13%	58%	57%	19%	52%	53%	27%
		Direct via gully	4%	0%	6%	8%	9%	5%	9%	9%	8%	6%	3%	7%
		35% Delivery	10%	9%	10%	24%	23%	9%	22%	20%	17%	16%	14%	13%
		10% Delivery	34%	19%	35%	13%	14%	7%	10%	8%	8%	23%	12%	25%
		Stream Parallel	2%	0%	5%	0%	0%	1%	2%	0%	4%	2%	0%	5%
	7. Percent of road in each road rutting category (based on total length delivering)	Not Rutted	92%	100%	20%	88%	98%	22%	98%	100%	4%	94%	100%	16%
		Rutted - Interfering	4%	0%	6%	8%	0%	16%	1%	0%	2%	3%	0%	7%
		Rutted - Not Interfering	2%	0%	5%	4%	0%	6%	1%	0%	2%	1%	0%	4%

Monitoring Question or Hypothesis		East of Crest n=30			Coast/Spruce n=7			West of Crest n=23			Statewide n=60		
		Mean	**Median**	**Stand. Dev.**	**Mean**	**Median**	**Stand. Dev.**	**Mean**	**Median**	**Stand. Dev.**	**Mean**	**Median**	**Stand. Dev.**
,MQ3, H5a H5b	1a. Un-weighted miles of forest road delivering per miles of stream (road hydrology performance measure)	0.16	0.11	0.16	0.11	0.13	0.06	0.24	0.18	0.14	0.19	0.14	0.14
	1b. Weighted miles of forest road delivering per miles of stream (road hydrology performance measure)	0.09	0.07	0.11	0.05	0.06	0.03	0.11	0.09	0.07	0.09	0.07	0.09
	2. WARSEM modeled tons of road sediment delivered per miles of stream per year (sediment performance measure)	0.30	0.12	0.39	4.44	3.77	5.38	4.25	2.54	4.73	2.30	0.56	3.47
MQ4	1a. Un-weighted miles of forest road delivering per miles of stream (road hydrology performance measure) divided by the performance target by target region	1.33	0.88	1.34	0.46	0.51	0.23	0.96	0.71	0.57	1.09	0.67	1.01
	1b. Weighted miles of forest road delivering per miles of stream (road hydrology performance measure) divided by the performance target by target region	0.76	0.60	0.89	0.21	0.23	0.12	0.42	0.36	0.30	0.56	0.35	0.65
	2. WARSEM modeled tons of road sediment delivered per miles of stream per year (sediment performance measure) divided by the performance target by target region	0.10	0.04	0.13	0.44	0.38	0.54	0.71	0.42	0.79	0.37	0.15	0.53

4.3.1 Monitoring Question 1

What is the condition of forest roads at each sample event, specifically those attributes management can change relative to sediment production and delivery?

MQ 1.1 and H2a1 Total road length draining to streams per unit area (road miles/sq mi)

The total length of road draining to streams per unit area sampled provides a measure of the length of road that is hydrologically connected to streams in each sample unit (Figure 10). Across the state, half of the sample units had less than 0.4 miles/sq mi of delivering road. Units in the West of Crest and Coast/Spruce geographic zones had more weighted road length delivering than those in the East of Crest zone, as would be expected given the overall lower stream densities in the East of Crest sample units.

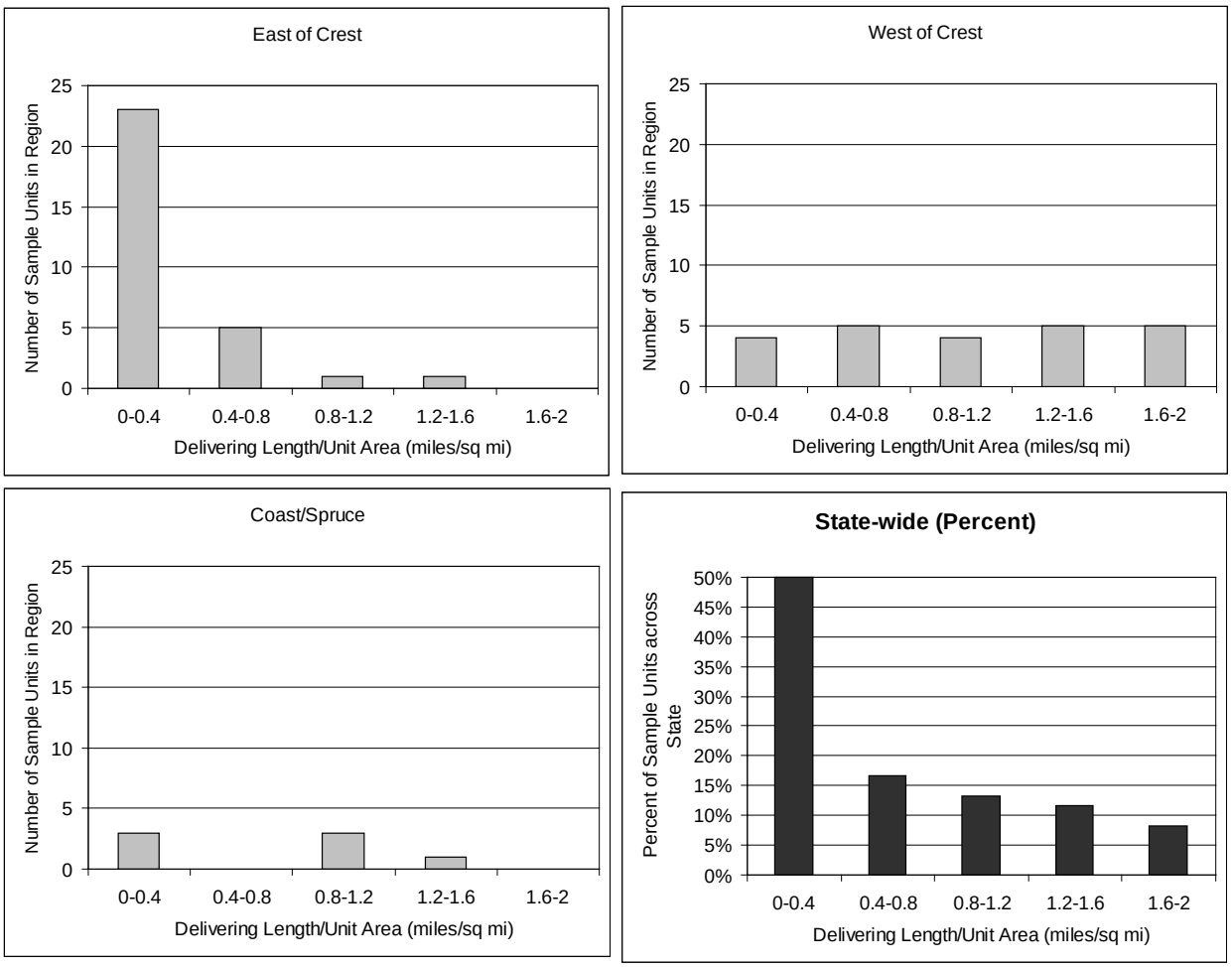


Figure 10. MQ1.1 Total Road Length Delivering to Streams per Unit Area.

MQ 1.2 and H2a2 Percent of road network draining to streams

The percentage of the road network draining to streams was calculated using the un-weighted length of road delivering. Across the state, more than half of the sample units had less than 10

percent of the road network delivering (Figure 11). As with MQ1.1 (length of road delivering), units East of the Cascade Crest had the lowest percent of the road network delivering to streams and units in the West of Crest region had the highest percent of the road network delivering.

State-wide, an average of 11 percent of the surveyed road length delivered to streams, either directly, via a gully, or partially through an indirect sediment plume pathway across the forest floor. The percent of connected road network can vary greatly among watersheds based on values reported in other studies, although it should be noted that the survey methods and definitions of road connectivity between different studies make direct comparison of results complex. Martin (2009) reports an average of 12 percent of the surveyed road length on private timberlands was hydrologically connected in a landowner-conducted survey of 1,047 miles of road throughout Washington. Watershed Analyses conducted in the 1990's on timberlands owned by Boise Cascade ranged from 3 to 57 percent connectivity, with watersheds in drier, inland areas of Washington, Oregon, and Idaho having 3 to 17 percent connectivity and watersheds in wetter areas (e.g., Coast Range of Oregon) having 45 to 57 percent connectivity (Domoni Glass, personal communication). A survey of roads in the Sierra Nevada Mountains of California by Coe (2006) found 25 percent connectivity, and a survey in the Oregon Cascades by Wemple et al. (1996) reports 57 percent connectivity (although 24 percent of this was connected via gully without confirmation that the gully reached a stream).

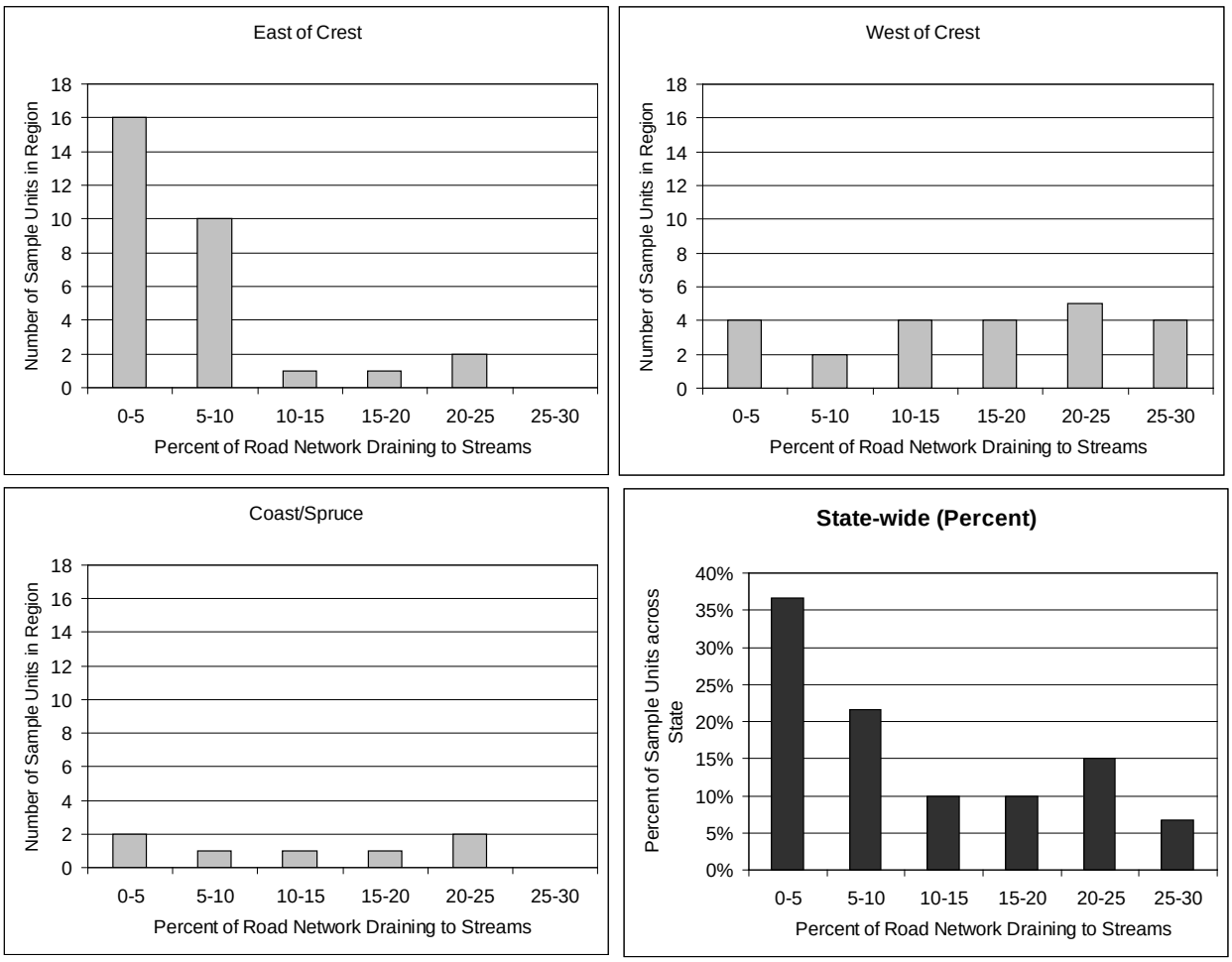


Figure 11. MQ1.2 Percent of Road Network Draining to Streams.

MQ 1.3 Percent of road in each surface category

The type and quality of road surfacing has an effect on how well the road holds up to traffic use and the amount of sediment produced from the road surface. Good quality, durable gravel surfacing holds up better and produces less sediment than native-surfaced roads. Poor quality or pitrun gravel produces an intermediate amount of sediment.

Surfacing on inventoried roads was classified into one of four categories: Asphalt, good gravel, pitrun/worn gravel, and native (unsurfaced). Since gravel surfacing breaks down with traffic use, there is a continuum from good gravel to worn gravel. This phenomenon is particularly evident on heavily traveled mainline gravel roads. Good gravel was differentiated from pitrun by visually determining if the gravel particles were clearly visible or if the gravel particles were embedded in fines. For consistency, at the start of each sample unit field crews assessed the condition of surfacing on the roads and agreed upon the surface type.

Surfacing in the East of Crest units was primarily native; pitrun/worn gravel dominated Coast/Spruce and West of Crest roads (Figure 12).

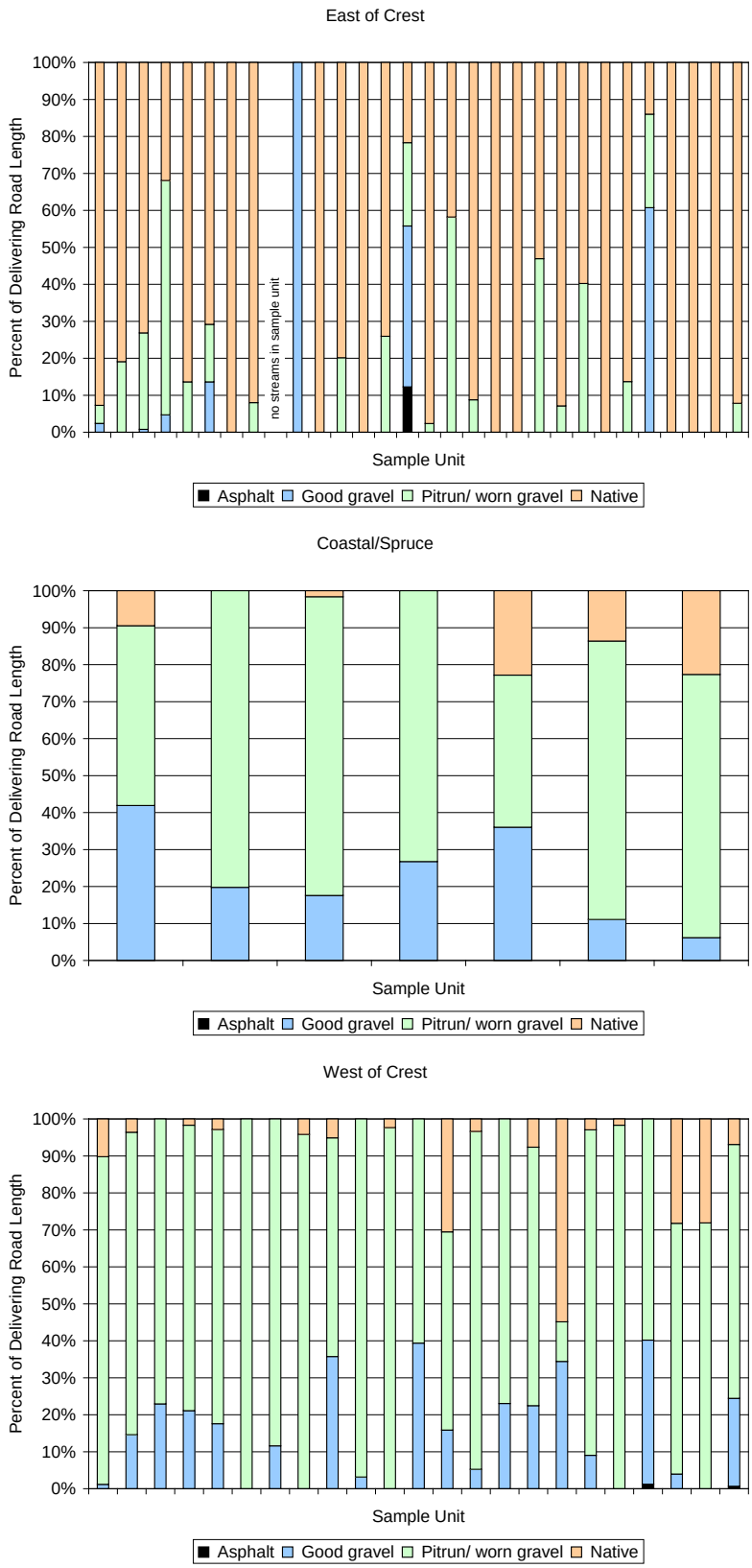


Figure 12. MQ1.3 Percent of Delivering Road Network by Surfacing.
This information is also included in numerical format in Table 7.

MQ 1.4 Percent of road in each traffic category

Traffic on unpaved roads has been shown to increase sediment production as the road surface is broken down into fine-grained particles that can be eroded by surface runoff, and passage by truck tires produces a pumping action that brings fines to the surface.

Landowners reported the daily average traffic use over the past year on their roads within each sample unit. Traffic levels were grouped into categories based on WARSEM traffic categories (none, occasional, light, moderate, moderately heavy, heavy, and very heavy). Traffic use on most of the road network sampled across the state was occasional, with a daily average of less than 1 log truck/day (Figure 13). Several of the units in each of the geographic areas had some roads with light use (1-2 log trucks/day) or moderate use (3-4 log trucks/day). Two of the units in the Coast/Spruce zone had roads with heavy traffic (4-5 log trucks/day); these units included mainline roads in areas of recent active hauling. Traffic use in each sample unit is open to change between sample periods as traffic patterns shift to accommodate haul from different harvest units through time.

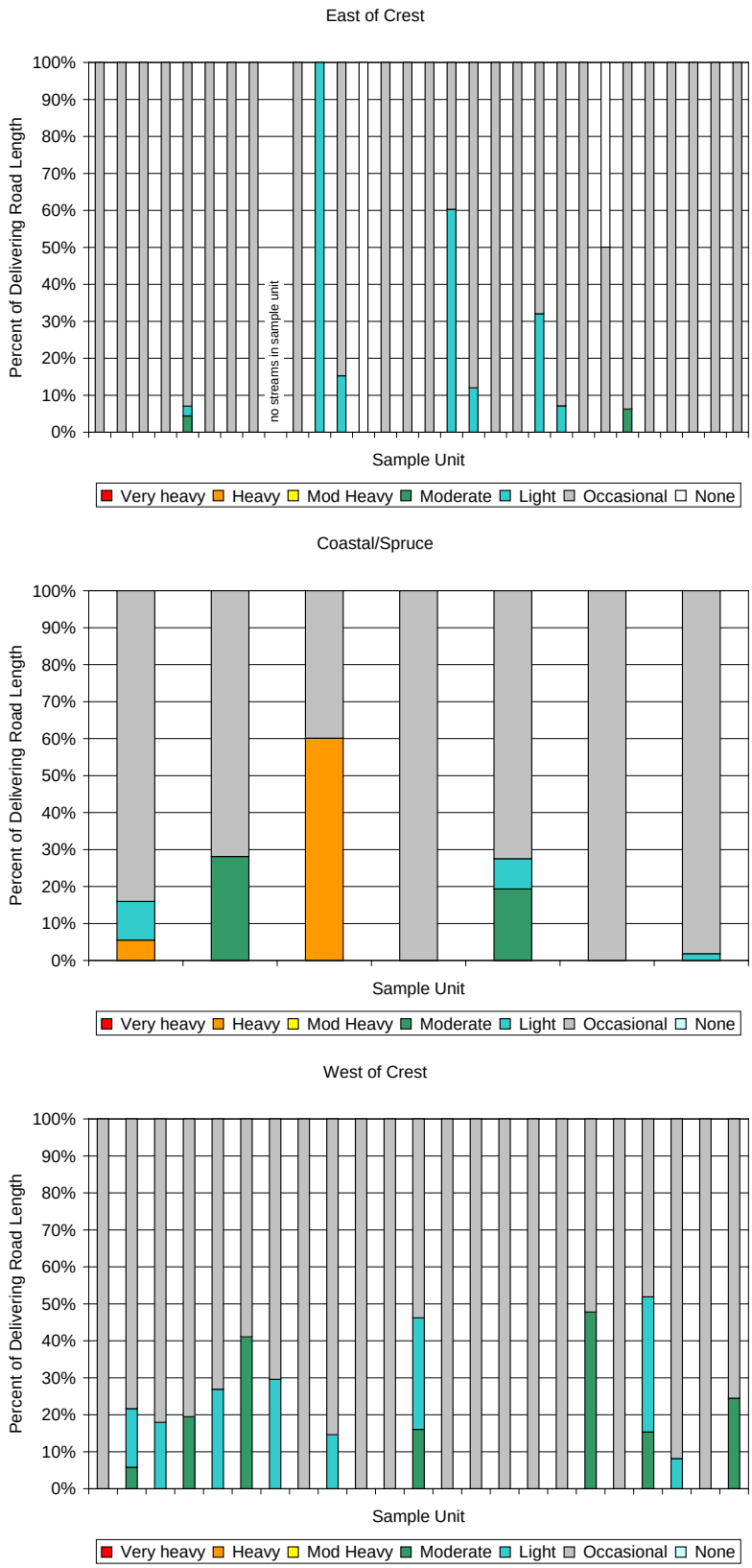


Figure 13. MQ1.4 Percent of Road Network by Traffic Level.
This information is also included in numerical format in Table 7.

MQ 1.5 Percent of road in each cutslope cover category

The percent of rock or vegetative cover on cutslopes influences the erosion rates on the cutslope. Cutslope cover data was collected in 10% cover increments on all delivering road segments. Higher cover categories would be expected to have less erosion from the cutslope. Some road segments had no cutslope; a cover rating of 'n/a' was used for these segments.

Cutslope cover varied widely, but in general there was more cover on West of Crest and Coast/Spruce road segments than on East of Crest segments (Figure 14). This is expected due to denser vegetation coverage from higher precipitation rates on the West side of the state.

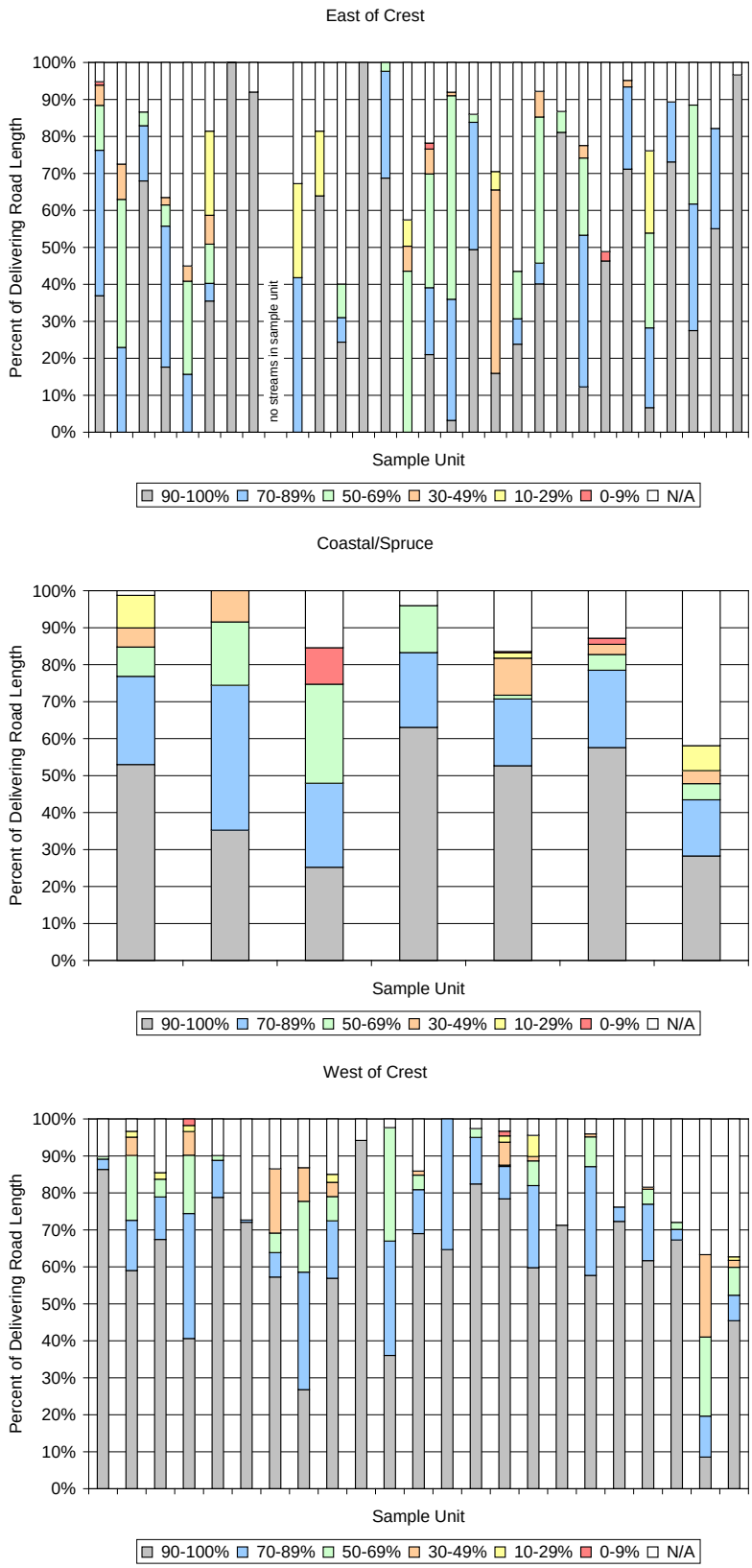


Figure 14. MQ1.5 Percent of Road Network by Cutslope Cover Category.
This information is also included in numerical format in Table 7.

MQ 1.6 Percent of drainage points by connectivity category

Forest roads can alter surface runoff patterns in a watershed by intercepting groundwater via the cutslope or collecting runoff on the relatively impermeable tread. If the drainage from the road prism enters a stream or wetland, it can route water and sediment to the stream at a different rate than without the road in place.

Road drainage connectivity was assessed by assigning each road segment into one of six delivery categories:

- No delivery (no data was collected on these segments)
- Drains directly into stream channel/typed wetland
- Drains directly into stream channel/typed wetland via a gully
- 35% delivery to a stream/typed wetland
- 10% delivery to a stream/typed wetland
- No evidence of delivery, but road parallels within 20 feet of a stream/wetland

The percent of delivering points in each delivery category varied widely between sample units (Figure 15). The location of roads in relationship to streams in each sample unit played a large role in delivery, as did grading and placement of cross-drain culverts and drainage dips near streams.

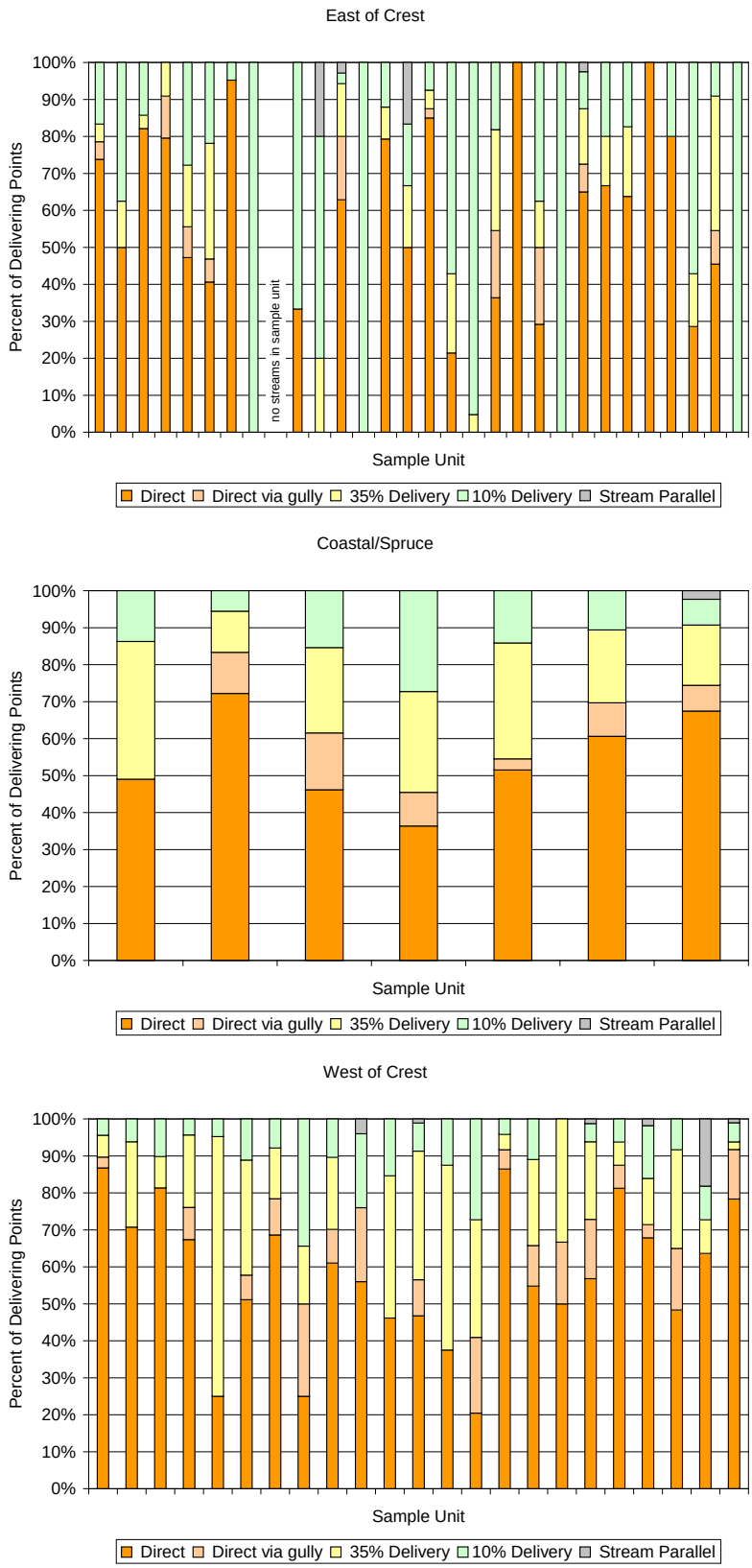


Figure 15. MQ1.6 Percent of Drainage Points by Connectivity Class.
This information is also included in numerical format in Table 7.

MQ 1.7 Percent of road in each road rutting category

Ruts in the road surface can result in increased erosion by channeling drainage into a concentrated flow path. In addition, ruts can change the intended drainage pathway on the road surface by capturing water on an outsloped or crowned road and directing it down the road tread to a stream instead of allowing it to disperse across the fillslope. Ruts can be formed by traffic on soft road surfaces, such as saturated native surfaced roads, or by concentrated water running down long, steep (over 10%) sections of road. For the present study, ruts were defined as wheel indentations over two inches deep and classified into three categories: no ruts, ruts not interfering with tread drainage, and ruts interfering with drainage.

Over half of the sample units had no ruts on delivering road segments, and most of the units had no ruts on over 90% of the delivering road length (Figure 16). A few units in the East of Crest and one in the Coast/Spruce regions had ruts on up to 40% of the delivering road length.

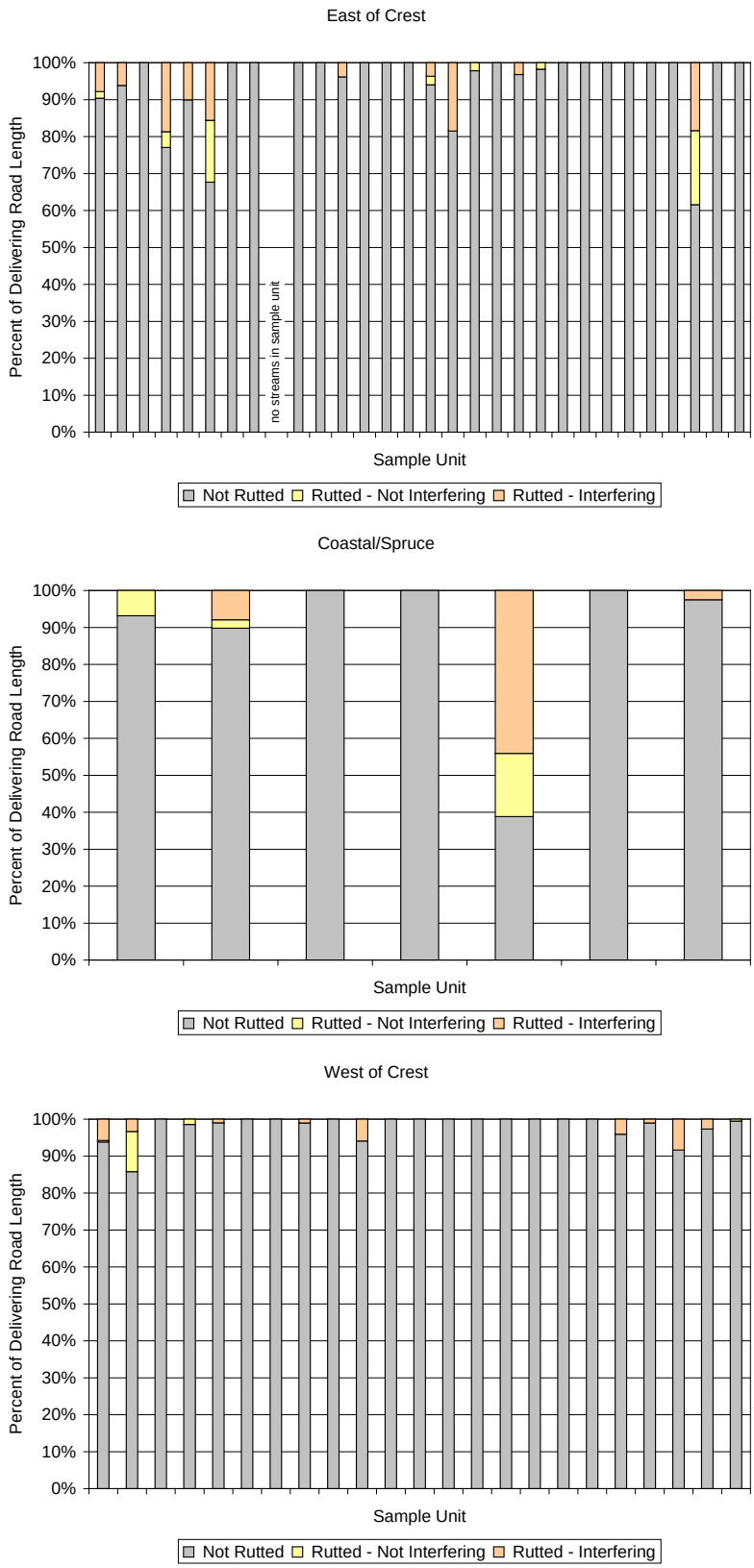


Figure 16. MQ1.7 Percent of Delivering Road Length by Rutting Category.
This information is also included in numerical format in Table 7.

4.3.2 Monitoring Question 2

Have road attributes that affect sediment production and delivery improved over time?

Note that this question will be answered through comparison with future sample event results; data for the first sample event are presented here.

H2b1 Road surfacing index

The road surfacing index is a combined/weighted metric that is computed by multiplying the length of road in each road surfacing class by the WARSEM numerical factor for that class and dividing the sum of all classes by the total length x lowest numerical WARSEM factor for road surfacing. A lower road surfacing index indicates that relatively less erosion would come from roads in that unit based on the distribution of surfacing in the unit.

Units in the East of Crest region had higher surfacing indices than units in the Coast/Spruce or West of Crest regions. This distribution reflects the higher percentage of roads with native surfaces in East side units (Figure 17).

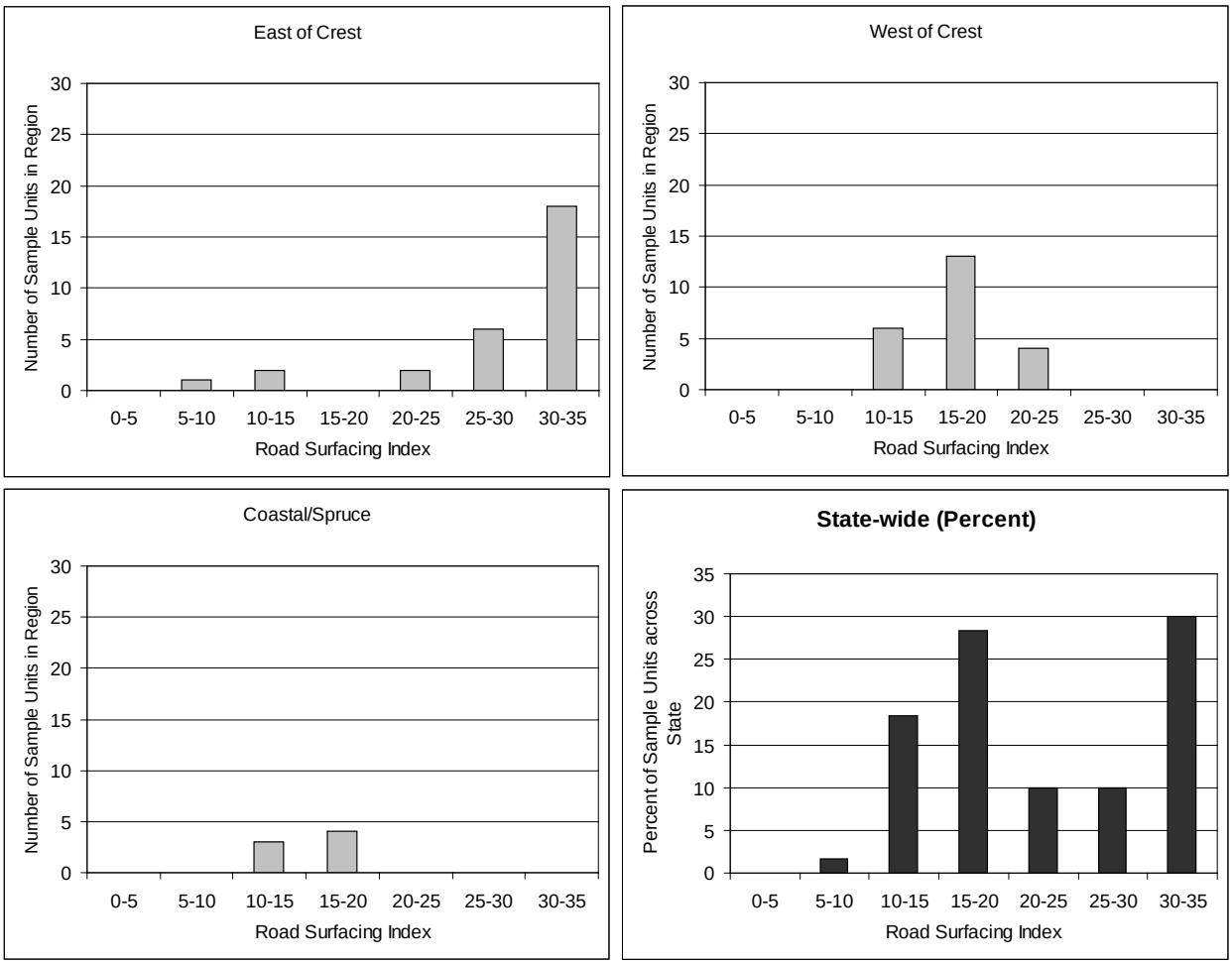


Figure 17. H2b1 Road Surfacing Index.

H2b2 Road traffic index

The road traffic index is a combined/weighted metric that is computed by multiplying the length of road in each road traffic class by the WARSEM numerical factor for that class and dividing the sum of all classes by the total length x lowest numerical WARSEM factor for traffic. A lower road traffic index indicates that relatively less erosion would come from roads in that unit based on the distribution of traffic use in the unit.

Units in the East of Crest region had lower traffic indices than units in the Coast/Spruce or West of Crest regions (Figure 18). One unit in the Coast/Spruce zone had very heavy traffic on a mainline with many delivery points.

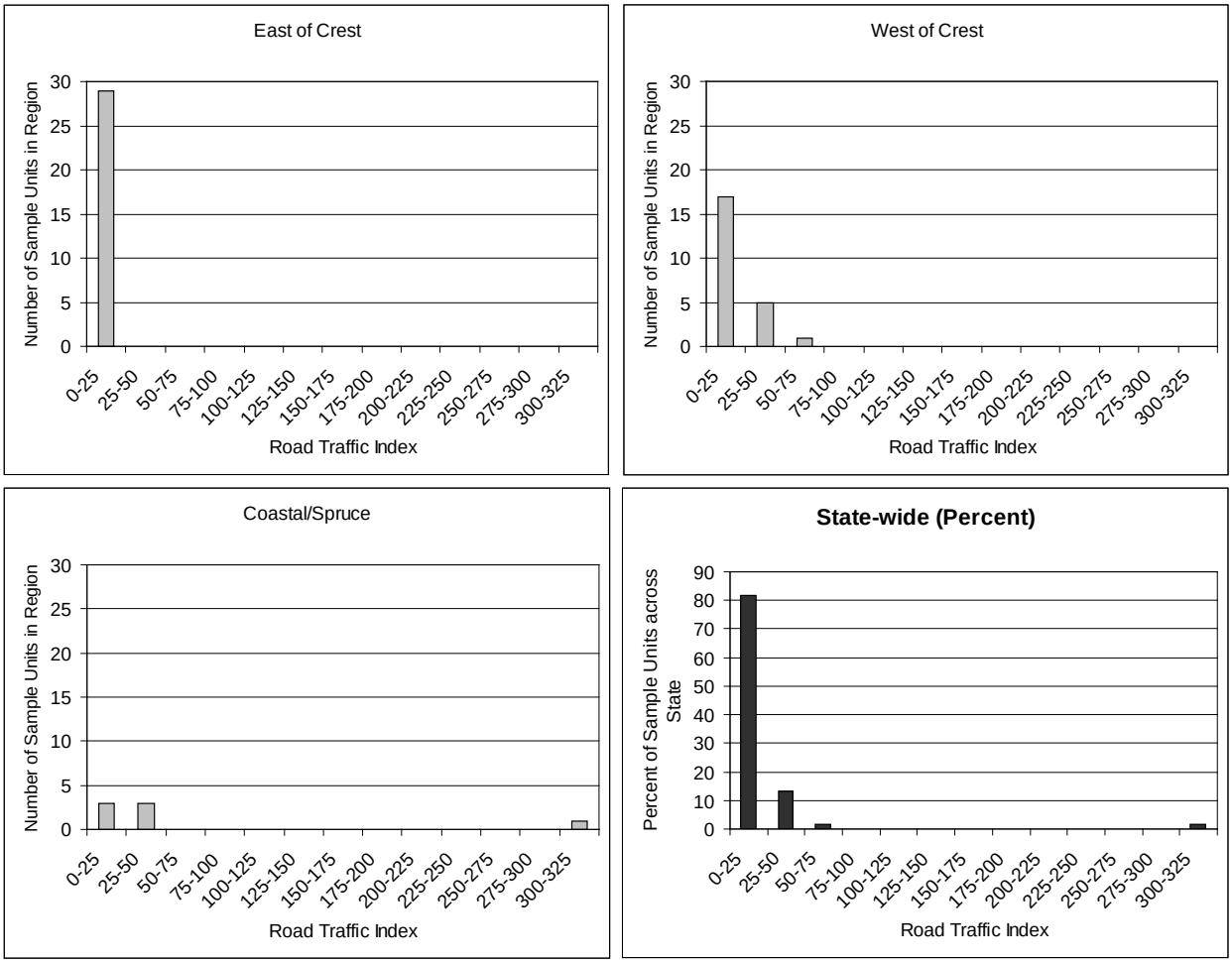


Figure 18. H2b2 Road Traffic Index.

H2b3 Cutslope cover index

The cutslope cover index is a combined/weighted metric that is computed by multiplying the length of road in each cutslope cover class by the WARSEM numerical factor for that class and dividing the sum of all classes by the total length x lowest numerical WARSEM factor for

cutslope cover. A lower cutslope cover index indicates that relatively less erosion would come from roads in that unit based on the distribution of cutslope cover in the unit.

Cutslope cover indices are relatively evenly distributed between 0.5-3 in East of Crest units, with a slight skew toward lower indices in Coast/Spruce and West of Crest units (Figure 19).

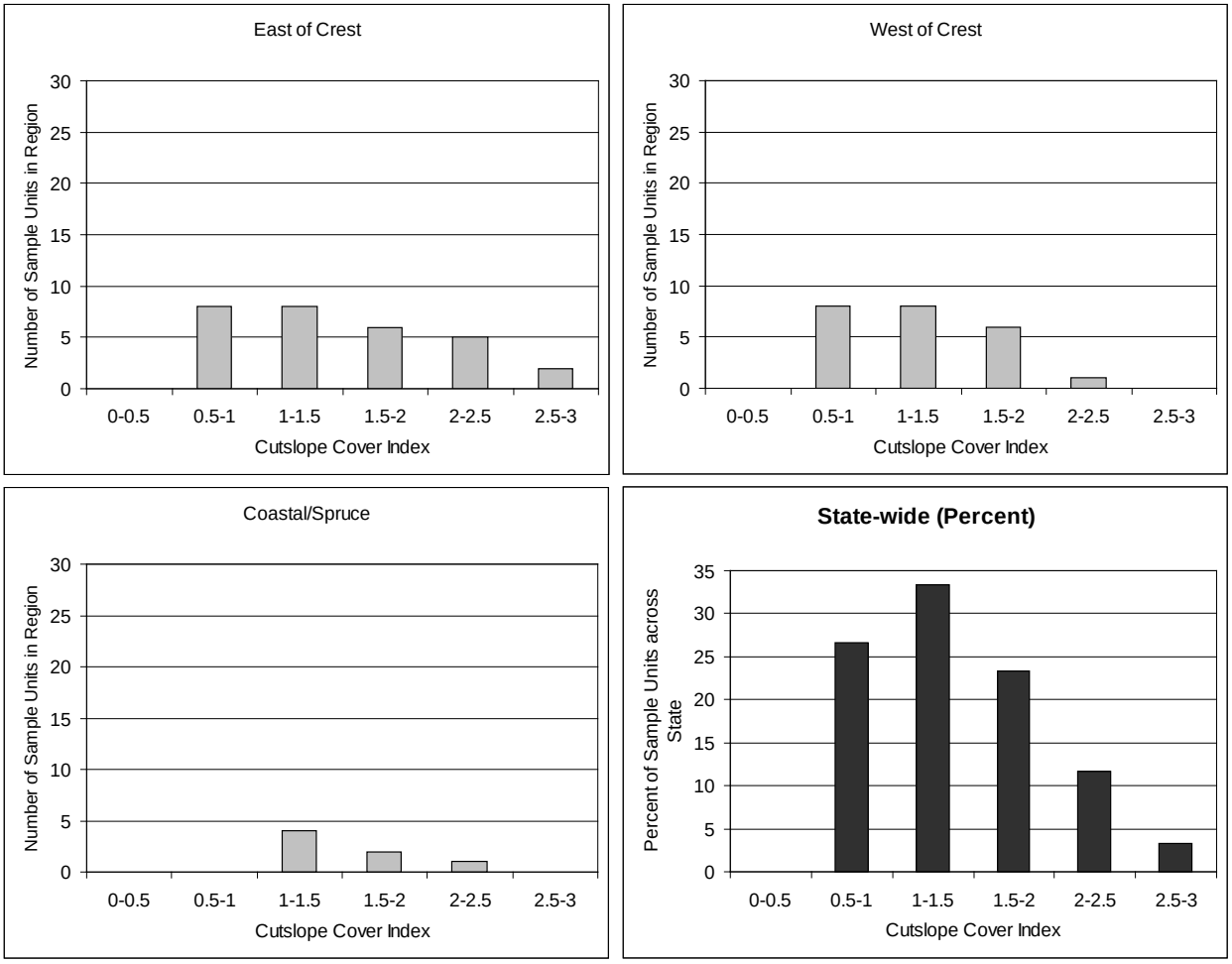


Figure 19. H2b3 Cutslope Cover Index.

H2b4 Miles of delivering road with ruts interfering with drainage

The distribution of units by miles of delivering road with ruts that were determined to interfere with drainage is shown in Figure 20. Less than 0.2 miles of roads with interfering ruts were found in the majority of units, although a few units had a relatively large number of rutted roads.

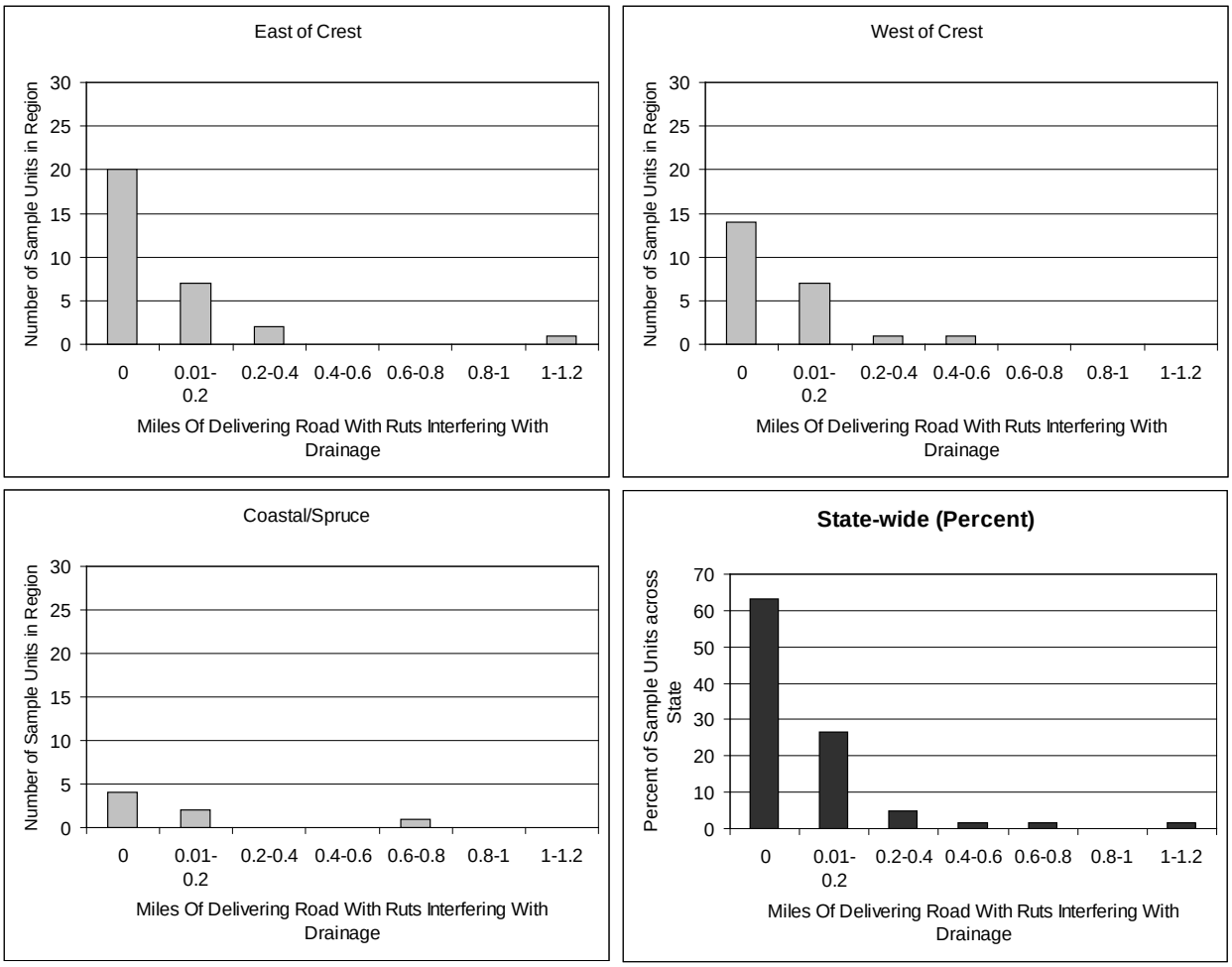


Figure 20. H2b4 Miles of Delivering Road with Ruts Interfering with Drainage.

4.3.3 Monitoring Question 3

What is the status of road performance measures for drainage connectivity and sediment delivery to streams at each sample event?

MQ 3.1 and H5a Miles of forest road delivering to streams per miles of stream (road hydrology performance measure)

Schedule L-1 of the Forests & Fish Report provides two performance measures with targets for existing roads. The first is a measure of the hydrologic connectivity of the road system in relation to the stream density in a sample unit (RLEN), and was calculated as the miles of road delivering to streams per miles of stream in each sample unit. The higher this number is, the more road network drains to streams in the unit. Sample units in the East of Crest region had less road delivering/miles of stream than units in the Coast/Spruce or West of Crest regions.

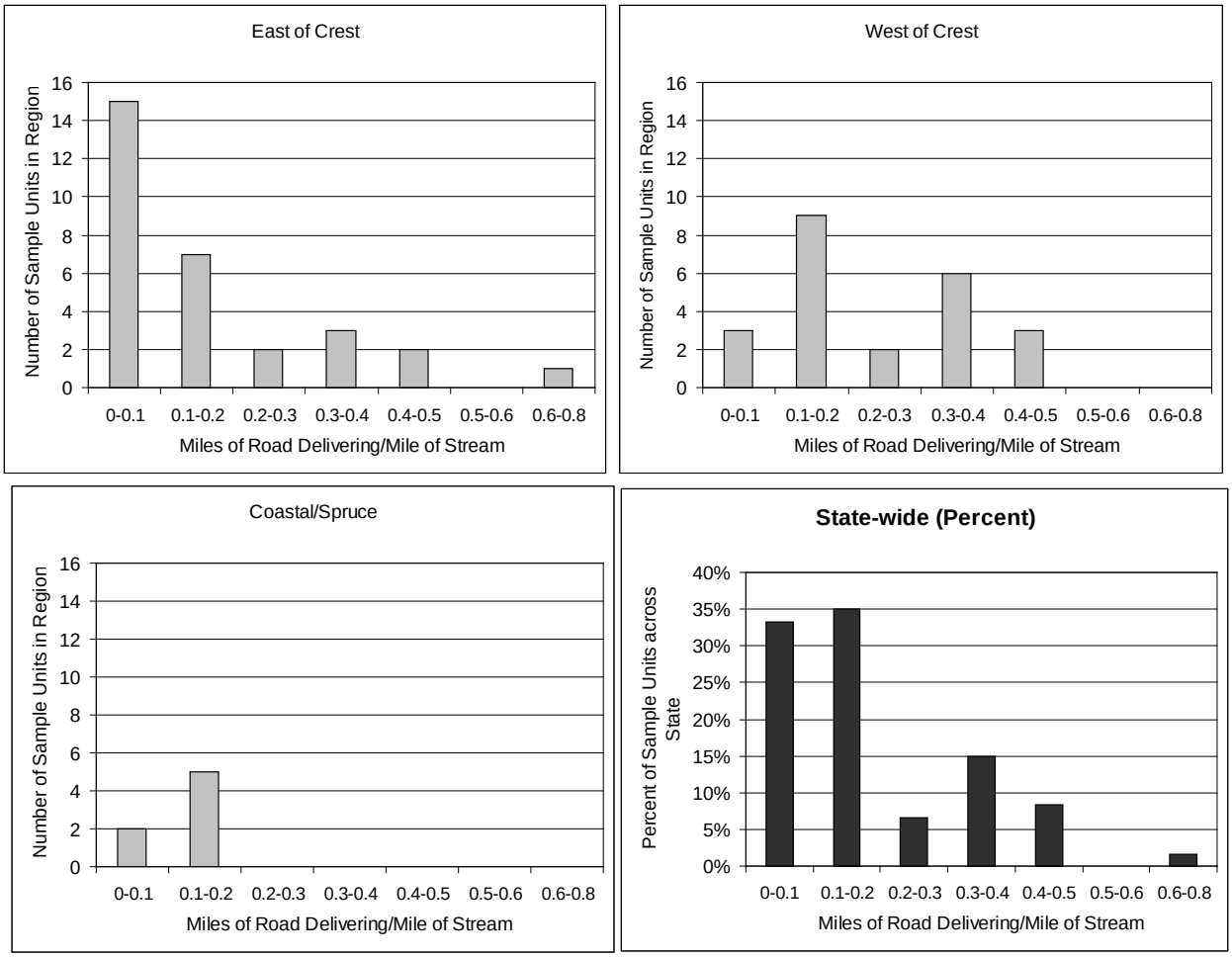


Figure 21. MQ3.1 Miles of Delivering Road per Mile of Stream (FFR Hydrology Performance Measure).

Figure 22 shows the cumulative percentages of sample units as a function of miles of road connected to the stream network per mile of stream (RLEN) for each region and statewide. These graphs can be thought of as counting the number of units (or percentage of study units) that have RLEN values less than or equal to the value on the x-axis. For instance, it can be seen in Figure 22 that 68 percent of the sample units had RLEN values of less than 0.2 mile of delivering road per stream mile. The range of each regional target (upper and lower target) is also shown on the graph so that the percentage of sample units meeting either the upper or lower end of the target range can be determined for each region.

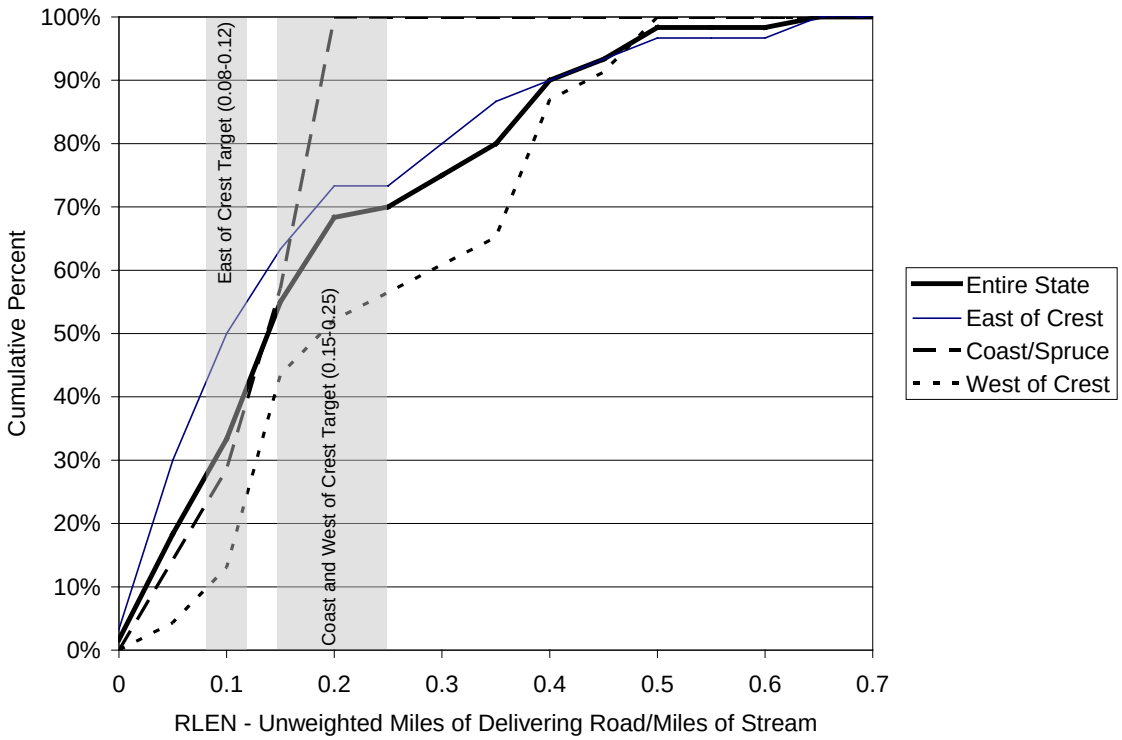


Figure 22. MQ3.1 Cumulative Percent of Sample Units, Miles of Delivering Road per Mile of Stream (RLEN).

MQ 3.2 and H5b WARSEM modeled tons of road sediment delivered to streams per mile of stream (road sediment performance measure)

The other FFR performance measure is tons of road sediment delivered to streams per mile of stream. This metric focuses on the relative amount of sediment delivered to streams in the sample unit. All sample units in the East of Crest region had less than 2 tons of sediment/mile of stream (Figure 23). Most units in the Coast/Spruce zone had less than 6 tons/mile, and units in the West of Crest region varied the most, with up to 16 tons/mile

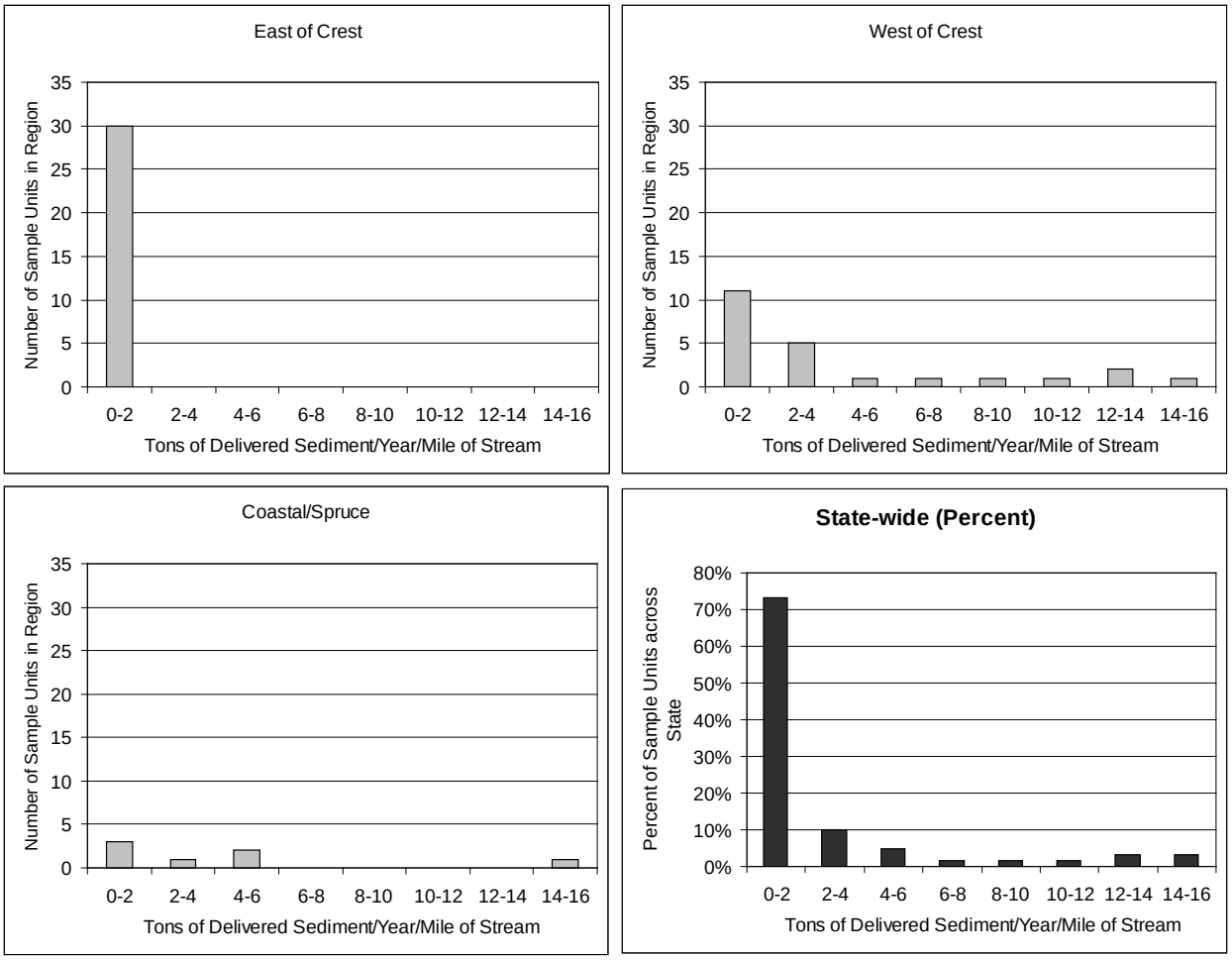


Figure 23. MQ3.1 Sediment Delivered/Year per Mile of Stream (FFR Sediment Performance Measure).

Figure 24 shows the cumulative percentages of sample units as a function of sediment delivered per mile of stream (RSED). The range of each regional target (upper and lower target) is also shown on the graph so the percent of sample units meeting either the upper or lower end of the target range can be determined for each region.

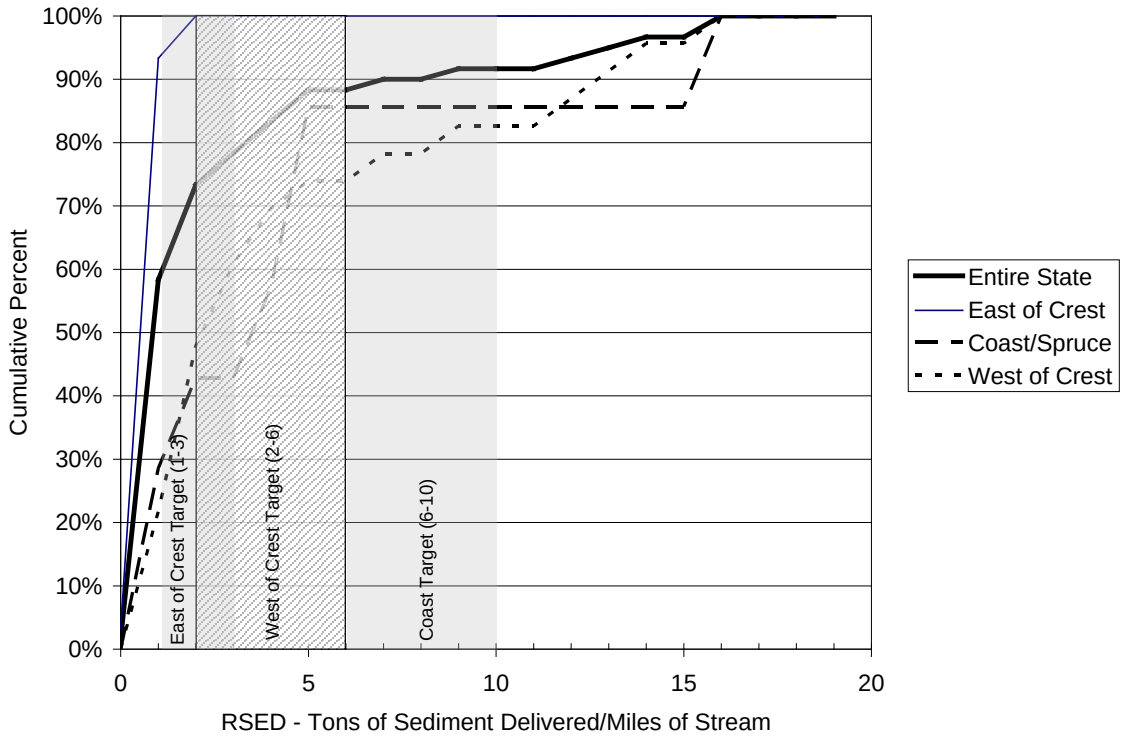


Figure 24. MQ3.1 Cumulative Percent of Sample Units, Tons of Delivered Sediment per Mile of Stream (RSED).

4.3.4 Monitoring Question 4

What is the status of road performance measures relative to their targets, by performance target region, at each sample event?

MQ 4.1 Miles of forest road delivering to streams per miles of stream (road hydrology performance measure) divided by the performance target by target region

Monitoring Question 4.1 looks at the road hydrology performance measure in each unit in relationship to the upper limit of the performance target for that region. If the ratio of the measure/target is one or less, the unit meets the target. If the ratio is greater than one, the unit exceeds the target.

Across the state, 62 percent of the sample units meet the sample hydrology target (Figure 25). The sample unit that far exceeded the standard (East of Crest Unit, 5-6 bin) had a large number of mid-slope roads resulting in numerous stream crossings despite the moderate overall stream density.

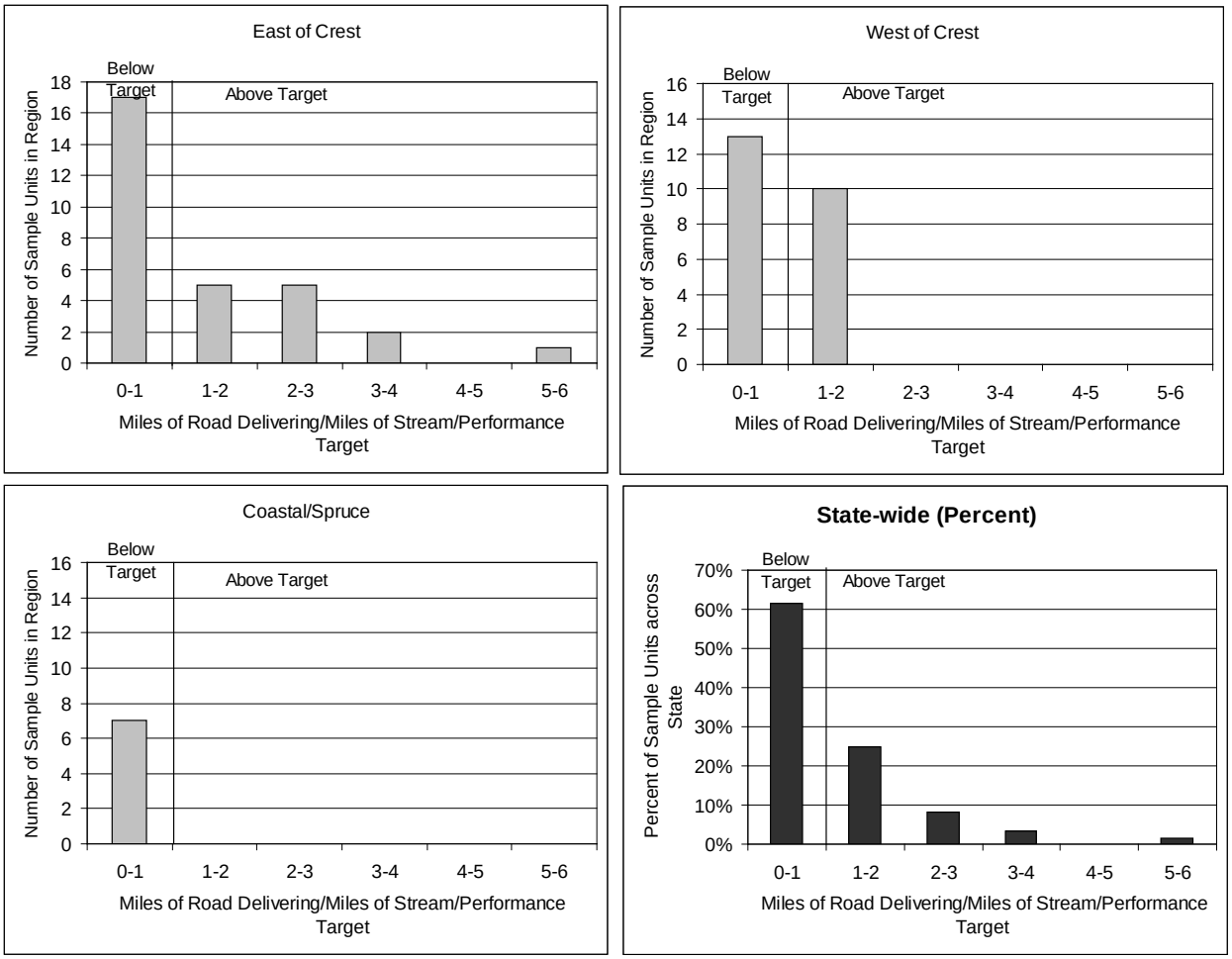


Figure 25. MQ4.1 Miles of Delivering Road per Mile of Stream Divided by Regional Performance Target.

MQ 4.2 WARSEM modeled tons of road sediment delivered to streams per miles of stream (road sediment performance measure) divided by the performance target by target region

Monitoring Question 4.2 looks at the road sediment performance measure in each unit in relationship to the upper limit of the performance target for that region. If the ratio of the measure/target is one or less, the unit meets the target. If the ratio is greater than one, the unit exceeds the target.

Across the state, 88 percent of the sample units meet the sediment target (Figure 26). All sample units in the East of Crest region meet the target. One unit in the Coast/Spruce region and several in the West of Crest region do not meet the target.

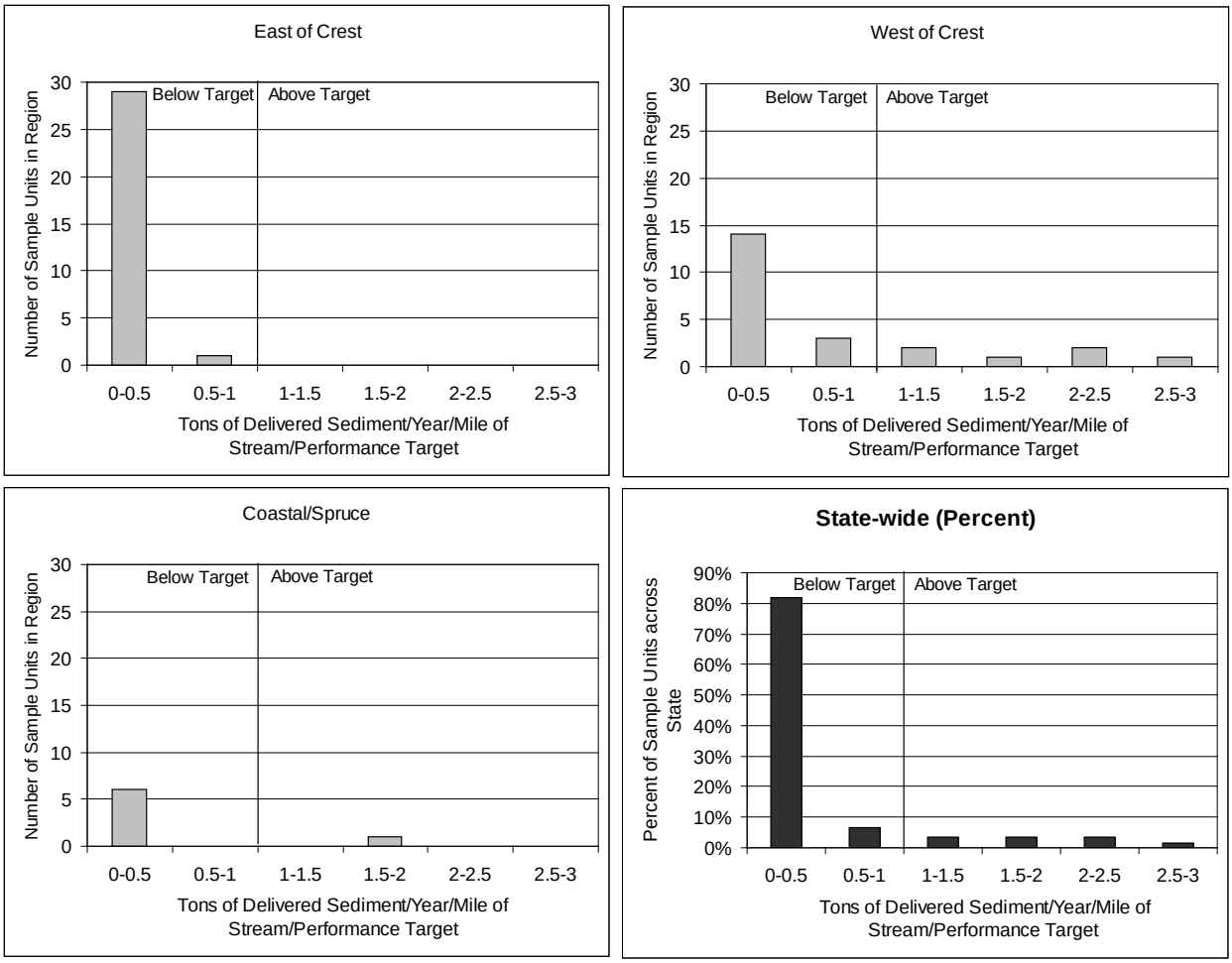


Figure 26. MQ4.2 Sediment Delivered/Year per Mile of Stream Divided by Regional Performance Target.

4.3.5 Monitoring Question 5

Have measures of road sediment performance improved over time?

This monitoring question will be assessed through comparison of the initial and subsequent sampling events.

4.3.6 Monitoring Question 6

Will roads judged to meet FFR road standards meet the performance targets?

In order to assess how much of the road system in each sample unit either had RMAP work completed or was up to current road standards, each landowner was asked the percentage of roads on their lands in the unit that they consider to be up to standards. This measure was compiled for each sample unit by combining the percent of roads reported to meet standards from each landowner by the percent of the road system on their land in each sample unit. As this metric was based on landowner judgment and was not verified by field crews, it should be regarded as an estimate rather than a precise measurement. Figure 27 shows the percent of

sample units by the percent of roads reported to be up to current road standards for each region and across the entire state. Overall, a high percentage of roads were reported to be up to standards, with over half of the units reporting that at least 85 percent of the roads meet standards. Percent of roads meeting standards was highest in the East of Crest region, and lowest in the West of Crest region.

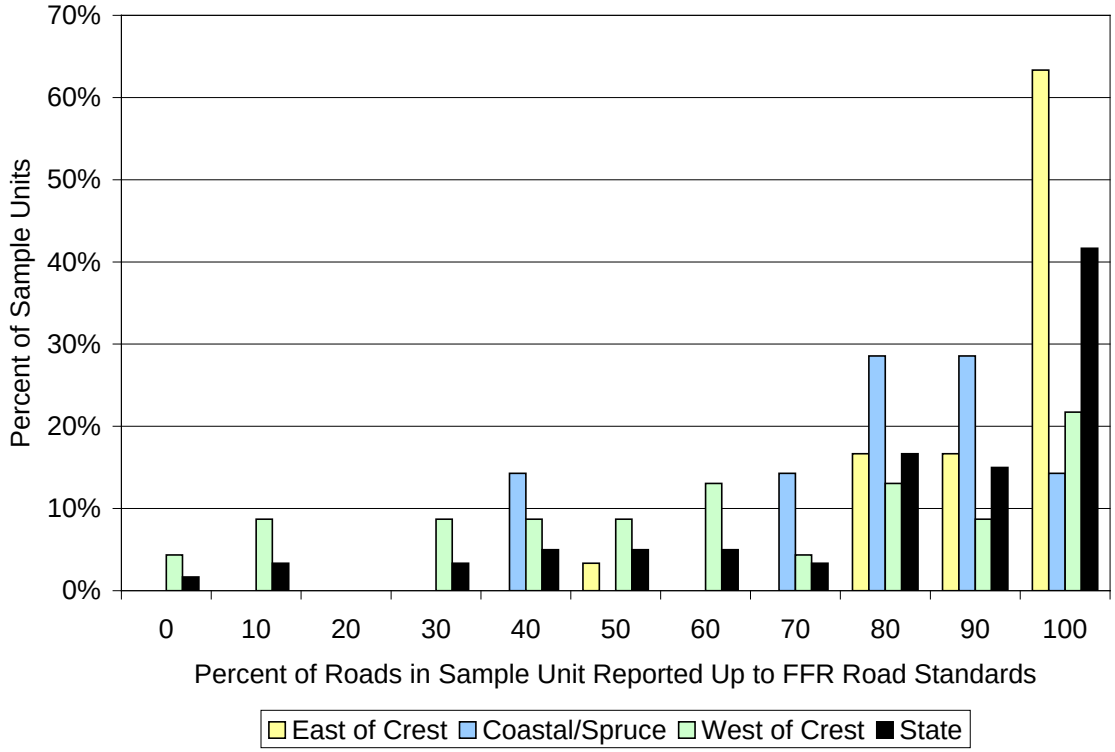


Figure 27. Percent of Sample Units with Percent of Roads Reported to be up to FFR Road Standard

H6a Miles of forest road delivering to streams per miles of stream (road hydrology performance measure) vs. percent of road length in each unit meeting FFR road standards

H6b WARSEM modeled tons of road sediment delivered to streams per miles of stream (road sediment performance measure vs. percent of road length in each unit meeting FFR road standards

The FFR performance metrics for road length delivering (MQ3.1) and sediment delivery (MQ3.2) were calculated for each sample unit, divided by the FFR performance target for that metric (shown in Table 1) and plotted by the percent of roads within the sample unit that were reported by landowners to meet current road rule standards. This relationship between percent roads meeting performance standards and the ratio of each sample unit performance measure to the regional target are displayed in Figures 28 and 29.

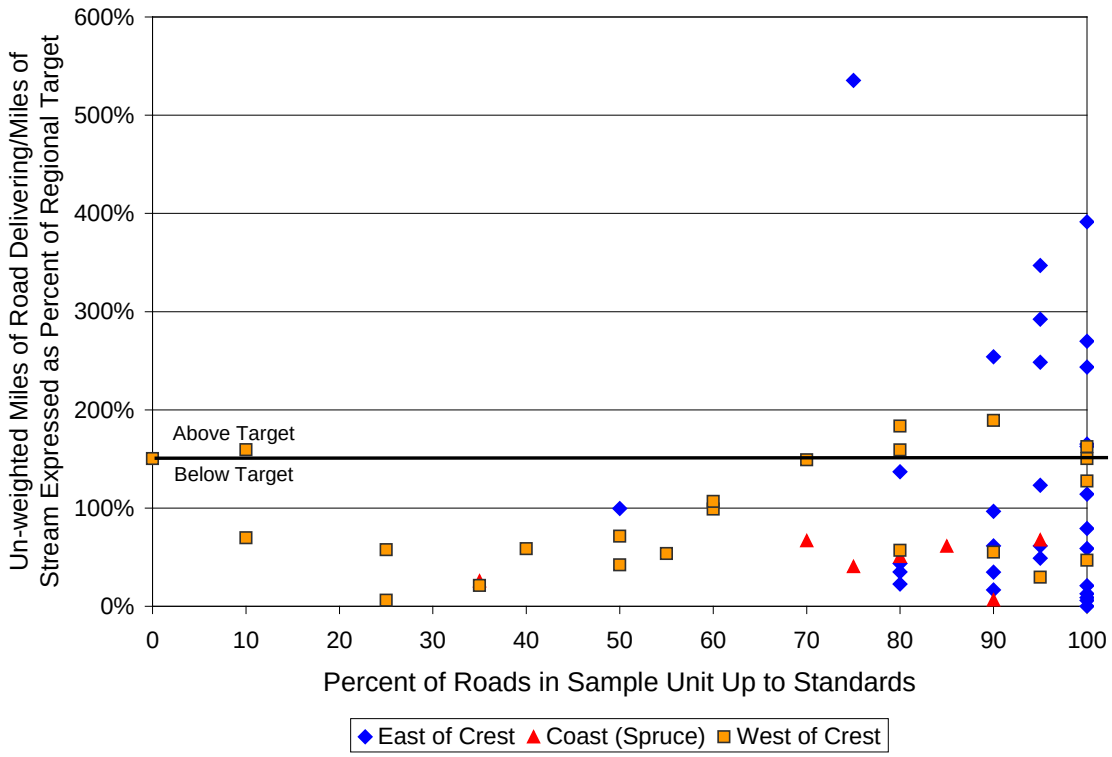


Figure 28. Miles of Delivering Road/Miles of Stream (MQ3.1) Expressed as Percent of Regional Target vs. Percent of Road Miles in Sample Unit Reported to be up to Standards.

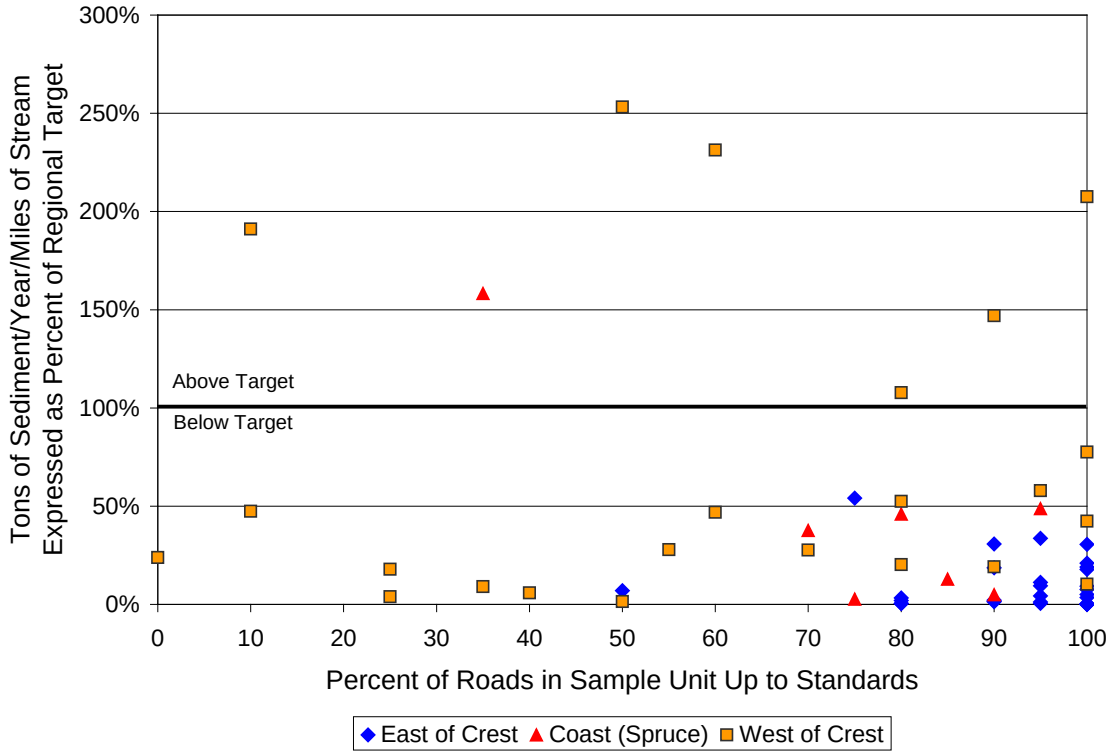


Figure 29. Tons of Delivered Sediment/Miles of Stream (MQ3.2) Expressed as Percent of Regional Target vs. Percent of Road Miles in Sample Unit Reported to be up to Standards.

For MQ3.1, length of road delivering, there is no evidence of a relationship with the reported percent of roads up to standards (one-tailed Kendall's trend test $p = 0.327$). This may reflect a general RMAP objective of prioritizing road maintenance on road sections with the highest need for maintenance/repairs.

All of the Coast/Spruce zone units met the road mile delivering target, but in the East of Crest and West of Crest zones only 57 percent of the units met the road mile delivering target. This was true even though 93 percent of the units reported more than 80 percent of their roads were up to standards. The reported percentage of roads up to standards varied widely for the Coastal and West units. The sample unit that far exceeded the standard (East of Crest Unit, 535 percent of standard) had a large number of mid-slope roads resulting in numerous stream crossings despite the moderate overall stream density.

For MQ3.2, sediment delivered, there is a significant monotonic decrease in sediment delivered with increase in percent of roads meeting performance standards within the sampling unit ($p = 0.0028$). The regional results for sediment delivery are quite different than the road length delivering results, with all of the East of Crest units meeting the sediment performance target, and 74 percent of the West of Crest units and 86 percent of the Coastal/Spruce units meeting the target.

The reasons for the differences in meeting delivering length and sediment targets between the different geographic regions are likely due to a combination of factors, including a much lower stream density on East of Crest units (metrics are divided by miles of stream in each unit), and much lower precipitation on the Eastside resulting in lower calculated sediment production in those units. It is also possible that landowners prioritized roads for RMAP work and upgrades to first work on those areas with roads that they judged to need the most improvements.

It should also be noted that in some of the selected sample unit areas it will be much more difficult to meet the target values as beacause of the location of the existing roads in relation to the stream network. For example, areas with a ridgetop road system (Figure 30 left side) have a much lower density of road/stream crossings (blue squares) and therefore a much lower chance for long lengths of delivering roads than a unit with primarily midslope and valley bottom roads (right side of Figure 30). Moreover, some geologic materials and soils are more erodible than others and/or more prone to gullying.

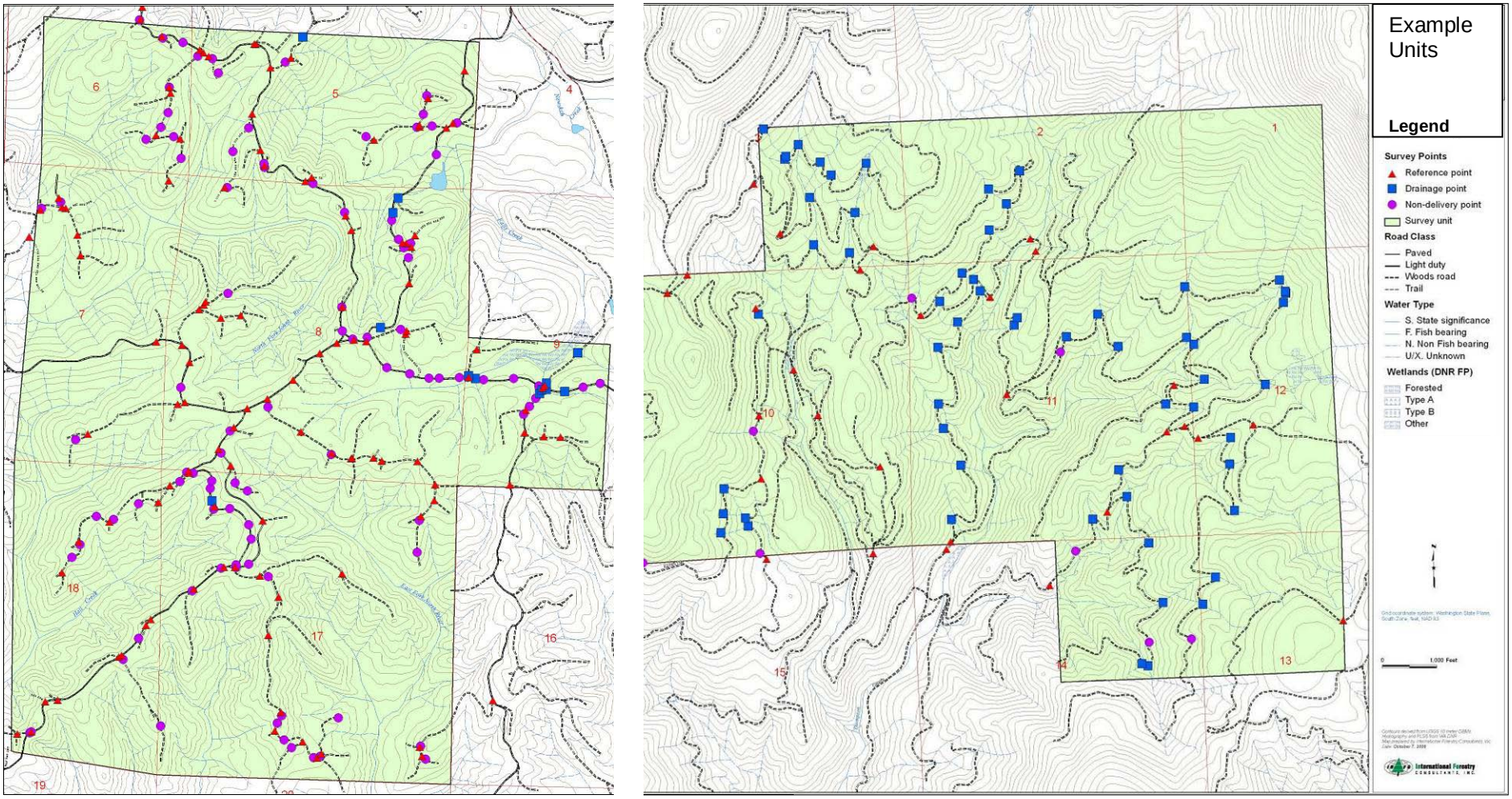


Figure 30. Sample Unit Maps Showing Difference in Delivery Between Units with Ridgetop Roads (left) and Midslope/Valley Bottom Roads (right)

4.4 New Roads Evaluation

While the majority of this study evaluates road effects collectively at the sub-basin scale, the data set also allows site-scale evaluation of new roads. A total of 11.8 miles of new roads, defined as being constructed within two years of the field sampling, were present within 11 of the sample units; these units were distributed among all three rule regions. The FFR performance target for sediment delivered to streams from new roads is "virtually none" (see Table 1 in Section 2 above), which does not lend itself to objective determination of success or failure. Despite the subjective target and small sample size, this evaluation provides an initial assessment of new roads built under post-2000 Forest Practices Rules.

Table 8 summarizes the miles of new road length and estimated sediment delivery from new roads in each of the 11 sample units. Delivery from the new roads varied greatly – five of the units with new roads had no delivery from the new road segments, while between 1 and 79 percent of the new road length in the remaining six units delivered to streams. The specific location of the new roads determined whether or not the road segments were hydrologically connected to streams. In several units, the new roads were a number of short spurs off existing roads and did not cross streams. However, in four of the units the new roads were midslope roads constructed in locations without previous road systems; these crossed many small drainages and did drain to streams. In the two units with the highest percentage of new delivering (72 and 79 percent), it did not appear that measures were taken to minimize the length of new road connected. The average length of road connected at drainage structures in these two units was 838 and 942 feet respectively. It did appear that an effort was made to minimize connectivity in the other units, with the average length of connected road at each crossing less than 250 feet in all but one of the remaining units. Across all the sample units, 17 percent of the new roads were hydrologically connected compared to 11 percent of all roads sampled across the state.

Sediment delivery from the new road segments was relatively minimal in eight of the 11 units (Table 8). Sediment delivery ranged from 0 to 136.5 tons/year. Note that the WARSEM computation of sediment delivery for new roads includes a factor for road age, with 10 times more sediment produced from a road constructed within the past year as an old road (see WARSEM technical documentation, Dubé et al. 2004). Therefore, the average of 18 tons/mile/year of sediment computed from new roads cannot be directly compared with the average of 2.1 tons/mile/year of sediment produced from existing roads across the state. Measurements of sediment produced from newly constructed roads shows erosion rates decline quickly and reach rates similar to older roads within 2-3 years (Ketchenson et al. 1999, Luce and Black, 1999, Grace 1999, Swift 1984, Dryness 1975, Megahan 1974, Megahan and Kidd 1972).

Table 8. Summary of New Roads by Sample Unit

Region	Miles of New Road in Unit	Un-weighted New Road Length Delivering (mi)	Percent of New Road Length Delivering	WARSEM Calculated Delivered Sediment (tons/yr)	Delivered Sediment per mile of new road (tons/mile/yr)
East of Crest	0.86	0	0%	0	0
	1.54	0.02	1%	0.04	0.03
Coastal Spruce	1.59	0.17	11%	17.3	10.9
	0.40	0	0%	0	0
	2.79	0	0%	0	0
	0.23	0.18	79%	4.9	21.5
	1.04	0.37	35%	51.2	49.4
West of Crest	0.44	0.32	72%	1.8	4.1
	0.43	0	0%	0	0
	2.29	1.00	44%	136.5	59.7
	0.19	0	0%	0	0
Total for New Roads	11.8	2.1	17%	211.8	18.0
Total for All Roads Evaluated Across State (new and old)	1283	147	11%	2,699	2.1

5 Summary and Recommendations for Future Sample Events

New Forest Practice Rules for forest roads under WDNR jurisdiction were adopted in Washington State in 2001. Implementation of the road rules for existing roads is planned over a 15-year time horizon through the completion of Road Maintenance and Abandonment Plans. The objectives of the Road Sub-Basin Scale Effectiveness Monitoring Program are:

- to determine if the road characteristics that affect runoff and sediment delivery to streams are improving through time as RMAP are implemented between 2001 and 2016; and
- to determine the extent to which roads on lands subject to WDNR forest practice rules meet the FFR performance targets (Table 1).

Characteristics of forest roads in a total of 60 four-square-mile stratified random sample units across the state were inventoried between 2006 and 2008, five to seven years after the forest practice rules were adopted. This is the first sample event planned for the Road Sub-Basin Scale Effectiveness Monitoring Program, and provides the first look at the current status of forest roads.

A high percentage of roads in the sample units across the state were reported to either have RMAP work complete, or already be up to current road rule standards, with over half of the sample units reported to have at least 85 percent of road length meeting standards. An average of 11 percent of the road network delivered to streams or wetlands. Across the state, 62 percent

of the sample units met the Forests & Fish Report (FFR) hydrology performance target (miles of delivering road/miles of stream) and 88 percent of the units met the FFR sediment performance target (tons of delivered sediment/year/miles of stream). These results include new and existing roads combined.

Statewide, no relationship was found between length of road delivering and percent of unit reported to be up to standards; however, a statistically significant decreasing relationship was found between sediment delivery and percent of roads up to standards. These findings suggest that, while there may be locations where there is higher hydrologic connectivity or sediment input from roads, overall across the state there is a trend of decreasing sediment input as roads are brought up to modern standards.

Based on the findings of this first sample period, the following recommendations are made for future monitoring events:

- Continue with the next planned monitoring period (planned interval of 5 years) to confirm the findings that, in general, roads conditions improve and they meet the performance metrics as they are brought up to FFR standards.

- Continue to stress measures to minimize observer variability. To help reduce observer variance, minimize the number of field crew members to the extent practical. Five crew members were used during this first sample event. Training, duplicate crew surveys, and continued training through the field season are important to help minimize variance due to observer error.

- Re-evaluate the most effective method to conduct observer variability testing prior to the next sampling period. One addition to the existing methods that may prove helpful is to have the project leader survey the test road segments to use as a benchmark for observer variability. This could point out which attributes are most prone to variability, indicating that more training or guidance is needed.

- As specified in the Study Design, the same study units are planned to be re-sampled during subsequent sample events. Prior to the next sampling period, methods for re-sampling units (Field Protocol, Section 4.11) should be thoroughly discussed to specify how field crews will determine if each road attribute has changed (or not) between sampling events.

- Prior to the next planned sample event, methods for sampling should be discussed with a statistician regarding variables that will be or will not be re-sampled for each unit. Also, the method for analyzing changes should be carefully considered prior to sampling.

- Consider developing a companion study specifically designed to obtain access to and evaluate road conditions on small forest ownerships.

6 References

Bilby, R.E., K. Sullivan and S.H. Duncan, 1989. The Generation and Fate of Road-surface Sediment in Forested Watersheds in Southwestern Washington. Forest Science Vol. 35, No. 2, pp. 453-468.

Bohle, Todd, 2010, personal communication via e-mail.

Brake, D., M. Molnau, and J.G. King, 1997. Sediment Transport Distances and Culvert Spacings on Logging Roads within the Oregon Coast Mountain Range. Presented at the 1997 Annual International ASAE Meeting Minneapolis, MN. Paper No. IM-975018.

Burroughs, E.R. Jr., and J.G. King, 1989. Reduction of Soil Erosion on Forest Roads. General Technical Report INT-264. USDA Forest Service, Intermountain Research Station, Ogden, Utah.

Coe, D., 2006. Sediment Production and Delivery from Forest Roads in the Sierra Nevada, California. Master of Science Thesis, Colorado State University, Spring 2006.

Conover, W.J., 1980. Practical Nonparametric Statistics. Second Edition. John Wiley and Sons, Inc., New York, New York.

Dubé, K., C. Luce, T. Black, and M. Riedel, in preparation. Comparison of Road Surface Erosion Models with Measured Road Surface Erosion Rates. Paper prepared for the National Council for Air and Stream Improvements, December 11, 2008.

Dubé, K., W. Megahan, and M. McCalmon, 2004. Washington Road Surface Erosion Model. Prepared for the Washington Department of Natural Resources. February 20, 2004. available online at http://www.dnr.wa.gov/BusinessPermits/Topics/ForestPracticesApplications/Pages/fp_warsem.aspx

Dryness, C.T., 1975. Grass-legume Mixtures for Erosion Control along Forest Roads in Western Oregon. Journal of Soil and Water Conservation, July-Aug. 1975, pp. 169-173.

Foltz, R.B., 1996. Traffic and No-Traffic on an Aggregate Surfaced Road: Sediment Production Differences. Paper presented at Food and Agriculture Organization Seminar on Environmentally Sound Forest Road and Wood Transport, Sinaia, Romania, June 17-22, 1996.

Foltz, R. B., and E.R. Burroughs. Jr., 1990. Sediment Production from Forest Roads with Wheel Ruts. In: Watershed Planning and Analysis; Proceedings of a Symposium, July 9-11 1989, Durango CO. ASCE, pp. 266-275.

Glass, Domoni, 2001, personal communication. File transmitted by diskette.

Grace, J.M., III. 1999. Erosion Control Techniques on Forest Road Cutslopes and Fillslopes in North Alabama. In: Seventh International Conference on Low Volume Roads. Volume 2; Transportation Research Record 1652; pp. 227-234.

Haupt, H.F., and W.J. Kidd, Jr., 1965. Good Logging Practices Reduce Sedimentation in Central Idaho. Journal of Forestry, Vol. 63, pp. 664-670

Haupt, H.F., 1959. Road and Slope Characteristics Affecting Sediment Movement from Logging Roads. Journal of Forestry, Vol. 57, pp. 329-332.

Ketcheson, G.L. and W.F Megahan, 1996. Sediment Production and Downslope Sediment Transport from Forest Roads in Granitic Watersheds. USDA Forest Service, Intermountain Research Station. Research Paper INT-RP-486.

Ketcheson, G.L., W.F. Megahan, and J.G. King, 1999. "R1-R4" and "BOISED" Sediment Prediction Model Tests using Forest Roads in Granitics. Journal of the American Water Resources Association Vol. 35, No. 1, pp. 83-98.

Kochenderfer, J.N., and J.D. Helvey, 1987. Using Gravel to Reduce Soil Losses from Minimum Standard Forest Roads. Journal of Soil and Water Conservation, Vol. 42 pp. 46-50.

Luce, C.H., and T.A. Black, 1999. Spatial and Temporal Patterns in Erosion from Forest Roads. In: M.S. Wigmosta and S.J. Burges, eds. The Influence of Land Use on the Hydrologic-Geomorphic Responses of Watersheds. AGU Monographs.

MacDonald, L., and D. Coe, 2008. Road Sediment Production and Delivery: Processes and Management. *In* Proceedings of the First World Landslide Forum. International Programme on Landslides and International Strategy for Disaster Reduction. United Nations University, Tokyo, Japan, pp 381-384.

Martin, D., 2009. Forest Road Runoff Disconnection Survey of Private Timberlands in Washington. Washington Forest Protection Association, Olympia WA. January 30, 2009.

Megahan, W.F., 1974, Erosion over time on severely disturbed granitic soils: a model: USDA Forest Service Intermountain Research Station General Technical Report INT-156, Boise, ID.

Megahan, W.F., and G.L. Ketcheson, 1996. Predicting Downslope Travel of Granitic Sediments from Forest Roads in Idaho. Water Resources Bulletin, Vol. 32, No. 2, pp. 371-382.

Megahan, W.F., and W.J. Kidd, 1972. Effect of Logging Roads on Sediment Production Rates in the Idaho Batholith. USDA Forest Service, Research Paper INT-123, 14p. Intermountain Forest and Range Experiment Station, Ogden, Utah.

Meyers, E., 2007. Reducing Sediment Production from Forest Roads during Wet-Weather Use. PhD. Dissertation, Oregon State University, August 3, 2007.

Mills, K, L., Dent, and J. Robben, 2003. Wet Season Road Use Monitoring Project. Oregon Department of Forestry, Forest Practices Monitoring Program, Technical Report #17. June 2003.

NCASI, 2000. Handbook of Control and Mitigation Measures for Silvicultural Operations. Unpublished Draft Technical Bulletin. Research Triangle Park, NC:National Council for Air and Stream Improvements, Inc.

Raines, M., R. Conrad, J. Clark, D. Coe, R. Palmquist, C. Veldhuisen, 2005. Road Sub-Basin Scale Effectiveness Monitoring Design. Prepared for the Washington Department of Natural Resources. May 9, 2005.

Reid, L.M,, 1981. Sediment Production from Gravel-surfaced Forest Roads, Clearwater Basin, Washington. M.S. Thesis, University of Washington.

Reid, L.M., and T. Dunne, 1984. Sediment Production from Forest Road Surfaces. Water Resources Research Vol. 20, No. 11, pp. 1753-1761.

Rogers, L. W. and A. G. Cooke, 2009. The 2007 Washington State Forestland Database, Final Report. Rural Technology Initiative, University of Washington. March 24, 2009. pp. 47 and appendices.

Rural Technology Institute, 2007. Washington Statewide Parcel Database, 2007 9.2 version 1. Restricted Distribution GIS (ArcGIS 9.2) dataset compiled by the University of Washington Rural Technology Initiative for the Washington State *Parcels Working Group*. Seattle, WA.

Sugden, B.D., and S.W. Woods, 2007. Sediment Production from Forest Roads in Western Montana. Journal of the American Water Resources Association. Volume 43, No. 1, pp. 193-206.

Sullivan, K.O., and S.H. Duncan, 1981. Sediment Yield from Road Surfaces in Response to Truck Traffic and Rainfall. Weyerhaeuser Technical Report 042-4402.80. Weyerhaeuser Company, Technical Center, Tacoma, WA 98477.

Swift, L.W. Jr., 1984. Gravel and Grass Surfacing Reduces Soil Loss from Mountain Roads. Forestry Science, Vol. 30, pp. 657-670.

Swift, L.W. Jr., 1985. Soil Losses from Roadbeds and Cut and Fill Slopes in the Southern Appalachian Mountains. Southern Journal of Applied Forestry, Vol. 8 No. 4, pp. 209-213.

Trimble, G.R., and R.S. Sartz, 1957. How far from a Stream Should a Logging Road be Located? Journal of Forestry, Vol. 55, pp. 339-341.

Washington Department of Natural Resources, 2005. "CMERlands" Available GIS Data. August 2005. http://www.dnr.wa.gov/AboutDNR/BoardsCouncils/CMER/Pages/Home.aspx.

Washington Forest Practices, 1997. Standard Methodology for Conducting Watershed Analysis, Version 4.0. Washington State Department of Natural Resources, Forest Practices Division. November 1997.

Watershed Professionals Network, 2006. Washington Road Sub-Basin Scale Effectiveness Monitoring Field Protocol. Prepared for the Washington Department of Natural Resources. Version dated February 28, 2006 (used for 2006 field work).

Watershed Professionals Network, 2008. Washington Road Sub-Basin Scale Effectiveness Monitoring Field Protocol. Prepared for the Washington Department of Natural Resources. Version dated April 17, 2008 (also updated at end of 2008 field season – version dated March 31, 2009).

Wemple, B., J. Jones, and G. Grant, 1996. Channel Network Extension by Logging Roads in Two Basins, Western Cascades, Oregon. American Water Resources Association, Water Resources Bulletin. Vol. 32 No. 6, pp. 1995-1207.

Appendix A. Meeting Minutes

WDNR Road Sub-Basin Scale Effectiveness Monitoring
Phase I Kickoff Meeting Minutes
Wednesday, December 14 1 PM

Attendees: Kathy Dubé, Mary Raines, Curt Veldhuisen, Jenelle Black
By conference call: Karen Kuzis, Julie Dieu, Dawn Hitchens, Nancy Sturhan

1. Introduction and Meeting Purpose (Jenelle Black)

Purpose of meeting was to review the objectives and proposed methodology for the road sub-basin scale effectiveness monitoring project so contractor (WPN) and UPSAG committee members have shared vision of how project will proceed.

2. Goals and Objectives of Monitoring Project (Jenelle Black)

Jenelle reviewed goals in RFQQ:

- Provide a way to measure if the FFR road rules are having a positive effect on the condition of roads on forested lands in Washington.
- This assessment will establish baseline condition – assessment will be repeated at 5 year increments (separate project)
- Current project will develop methods and protocols for assessment methodology (spelled out in the monitoring design), develop crew training materials, train crew, and survey the first group of road sites (anticipated to be 17 – 20 sites)
- QA/QC procedures are quite involved and are also part of this project – objective is to provide statistically valid sample and be able to judge amount of error associated with field measurements
- Survey protocol will also include fish passage questions from ISAG. For reporting purposes, these can be included in a separate table in the database, linked to information collected on that drainage structure.

The group discussed the responsibilities of team members:

- On the DNR/CMER side, Jenelle Black is project manager. UPSAG team members will be available to review documents and guide process, as requested by the DNR/CMER project manager. Dawn Hitchens will handle contract/billing questions. Nancy Sturhan will be available to answer any policy questions.
- The WPN team will divide work among Kathy Dubé (technical issues, protocol, training), Karen Kuzis (contract/billing and logistics), and Mary Raines (QA/QC lead).

3. Field Protocol and Training (Kathy Dubé)

Kathy said that most of the data collection methods are clear, but she had a few questions on ways to deal with segments that are unusual, particularly as far as delivery and road drainage configuration. After some discussion, the group decided Kathy is going to list the questions/special situations and provide a suggested approach to addressing the situations.

UPSAG will review these and come to agreement about how they will be handled. Question also came up as to how the FFR metric of miles of road/miles of stream will deal with partial delivery segments. This is really a policy issue. WPN will provide suggestions and UPSAG again will make final determination.

Marking segments in field – Kathy has a list of several methods for monumenting the starting points of road segments that could be used depending on landowners and logistics: GPS, odometers, paper base maps with notes, and photos. Julie Dieu said that after experimenting with several different methods, they have found that putting metal tags on the base of trees/stumps near the top of the cutslope is a good way to deal with this. Use aluminum nails. Scratch info into tag, or use plastic tags and write on both sides (UV rays fade writing). The group decided that several methods should be available and planned for, depending on the site and any landowner constraints. As long as the method used for each monument is well-documented, the use of multiple methods is not expected to be a problem.

Nancy Sturhan asked who will make changes to the WARSEM code. Marc McCalmon has been designated as the person to do this. He worked on the original WARSEM code and currently works for WDFW. There were some concerns about whether or not he would have time to do this or if we needed an alternate person available. Nancy/Jenelle will work on scheduling this with Tim Quinn (Marc's boss). We anticipate 1 week of Marc's time, in January or February.

Training Location - Several sites were discussed. Jenelle said it would be nice if we could use one of the actual monitoring sites – maybe Skagit or Grays Harbor area – for training. Kathy also mentioned a DNR-owned property near Seattle that was used for the WARSEM project that might be available. We discussed that it might be hard to find an Eastside site that was snow-free at time of training – may need to hit that site later in Spring, as the crews start to work on that side.

Training Date – targeting February

How many DNR/UPSAG members and who can or do want to attend the training? Jenelle was going to do some inquiry about who (from where) may want to attend. Curt said he may attend one day. Kathy said up to 12 people in the field portion of the training is OK – as many as want can attend the first day (classroom).

Julie Dieu mentioned that in her experience, having crew members train the next generation of field crew was the most effective method (job shadow for a week or so – worked well).

4. QA/QC Program (Mary Raines)

Mary asked if the training materials and field protocols part of the QA/QC document or if they are separate documents that reference the other? The group said either way would work as long as they are integrated. Kathy and Mary are going to coordinate development, probably separate documents which will tier off of each other. In response to a question by Mary, the QAPP should include the methods and statistical tests for evaluating operator variability.

Jenelle expressed an expectation that the QA process will include methods for identifying and correcting problems and errors at all levels in data collection and processing, and that tracking operator variability throughout the process rather than just at the beginning and end of data collection would be one way to do this. This suggestion will be taken into consideration during the QAPP development if practical, but it is not specified in the contract.

5. Field Surveys (Kathy Dubé)

Site selection status – Sites have been identified but status is unclear. Jenelle is working on this. UPSAG members may be able to help? Jenelle is going to follow-up and determine what needs to happen to let WPN know which sites to sample.

Protocol needs to include how to determine what's a road vs. a cat road. On Westside, rocked vs. native can be helpful. Also, gradient over 18%.

6. Reporting

Kathy will have Specific Questions on Field Approach by Jan 3rd – maybe a Draft Field Protocol. Mary will have outline of QAPP by Jan 3rd. UPSAG Meeting is Jan 10th

Jenelle will provide a list of UPSAG meeting dates so we can plan monthly report submissions (normally meetings are first Tuesday of each month; will let us know if different). We will plan to send in monthly reports on last Tuesday of month.

WDNR Road Sub-Basin Scale Effectiveness Monitoring
Phase I Second Unit Meeting Notes

Date:	March 31, 2006
Time:	8 AM
Purpose:	Discuss any changes to field protocol necessary based on first two field sites (S001, S002)
Personnel:	Dan Thomas, Jesse Saunders, Matt Rourke, Brett Winterowd, Pete Malinak, Karen Kuzis, Jenelle Black, Mary Raines, Kathy Dubé

Discussion Topics:

1. Any comments on delivery flowchart (attached file)? No – looks good, is working.

2. Changes to either the GPS data dictionaries or the paper data forms made during first week (see Road Form short.xls):
   - New form (Road log) for keeping track of mileage on each road segment completed.
   - Added Begin Reference Point ID on Reference Point form (GPS and paper)
   - Changed Road Name to text on Williamson's data dictionary
   - Minor tweaks such as adding S002
   - Begin RP is set to "1" as a default – it can be easy to forget to input this number. Dan will look into flagging this in a manner so it can't be missed as easily.

3. GPS road line feature does not need to be run if there is a good road layer in GIS. Only run a road line if it's a new road (not on map) or the road layer looks obviously wrong based on the orthophotos.

4. One item that is hard to rate in the field is road maintenance category. It's hard sometimes to determine what is "recent" in the field – within the last year? Last 5 years? Discussed that recent maintenance is within last five years, and the objective is to break out RMAP activities from normal maintenance. Normal maintenance includes surfacing, grading, and ditch cleaning. Other BMPs includes anything else such as new culverts, sediment traps, etc. that are intended to bring roads up to current/RMAP standards.

5. Vegetated ditchlines tended to trap quite a bit of sediment, but "vegetated" is not called out as a separate ditch condition. (Kathy noted we don't have much research data on this, but it's obvious if there is future research, this could change. Vegetated ditches do not transport sediment in the same way as "stable" ditchlines. We group these conditions under "stable" in the database.) Keep as "stable" for now.

6. At road intersections, note road junction road name in comments field on Reference Point form.

7. At stream crossings with no culvert apparent (e.g. Humboldt, or a ford) we will put in a new field called “Other/unknown” in the Drainage Structure Type field.

8. Brett asked about surfacing in the situation where there was a lot of leaf litter/organic matter over what used to be gravel. He recorded it at pitrun. We all agreed.

9. QA/QC going OK; checking records at end of day, and correcting items after conversion to a shape file.

10. Deliverables for each unit includes (2 hard copies of each plus an electronic file where applicable):

    - Copy of field maps (marked with your name - small maps as well as large map)
    - Copy of any data sheets (we used a few where the GPS wasn't recording satellites)
    - Report of road miles surveyed (road log - field form) so we know how many miles we get paid for
    - Photos (hard copies and electronic)
    - Final Maps (Ted produces these - marked with intersection/culvert/drainage points, streams, roads, project boundary - electronic file as well as hard copies)
    - Copy of GPS data files (convert to shape files and give to Ted - and Kathy will load the road segment data into the road database to run WARSEM - only need electronic copy of these)

11. It is most helpful to have completed the landowner interview ***BEFORE*** the field crews are on the ground to get:

- Keys,
- Directions to field unit,
- ROAD NAMES and GIS layers
- Any access issues (or areas to be careful of)
- Road maintenance levels

WDNR Road Sub-Basin Scale Effectiveness Monitoring Phase II
Kickoff Meeting Minutes
Friday March 21, 2008 9 AM

Attendees (conference call): Jenelle Black, Julie Dieu, Kathy Dubé, Dawn Hitchens, Karen Kuzis, Laura Vaugeois, Curt Veldhuisen

7. Introduction and Meeting Purpose

Purpose of meeting was to review the objectives and proposed methodology for Phase II of the road sub-basin scale effectiveness monitoring project so contractor (WPN) and UPSAG committee members have shared vision of how project will proceed.

8. Updates to Methods

a. Wetlands
Current protocol includes delivery to Type A and B wetlands. Jenelle mentioned that there is a new wetland layer on the DNR website (date 12/2007) that Ted should use for making field maps. She also mentioned that we should continue to note delivery to any un-mapped but obvious typed wetlands that are apparent from the orthophoto coverage. and make notes in the comment field during the inventory if there are wetlands we see on the ground and we cannot easily determine if they are Typed wetlands or not. We will discuss wetlands during training. There is interest from other CMER groups on the data we are collecting pertaining to wetlands.

b. Roads next to streams
There was discussion of how to handle roads that run alongside streams and wetlands (and through wetlands). Currently the protocol looks at delivery from drainage structures (constructed or not) and includes some amount of delivery if a sediment plume ends within 20 feet of a stream. Jenelle was concerned that we are not capturing any information on roads that have dispersed runoff and may deliver some sediment. Existing research and observations suggest that dispersed runoff may only transport sediment 10-15 feet from road. Concern was also expressed that we remain consistent with methodology from first set of sample units (Phase I). Consensus: An additional delivery category will be added – category 5 – for road segments that would have been classified no delivery (category 0) but that are along or through a stream or wetland and the edge of the road tread is within 20 feet of a stream/floodplain or typed wetland. The begin/end station and all other road attributes will be collected for these segments.

9. Site Selection/Landowner Contacts

a. Blue Mountains
Jenelle said that sites in the Blue Mountains area are tricky because there are cultivated ag lands on the hilltops and some inter-fingering forest lands in valleys/north slopes. Farm roads are commonly used for hauling, and are not being maintained as forest roads. UPSAG decided not to include the forest roads in the monitoring; consequently many of the potential units in this are have dropped out.

b. Description of site selection process for report
   Jenelle will write a description of the site selection process that will be included either within the text or within an appendix of the final report.

c. Procedure for sites in "December 2007 storm" area
   We discussed the procedure to use for units within the area hit hard by the December 2007 storm (SW Washington). There is concern that there may be enough damage that access either to the unit or on roads within the unit would be extremely difficult and beyond what any of us had budgeted for (both slides and blowdown). The pros and cons of sampling sooner (capture storm damage – but is it representative?) vs. later (storm damaged roads may be freshly fixed, but more likely to be able to access roads) were discussed. Consensus: Check with landowners of these sites soon to determine extent of problem (possibly also look on the air photos UPSAG will be obtaining for the mass wasting study?). Then check back with UPSAG to determine how to proceed.

d. Landowner letter
   Jenelle has updated the landowner letter and will send it out Monday to landowners in units we anticipate to be sampling soon – rest of landowners as they are identified.

10. Field Training/Testing Methods and Plan

a. Training currently scheduled for April 14-15 in units near Ellensburg
b. Then crews will complete 1-2 units in Ellensburg area (April 16-17…) with Kathy/Jenelle/Karen available for ride-alongs.
c. Second Unit meeting after they complete these units and are still in area
d. First Variability Test at end of week (April 18). Kathy will run through roads selected for testing to make sure any ambiguous items (such as weird road junctions) are cleared up before test and will flag start/stop locations.
e. Training will include PowerPoint presentation and then field training (first with paper, then GPS/Jamars). Will include last year's training plus discussion of wetlands, roads next to streams, any other revisions. OK to discuss differences between observers we found from last year's variability testing.

11. QA/QC Methods and Plan

a. Field Assistance Program (ride-alongs). Kathy/Jenelle/Karen – will ride along first week of field sampling, also available second week as needed if questions (Jenelle gone April 21-22). Then ride along one day/sampling month with each crew member for duration.
b. Replicate surveys. Try to schedule several during first 2 weeks of inventory, plus others throughout sampling season. Williamson crew found it helpful to do a quick check at the beginning of each unit to confirm surfacing calls, any other issues they saw with unit.
c. Third Party QA/QC. Karen will do this. Objective is to make sure crew members are following field methods and QA/QC protocols.
d. Data Cross-check. Each member checks data at end of day (issue with INFO not having enough time last year – going too fast – will discuss with them). Also check at end of

each unit; Kathy performs final QA/QC to ensure all data fields filled in, all photos, maps, etc. are in file. Will use unit check off list as we did last time.

e. Will provide general feedback to crew members on areas of concern following first variability test (no site specific information that could bias test at end of project).
f. Will make sure notes from each replicate survey and field assistance program visit are distributed to all crew members so they can see what items were discussed.

12. Field Crews

Have 5 field crews from Phase I plus one new field crew member (Amanda Ogden) to help Williamson crews.

13. Sampling Plan

Sampling will start in units near Ellensburg in conjunction with training/testing. Many west side units should be clear of snow and ready to go – we'll be checking on access as noted in 3.c. above. Williamson would like to go on Eastside units – will be checking snow/mud conditions.

WDNR Road Sub-Basin Scale Effectiveness Monitoring Phase II
Second Unit Meeting Minutes
Friday May 23, 2008

Attendees Kathy Dubé, Scott Hotchkiss, Pete Malinak, Jesse Saunders, Dan Thomas, Brett Winterowd, (via conference call) Jenelle Black, Karen Kuzis

14. Introduction and Meeting Purpose

Purpose of meeting was to review the data collection procedures from the training, first, and second sample unit, and discuss any issues that came up in field assistance visits or replicate crew surveys to date.

15. Discussion

a. When is a skid trail vs. a road?
Contract says inventory roads, not skid trails (roads = driven on by logging trucks). Sometimes the distinction is not obvious, particularly if skid trails are constructed/graded. Came across examples in E008 - unit was harvested in last 5 years – making skid trails more obvious. The skid trail in question was graded and looked like a road for part of its length, but like a skid trail for last half of length and has not received maintenance. The manual says the roads only count if constructed to be driven on by trucks. Brent noted a reliable factor for identifying roads vs. skid trails are the location of landings and slash piles. If there is no landing at the end it probably wasn't used as a road. Crews will have to use some judgment in gray areas. The situation is most likely to arise in areas that have been harvested more recently.

b. Spring and seeps were observed in numerous cut slopes – when is it a stream?
Some are obvious seeps starting in a cut slope and forming a channel downstream (often in road ditch) which is then defined as gully. If there is an obvious upslope channel then it is a stream.

The definition was more difficult where water was coming out a bit above the cut slope. Kathy and the crews decided if the water and channel looks like it start is within 10-20 ft of cut slope and is associated with the ground disturbance at the cut slope, it should be classified as a seep. If the channel is further up the hill and cut into native material call it a stream.

c. When should you note sediment delivery that is not necessarily related to the roads being surveyed?
Jenelle observed a landing constructed on a spring complex that delivered and mitigation measures weren't working. She also noted skid trails constructed and filled in type 5 channels diverting water down skid trail to the stream. She had made notes of these situations on the map. After consideration she thought we might not want to attribute non-road related problems to specific landowners in notes or on the maps. Better approach would be to keep a list of observed sediment delivery sources not necessarily related to roads in a separate notebook and pass the information onto Kathy and Jenelle. If needed this can be communicated to landowners informally by Jenelle and if similar conditions are

observed in more than one location may be noted in the report as item for future consideration. If we find sediment sources not related to roads keep notes on separate paper don't make specific notes on field forms and maps.

d. Surfacing
Remember to agree on surfacing at beginning of the unit.

e. Monuments
Put monument tags at the base of the trees – do not make them too obvious as targets.

f. Recording Non-Delivering Stream Adjacent Roads
Dan at INFO is concerned this was not included in the original Scope of Work and will likely take more time especially in units with a high density of streams and wetlands. May need a change on work order.

Kathy noted the 1st unit had two additional segments non delivering segments they had to stop at 3 other segments measuring the distance. The 2nd unit 2 segments which were delivering and would be recorded anyway.

Jenelle noted if crews are driving along a road adjacent to stream they should be getting out and looking at the stream/road adjacent to the stream to assess delivery anyway. Dan noted if the segment is not delivering you don't need to make a segment.

Dan also noted on the Westside wetlands dam water and there are lots of wetlands adjacent to roads where they will need to add segments. Kathy noted the difference is adding segments even if no delivery. Dan will evaluate how long it takes to add the new segments to address Jenelle's concern that they would have been stopping and looking at these areas anyway.

Kathy noted a quick review showed 2,000 to 4,000 feet stream adjacent road in each unit. The team will have a call with Dan after they complete the Hancock Unit (W043) next week to discuss how much work this actually added.

g. Consistency Issue - Delivery
Kathy told the crews they need to get in the habitat of following the delivery flow chart to stay consistent.

16. Other items discussed after meeting – field crews please note

a. Powerline roads
Issue came up in E037 regarding powerline roads. Roads used just for powerline access are not subject to DNR regulations. Roads used for both powerline and timber haul are. Jenelle suggested checking with landowners in these units to determine which roads to survey/not survey.

b. GPS point at end of road
Dan said he normally takes a GPS point at the end of the road to mark the end point of survey for Ted.

c. Blocked roads

Jenelle said field crews should remember protocol on blocked roads. Walk road out with string box for 200 meters. If 90% veg cover on road, can stop. If less than 90% veg cover, keep walking to end (or point where 90% cover for 200 meters). Note blockage, how far you went/if vegetated on map.

17. Next Steps

Today the crews are heading to complete variability testing in Unit E083.

Next week (5/27 to 30) Info will sample W043 (Hancock) and the following week (6/2 to 6/6) they will be in Forks for Units S007 and S005.

Williamson starting June 2nd at a unit to be determined. Possibly E069? depending on permissions.

WDNR Road Sub-Basin Scale Effectiveness Monitoring Phase II
Phase II Survey Completion Meeting Minutes
Thursday January 29, 2009
8 AM via conference call

Attendees Kathy Dubé, Jesse Saunders, Dan Thomas, Brett Winterowd, Jenelle Black, Karen Kuzis, Ted Hitzroth

Meeting Purpose

Purpose of meeting was to discuss sampling results, any issues that arose during the project and the analysis of sample data required to meet project goals.

Items Discussed

1. Phase II Field Sampling
    a. Any issues or suggestions for next time? Any specific things noted in field that should be considered for analysis of particular units (i.e., most roads were vegetated so we didn't survey very far; lots of pulled culverts)
        - Add a new field in the non-delivering point file for presence of BMPs (some non-delivering culverts didn't deliver because landowners had done a great job with BMPs – should be noted)
        - Add a method to show if a road is not surveyed, and why. Note these roads on field map, and why not surveyed (e.g., vegetated, blocked, not found on ground). Add this information into GIS road layer (line feature attributes) and map. One reason this is important is because number of stream crossings in GIS vs. found in field is used to adjust miles of stream in unit (FFR metric)
        - Logistics:
            - Fro efficiency, define unit, contact landowners as far in advance as possible.
            - Good for field crews to know contacts, who has keys. Can be an issue on Eastside where some landowners don't know which locks are on which gates.
        - Training/documentation ideas:
            - Describe Eastern/Western WA road differences and how to classify them (ditches, etc.)
            - Definition of a stream, disconnected stream.

2. Post-Sampling Processing
    a. Any issues or suggestions for next time?
        - Add Point ID for non-delivering culverts
        - Add lat/longitude into Access database for each point so that other users can download into hand-held GPS units

3. QA/QC Protocols (training, management visits, replicate surveys/ride-alongs, crew variability assessment)
    a. Any issues or suggestions for next time?

- Make sure that field crews review data at end of day to help catch incorrect entries (particularly stationing that's way off)
- Training:
    - Stress how to define a stream in the field, and include training on methods for disconnected streams.
    - Include differences between Eastern and Western WA road construction techniques and how to code (for example, presence/absence of ditches, outsloped roads).
    - Start field training with easier sites, then move onto more complex sites
    - Maybe base training in Cle Elum area and include one field day on E WA roads, one field day on W WA roads.
- Crew variability – might get better measure of variability if easier road segments were used instead of confusing ones.

4. Data Analysis
    a. Database Entry
        - How to deal with new fields – stream adjacent – with 2006 data? Add new field into Access database, populate with -8888 for 2006 data to indicate that no data was collected in 2006
    b. Selection of 4 square mile area for 2006 data units
        - Jenelle will send a clipped coverage to Ted, Ted will send Kathy a file listing points that are in/out of 4 square mile area.

At this point the call was ended and Kathy, Jenelle, and Karen discussed the following items on Thursday February 5.

    c. Database update:
        - add fields for latitude/longitude for each point
        - pull Qsecs 17-24 from 2006 data into separate data tables within database.
    d. Statistical analysis (operator variability, Monitoring Design Table 10 Questions – MQ1, MQ3, MQ4) – Alice is unpacking in her new location; she said that she can do analysis at end of February.

5. Report Outline, Schedule
    a. Proposed outline, work products
        - Audience: assume there will be readers who do not know about CMER, monitoring program.
        - Intro – can get text from study design, RFP
        - Map – just put dot for each unit to show general area
        - Do not report results by unit in report (unit by unit summary in Access database)
        - Summarize results by region, large vs. small landowners
        - Summary – list monitoring questions and summary answers
        - Recommendations for future sampling events, scientific study. No policy recommendations.

   b. Schedule – draft report end of February

6. CMER Presentation
   a. Abstract due 2/18
   b. Draft PowerPoint to Jenelle by 2/27 (review 3/2 in Olympia)
   c. Final PowerPoint to Jenelle by 3/16
   d. CMER Science Conference 3/18 – Kathy presents in Olympia
   e. CMER meeting presentation April 28?

Appendix B. Quality Assurance Project Plan (QAPP)

The Quality Assurance Project Plan is available electronically on the data CD.

Appendix C. QA/QC Report

The QA/QC Report is available electronically on the data CD.

Appendix D. Field Protocol

The 2008 field protocol is available electronically on the data CD.

Questions Leading to a Forests & Fish Policy
Adaptive Management Recommendation to the Forest Practices Board

Project title: **Road Sub-Basin Scale Effectiveness Monitoring Project**
Report title: **Washington Road Sub-Basin Scale Effectiveness Monitoring, First Sampling Event (2006-2008) Report**

1. Does the study inform a rule, numeric target, performance target, or resource objective?

The study evaluates whether forest roads meet the Numeric Targets for hydrologic connectivity and sediment delivery to streams

2. Does the study inform the Forest Practices Rules, the Forest Practices Board Manual guidelines, or Schedules L-1 or L-2?

This study informs the following Schedule L-2 questions:
1. "*Test the effectiveness of the roads program at disconnecting road drainage from the stream network,*" and
2. "*Determine the effectiveness of road maintenance BMP on a site-scale and sub-basin scale in meeting road sediment targets*" (note: this study addresses BMP effectiveness at the *sub-basin*, but not the *site* scale)

3. Was the study carried out pursuant to CMER scientific protocols (i.e., study design, peer review)?

The study design and report followed all protocols, including UPSAG, CMER and Independent Science Panel reviews & revision.

4. What does the study tell us? What does the study not tell us?

This study was designed to consist of three sequential sampling events to track the condition of forest roads through the 15-20 year implementation period of Road Maintenance and Abandonment Plans (RMAP) initiated in 2001. This report summarizes the findings of the first sampling event, which was conducted in 2006 and 2008.

Road managers reported that over half of the sample units had at least 85 percent of road length meeting post-RMAP standards. Across all samples, an average of 11 percent of the road length was hydrologically connected to streams or wetlands, though much variability exists between regions and blocks. Sixty two percent of the road samples met the regional performance target for hydrologic connectivity and 88 percent of the samples met the sediment target. These are all favorable results, given that they were observed less than half way through the RMAP implementation period. Sediment delivery performance by sample block was statistically correlated with progress toward RMAP standards. However, hydrologic connectivity was not statistically related to progress toward rule standards, reflecting that connectivity targets are difficult to achieve for roads located in areas of high stream density. The results of future

monitoring events (planned interval of five years) will identify what changes in road performance result from additional road improvements.

Due to the sample selection protocol, approximately ninety-five percent of the roads sampled were within large industrial and state or local government ownership. Although the study was intended to incorporate roads owned by small forest landowners, their fragmented ownership pattern seldom fit into the sub-basin-sized (i.e., 4 mi^2) sample blocks. Development of a companion study designed specifically to access and evaluate roads on small ownerships would be needed to provide such a characterization.

This study did not address the following related issues:

a. It did not directly measure actual eroded sediment quantities delivered to surface water, turbidity levels or biotic impacts, because they were outside the project scope.
b. It did not evaluate the effectiveness of road conditions at preventing sediment delivery from landsliding.
c. It did not evaluate the effects of individual road practices because a sub-basin scale sampling approach was chosen instead.
d. It did not evaluate the contents of or implementation of RMAP, which are regulatory issues.
e. It did not evaluate the implementation or effectiveness of fish passage at forest roads.

5. What is the relationship between this study and any others that may be planned, underway, or recently completed?

This study is one of several studies in the CMER Work Plan conceived to evaluate the effects of forest roads on watershed functions, as shown below:

Project Name	**Project Status**	**Anticipated knowledge to be gained**	**How will uncertainty be reduced?**
Mass Wasting Effectiveness Monitoring (Post Mortem)	Report in CMER review, ISPR expected in early 2011	Establish relationship between road maintenance conditions and road landslide initiation	Understanding of road sediment delivery via landslide initiation
Road Surface Erosion Model Validation Project	In CMER Work Plan, not yet scoped	Evaluate model's ability to predict sediment generation*	Improved confidence in sediment delivery values
Road Site-Scale Effectiveness Monitoring Project	In CMER Work Plan, not yet scoped	Quantify the effectiveness of individual road treatments	Improved understanding when applying individual road maintenance treatments
Intensive watershed monitoring.	In CMER Work Plan, not yet scoped	Establish biologically based performance targets for sediment and connectivity	Understanding of how road sediment affects biota
Extensive Fish Passage Monitoring	Delayed. Study was designed but needs revision.	Determine status and trend of fish passage on a regional scale	Field verification of fish passage work on forest roads

*The Washington Road Sediment Model used in this study was refined using all road research available in 2004. The validation project would be most valuable if additional erosion models or new field research was available.

6. What is the scientific basis that underlies the rule, numeric target, performance target, or resource objective that the study informs? How much of an incremental gain in understanding do the study results represent?

Performance Targets were developed using field data from Watershed Analyses and similar road studies. This study revealed some uncertainty in existing targets and indicated a wider range in road conditions than anticipated. Targets could be improved with results of intensive watershed monitoring and/or outside research. This study significantly improved knowledge of statewide forest road conditions, especially on industrial ownerships.

Type N Experimental Buffer Treatment Study: Baseline Measures of Genetic Diversity and Gene Flow of Three Stream-Associated Amphibians

Stephen Spear, Jason Baumsteiger, and Andrew Storfer

WASHINGTON STATE DEPARTMENT OF
Natural Resources
Peter Goldmark - Commissioner of Public Lands

Cooperative Monitoring
Evaluation & Research

May 2011

CMER 06-605

This page intentionally left blank

Washington State Forest Practices Adaptive Management Program
Cooperative Monitoring, Evaluation, and Research Committee (CMER)
Report

Type N Experimental Buffer Treatment Study: Baseline Measures of Genetic Diversity and Gene Flow of Three Stream-Associated Amphibians

Prepared by:

Stephen Spear, Jason Baumsteiger, and Andrew Storfer

Prepared for the

The Cooperative Monitoring, Evaluation, and Research (CMER) Committee
Washington State Forest Practices Board
Adaptive Management Program
Washington State Department of Natural Resources
Olympia, Washington

May 2011 **CMER 06-605**

Washington State Forest Practices Adaptive Management Program

The Washington State Forest Practices Board (FPB) has established an Adaptive Management Program (AMP) by rule in accordance with the Forests & Fish Report (FFR) and subsequent legislation. The purpose of this program is to:

> *Provide science-based recommendations and technical information to assist the FPB in determining if and when it is necessary or advisable to adjust rules and guidance for aquatic resources to achieve resource goals and objectives. The board may also use this program to adjust other rules and guidance.* (Forest Practices Rules, WAC 222-12-045(1)).

To provide the science needed to support adaptive management, the FPB established the Cooperative Monitoring, Evaluation and Research (CMER) committee as a participant in the program. The FPB empowered CMER to conduct research, effectiveness monitoring, and validation monitoring in accordance with WAC 222-12-045 and Board Manual Section 22.

Report Type and Disclaimer

This technical report contains scientific information from research or monitoring studies that are designed to evaluate the effectiveness of the forest practices rules in achieving one or more of the Forest and Fish performance goals, resource objectives, and/or performance targets. The document was prepared for the Cooperative Monitoring, Evaluation and Research Committee (CMER) and was intended to inform and support the Forest and Fish Adaptive Management program. The project is part of the Type N Amphibian Response Program, and was conducted under the oversight of the Landscape and Wildlife Advisory Group.

This document was reviewed by CMER and was assessed through the Adaptive Management Program's independent scientific peer review process. CMER has approved this document for distribution as an official CMER document. As a CMER document, CMER is in consensus on the scientific merit of the document. However, any conclusions, interpretations, or recommendations contained within this document are those of the authors and may not reflect the views of all CMER members.

The Forest Practices Board, CMER, and all the participants in the Adaptive Management Program hereby expressly disclaim all warranties of accuracy or fitness for any use of this report other than for the Adaptive Management Program. Reliance on the contents of this report by any persons or entities outside of the Adaptive Management Program established by WAC 222-12-045 is solely at the risk of the user.

Proprietary Statement

This work was developed with public funding, including contracts [**list all related contracts by number**]. As such it is within the public use domain. However, the concept of this work originated with the Washington State Forest Practices Adaptive Management Program and the authors. As a public resource document, this work should be given proper attribution and be properly cited.

Full Reference

S. Spear, J. Baumsteiger, and S. Storfer. 2011. Type N Experimental Buffer Treatment Study: Baseline Measures of Genetic Diversity and Gene Flow of Three Stream-Associated Amphibians. Cooperative Monitoring Evaluation and Research Report, CMER 06-605. Washington Department of Natural Resources, Olympia, WA.

Author Contact Information

Stephen Spear
University of Idaho, Department of Fish and Wildlife Resources,
P.O. Box 442282
Moscow, ID 83844-2282
sfspear2@yahoo.com

Jason Baumsteiger
University of California, Department of Natural Sciences
5200 N. Lake Rd.
Merced, CA 95343
cutthroat99@hotmail.com

Andrew Storfer
Washington State University, School of Biological Sciences
423 Neill Hall
Pullman, WA 99164-3140
astorfer@wsu.edu

Acknowledgements

This work was funded by and prepared for the Cooperative Monitoring, Evaluation and Research Committee under the Washington Department of Natural Resources Forest Practices Adaptive Management Program IAA 04-047 to WSU. We would like to thank Aimee McIntyre, Marc Hayes and Eric Lund for help with logistics in the field and advice on site selection, as well as insightful discussions regarding stream amphibians. Craig Steele provided much advice and assistance regarding giant salamander analyses. Sean Anderson, April Barreca, Karen Chojnacki, Nick Chojnacki, Jen Cotter, Tierra Curry, Jennifer Dhundale, Cristina Dressel, Robert Dyer, Jon Eastman, Andy Giordano, Marc Hayes, Tiffany Hicks, Denim Jochimsen, Scott Jones, Eric Lund, Doré Mangan, Jeffrey Marsten, Aimee McIntyre, Teresa Miskovic, Casey Richart, Maureen Thompson, Charissa Waters, Frithiof Waterstat, Nicholas Wenzel, Anna Yost, Kevin Young and Kayla Zaret all assisted with sample collection. Jon Eastman and Melanie Murphy assisted with laboratory work. Marc Hayes, Aimee McIntyre, Ivan Phillipsen, Teresa Miskovic, Chris Mendoza, Adrian Spidle, and Terry Jackson provided suggestions on the draft report.

Executive Summary

One of the major objectives of the Type N Experimental Buffer Treatment Study is to investigate how different forest buffers surrounding non-fish-bearing streams may influence stream-associated amphibian populations. Measures of genetic diversity and genetic differentiation within and among amphibian populations can allow for species identification, provide insight into trends in population size and identify the level of migration among sites. Herein, we report baseline, pre-treatment measures of genetic structure for three species of stream-associated amphibians: Coastal tailed frog (*Ascaphus truei*), Cope's giant salamander (*Dicamptodon copei*) and Coastal giant salamander (*Dicamptodon tenebrosus*). We addressed four main objectives:

1) Classify giant salamander individuals to the correct species and identify any hybrids
2) Estimate indices of genetic diversity for the three species at each sampled basin
3) Test for significant differences by region, block and anticipated treatment for any genetic diversity parameter
4) Determine genetic clusters for each species

We identified giant salamander individuals to species with high probabilities, and found 31 hybrids out of 1504 total sampled *Dicamptodon* individuals. However, no hybrid individuals represented the current F_1 generation, suggesting hybridization is not currently ongoing. Levels of genetic diversity (based on allelic richness and heterozygosity) were high in Coastal tailed frogs and intermediate for Cope's giant salamander and Coastal giant salamander. Fourteen sites (out of 47 possible) had high levels of inbreeding and/or evidence of declines, but no consistent patterns were found among the three species. Effective population sizes varied greatly among species, with very large sizes for Coastal tailed frogs (3000-7000), intermediate sizes for Cope's giant salamander (150-3150) and low sizes for Coastal giant salamanders (50-750). Genetic parameters revealed no differences that would suggest biased results based on site selection for any species, demonstrating that the approach for assigning treatments was successful. Generally, no differences existed by region or block, with the notable exception of Cope's giant salamander, which had lower values of allelic richness, heterozygosity, inbreeding coefficient and effective population size in the Olympics. Finally, we detected geographically large genetic clusters (at a regional scale) for Coastal tailed frogs and Coastal giant salamanders, consistent with their broad species range and increased ability to disperse terrestrially. In contrast, Cope's giant salamander had geographically restricted genetic clusters, whereby individual sites often were genetically distinct from nearby sites. However, degree of divergence varied by region, with a larger cluster in the Olympics and the most restricted gene flow in the South Cascades.

1. Introduction and Objectives 1
2. Materials and Methods 3
2.1 General Methods 3
2.1.1 Sample collection and study sites 3
2.1.2. Laboratory Methods 5
2.1.3. Genetic Analyses 5
2.2. Objective 1 Methods: Distinguishing giant salamander species 11
2.3. Objective 2 Methods: Providing baseline levels of genetic diversity 12
2.3.1. Summary indices of genetic diversity 12
2.3.2. Estimate of effective population size 12
2.3.3. Evidence for recent population declines 13
2.4. Objective 3 Methods: Testing for differences in genetic diversity among sites or blocks 14
2.5. Objective 4 Methods: Estimating genetic clusters 15
3. Results and Conclusions 16
3.1. Objective 1 – Dicamptodon species identification and hybrids 16
3.2. Equilibrium tests, presence of null alleles and sibling estimation 18
3.2.1. Coastal tailed frogs 18
3.2.2. Cope's giant salamanders 21
3.2.3. Coastal giant salamanders 23
3.3. Objective 2: Measures of genetic diversity and population size 25
3.3.1. Coastal tailed frogs 25
3.3.2. Cope's giant salamanders 27
3.3.3. Coastal giant salamanders 30

3.4. Objective 3: Comparison of genetic diversity measures among groups 32

3.4.1. Coastal tailed frogs 32

3.4.2. Cope's giant salamanders 35

3.4.3. Coastal giant salamanders 37

3.5. Objective 4: Genetic cluster size 39

3.5.1. Cluster site selection 39

3.5.2. Coastal tailed frogs 41

3.5.3. Cope's giant salamanders 43

3.5.4. Coastal giant salamanders 56

4. Final Conclusions and Synthesis **59**

4.1. Coastal tailed frogs synthesis 59

4.2. Cope's giant salamanders 59

4.3. Coastal giant salamanders 60

4.4. Final conclusions 60

5. Acknowledgments **61**

6. Literature Cited **61**

7. Glossary **66**

Appendix 1 ..... **68**

Appendix 2 ..... **69**

Appendix 3 ..... **70**

Appendix 4. ..... **71**

Appendix 5. ..... **72**

1. Introduction and Objectives

The Type N Experimental Buffer Treatment Study (hereafter the Type N Study) has the primary objective of identifying how different harvest treatments influence the biotic and physical resources and processes in non-fish-bearing streams, with the ultimate goal of informing forest management about which buffer prescription(s) is(are) the most effective in maintaining those processes and biotic resources. The Type N Study is investigating four main categories of response variables including amphibian abundance and genetic diversity, downstream fish abundance, export of nutrients, detritus and macroinvertebrates and water temperature. The inclusion of amphibian variables is due to several factors. Nearly 40% of all amphibian species are thought to be in decline and are thus of conservation concern. Due to their apparent greater sensitivity to environmental degradation when contrasted with other vertebrate groups, amphibians are recognized as important indicators of ecosystem health (Stuart et al. 2004). Especially pertinent to the goals of the Type N Study, amphibians are important trophic components in stream ecosystems, comprising a large percentage of the vertebrate biomass in Pacific Northwest streams (Bury et al. 1991). Therefore, forest practices that may influence amphibian populations could subsequently alter entire stream assemblages. Studies generally demonstrate a reduction in numbers of stream amphibians in managed stands relative to old-growth (Corn and Bury 1989, Welsh 1990). However, other studies have shown little effect of harvest on stream amphibians in second-growth forests (Diller and Wallace 1999, Wilkins and Peterson 2000. Although factors such as sedimentation, gradient, geology, and stream temperature have influenced abundance in the previous studies, there is little consensus on the reasons for the differences in harvest effects on stream amphibians. Therefore, there is need for carefully designed studies to test different mechanisms that moderate amphibian response to harvest. Additionally, previous studies have primarily been focused on assessing abundance of amphibians in streams. However, stream amphibians are most easily detected as larvae (Spear and Storfer 2008; Kroll et al. 2010), and larval numbers may not accurately represent adult population status (Goldberg and Waits 2009). A high number of larval individuals could be produced by only a few adults, and if this was the case, demographic studies would conclude a large population size, when in reality the breeding population was low. However, genetic data can assess levels of effective population size or reductions in population size that are not immediately obvious demographically (Luikart et al. 1998, Garza and Williamson 2001). As a result, we proposed molecular population genetic studies to assess pre-treatment amphibian population status, as well as post-treatment responses (Beebee 2005; Jehle and Arntzen 2005; Storfer et al. 2009). Coastal tailed frogs (*Ascaphus truei*) and giant salamanders (Cope's giant salamanders [*Dicamptodon copei*] and Coastal giant salamanders [*Dicamptodon tenebrosus*]) were chosen as the stream-associated amphibians for the focus of the baseline (pre-harvest) population genetics portion of the Type N Study. Coastal tailed frogs and Cope's giant salamanders (hereafter the two focal species) were the two species originally designated in the Type N Study design to be the focus of the genetic portion of this study as they were the only

stream-associated amphibian species occurring over a geographic range encompassing all study basins. However, identification and hybridization issues between the two species of giant salamanders, as later explained, resulted in the default incorporation of Coastal giant salamanders into the genetic analysis portion of the Type N Study.

Herein, we present the results of the amphibian genetic portion of the Type N Study, to provide baseline data for comparisons with future post-treatment analyses. Occurrence and abundance data are important to this study to provide an index of immediate population response and numbers, but the inclusion of genetic data is complementary because it provides a longer-term picture of population trends as well as insight into future population viability. Specifically, genetic data are pertinent to this study for the following reasons: (1) Both species of giant salamander are nearly impossible to identify using exclusively morphology during their larval stages, hence genetic markers are the best means of ensuring unambiguous identifications. (2) Amphibian population samples are usually larval-biased, which can lead to errant conclusions if one depends exclusively on abundance data. Multiple amphibian larvae can be the offspring of one parent, and larvae typically suffer high mortality rates prior to metamorphosis. Therefore, larval abundance estimates are unlikely to fairly represent adult population size. (3) Measures of genetic diversity provide insight into the long-term viability of a population. A population with low genetic diversity and a high level of inbreeding may not be sustainable at the scale of tens to hundreds of generations. Additionally, genetic diversity statistics can be used to detect changes in population size. (4) Finally, genetic data can estimate the degree of gene flow across a study area. The Type N Study design examines the effect of a treatment on a single basin, presumably reflected by the amphibians resident in that basin. In this design is the implicit assumption that individuals sampled post-treatment at that site were exposed to that treatment and thus are not immigrants from another site. Measures of genetic differentiation can be used to test this assumption. In addition, streams connected to Type N treatment basins can be affected by harvest in those basins; analyzing gene flow (and hence connectivity) among nearby streams can provide information of how the *local network* of streams may be affected.

The population genetics portion of the Type N Study had four main objectives:

1) Correctly identify individual giant salamanders to species (Cope's or Coastal) as well as identify any hybrids between the two species that might confound results.

2) Generate measures of genetic diversity for each species for each Type N treatment basin that would provide the baselines for comparisons to post-treatment data. Changes in measures of genetic diversity such as allelic richness, heterozygosity and effective population size can provide evidence for treatment effects that either would not be immediately apparent or would be undetectable through demographic measures.

3) Examine whether significant *a priori* differences exist among regions, blocks and treatments using measures of genetic diversity. As this study occurred before any harvest treatments were applied, we expect no initial differences to exist. However, it is critical that we determine that no initial differences exist (especially by treatment), so we can ensure that any potential differences seen in post-treatment analysis are not the result of any pre-existing conditions.

4) Identify the genetic cluster (spatial extent of gene flow) for each species using population clustering techniques. The genetic cluster will not only be useful for identifying the correct spatial or geographic scale of management for each species (i.e. stream, watershed, region etc.), but will also indicate whether treatments and blocks are independent or connected by gene flow.

2. Materials and Methods

2.1 General Methods

2.1.1 Sample collection and study sites

In collaboration with the Washington Department of Fish and Wildlife (WDFW), we collected tissues for genetic samples from three regions in western Washington: the Olympics, Willapa Hills and South Cascades (Fig. 1). These were the physiographic regions in the WDFW-developed Type N Study design; treatment basins were identified within each region. Each treatment is being applied to an entire non-fish-bearing basin, and four different treatments, grouped as study blocks, are being applied in the Type N Study: an unharvested reference (or control), and three prescriptions that vary in the length of the two-sided 50-foot wide stream buffer applied to each (one prescription has buffer along 100% of the stream length, one has no buffer, and one has the current Forest Practices Rules prescribed buffer (FF HCP, 2005), which approximates at least 50% of the stream length). For the purposes of this study, basins were 1st, 2nd, or 3rd order non-fish-bearing basins defined as the area from the non-fish-bearing point of the stream network upstream and including all headwaters of the stream network. One block (of the four aforementioned treatments) was sampled in the Olympics (sites 363, 1099, 1197, 1236; Fig. 2), 2.5 blocks were sampled in the Willapa Hills (sites 2260, 2468, 3074, 3098, 3110, 3111, 3437, 3576, 3914, 5785; Fig. 3) and 1 block was sampled in the South Cascades (sites 5378, 5595N, 5595S, 6000; Fig. 4). Table 1 lists all basins and their respective treatments. An extensive site selection process was conducted over two years to identify all non-fish-bearing basins meeting *a priori* selection criteria. Once all basins meeting these criteria were identified researchers worked with landowners to determine which basins were available for inclusion in the study as either a reference (not harvested) or a treatment (harvested according to a specified prescription) basin. For a detailed description of selection criteria and the site selection process see McIntyre et al. (2009). After determining a list of all suitable non-fish-bearing basins, sites were blocked

based on geographic location (location within the Olympic, Willapa Hill or South Cascade physiographic regions). The exception was in the Willapa Hills, where selection of basins for blocks was restricted by unavoidable logistical issues (land ownership and willingness of landowners to enact particular harvest treatments). As a result, inconsistencies exist among blocks with regard to spatial proximity. However, genetic cluster analyses will give insight into whether genetic differences exist within blocks that need to be accounted for. Within each block, treatments were randomized unless prevented by these same logistical constraints. We selected additional sites within the Olympic and South Cascade regions to serve as sites for the genetic cluster analysis to estimate genetic connectivity across the entire study area and within individual regions (see Objective 4 for details). We did not need to sample additional cluster sites in the Willapa Hills because we already had ten sites chosen for this region for the Type N Study.

Up to 50 individuals each of Coastal tailed frog and Cope's giant salamander were sampled from each treatment basin. We chose this sample size to attain the level of power required to confidently describe trends in genetic diversity, and in particular, to provide a robust baseline with which to compare to future post-harvest sampling. At nearby sites likely to be connected genetically ("cluster sites"), our desired sample size was 30 individuals, as these sites would not be directly impacted by application of the treatments to the study basins. Rather, we sampled cluster sites primarily to detect any changes in extent or directionality of gene flow that might be altered by forest practice harvest treatments applied to the study basins. Sampling occurred from 2006-2008. WDFW crews sampled each basin consistently each year, utilizing two sampling methods designed to detect both tailed frogs and giant salamanders. Sampling was conducted diurnally between 0700 and 1900 h, from 7 July – 1 November. Light-touch sampling was conducted along systematically selected stream reaches from the fish-end-point and upstream along every tributary to the headwall. A minimum of 25% of the stream length in each basin was sampled each year. Light-touch sampling involves turning all surface cover objects within the stream channel that are small cobble-sized (64 mm) and larger and visually searching for amphibians. Rubble-rouse sampling was conducted in 20 1-m sample plots randomly placed in each of the first 20 10-m intervals above the fish end point. Plots were restricted with block nets and intensively sampled by removing all cover objects larger than small gravel, (≥32 mm) and excavating down to 30 cm or bedrock. For both methods, substrates were returned to their original positions. Every effort was made to collect tissues from individuals distributed throughout the entire stream network from the fish-end-point upstream and including all tributaries. Where more than adequate numbers of amphibians were encountered, tissues were collected from the first individual of each genus encountered within each 10-m stream reach. This was important in the event that the two giant salamander species were differentially distributed throughout the stream network. This was also done in an attempt to minimize sampling full siblings. For example, when multiple tailed frogs were encountered in close proximity, only the first individual encountered in the area was sampled. WSU crews

also visited treatment basins that had low sample sizes in an attempt to increase sample size. WSU crews used a protocol that used kick sampling and visual surveys to locate individuals. WSU surveys started at the fish end point and continued upstream to the end of the basin or until sufficient sample size was reached. Therefore, sites with low sample size are likely to have actual low abundance as multiple sampling visits were conducted at each site. While inter-annual differences may influence the data it is unlikely to result in misleading genetic conclusions, as the pre-harvest sampling period (3 years) for the two focal species is considerably shorter than the generation time of either species (7-8 years). We obtained tissue from larval tailed frogs and giant salamanders using tail clips and we used mouth swabs (Goldberg et al. 2003) and toe clips to collect tissue from metamorphosed tailed frogs. Finally, because Cope's giant salamanders are difficult to distinguish from Coastal giant salamanders in the field (Nussbaum 1976, Good 1989), giant salamander samples included both Coastal and Cope's giant salamanders, as indicated by downstream genetic analyses. In fact, as previously stated, one of the four main objectives of the genetics portion of the study was to correctly identify individual giant salamanders to species. As a result, we included analyses for Coastal giant salamanders.

2.1.2. Laboratory Methods

We extracted DNA from all tissue samples using the Qiagen DNEasy Tissue Kit (Qiagen, Inc.). To develop microsatellite primers to use for genetic analyses, we sent extracted DNA from tailed frogs collected from the Olympic region and from Cope's giant salamander collected from the South Cascades region to Ecogenics GmbH molecular marker services. Ecogenics developed microsatellite primer sequences for 15 loci in each species. For the tailed frog samples, 13 loci could be easily amplified; the remaining 2 could not be amplified consistently (Spear et al. 2008). For Cope's giant salamanders, we used 11 of the loci in analyses; the other 4 loci contained a number of non-specific alleles that overlapped with alleles specific to Cope's giant salamanders, and thus led to unreliable scoring of genotypes. Additionally, nine of the Cope's giant salamander loci cross-amplified successfully in the Coastal giant salamander (Steele et al. 2008). We used multiplex PCR (polymerase chain reaction) amplification for both species using the Qiagen Multiplex PCR kit (Qiagen Inc.). Multiplex PCR allowed us to run PCR reactions for several microsatellite loci in the same mix, thereby reducing the overall number of PCR reactions required to conduct the analyses. We describe the specific PCR conditions and multiplex panels for both tailed frog and giant salamander loci in Appendices 1-2. All PCR products were run on an ABI 3730 sequencer at the Washington State University LBB1 core facility and then genotyped using GeneMapper 3.7 software (Applied Biosystems, Inc.).

2.1.3. Genetic Analyses

We formatted all genotypic data for the program CONVERT (Glaubitz 2004), which in turn reformats data for a number of popular genetic programs, including those used in

this study. We first tested that each locus conformed to Hardy-Weinberg expectations at each sampling site. One primary assumption of Hardy-Weinberg equilibrium is random mating across a population. For sites such as the individual streams for each Type N study basin, this is a reasonable assumption and any deviations from Hardy-Weinberg equilibrium might be due to processes such as inbreeding or substructure within streams, either of which would be important to identify. We tested for Hardy-Weinberg equilibrium using an exact test in GENEPOP v 3.4 (Raymond and Rousset 1995). Second, we tested for linkage disequilibrium among loci. Linkage disequilibrium is an estimate of statistical association among loci, which can be caused via physical linkage due to close proximity on a chromosome or population processes related to drift or selection. Tests of whether pairs of loci are in linkage disequilibrium are needed to ensure that each microsatellite locus can be considered statistically independent. If alleles at any two of the tested loci are significantly correlated with each other, then overall genetic results may be biased. We also used GENEPOP to perform tests of linkage disequilibrium.

Further, we tested for presence of null alleles, which can occur with microsatellite loci and produce misleading results. A null allele is an instance when a certain allele fails to amplify, and as a result, leads to individuals with missing data (when null alleles are homozygous) or an excess of apparent homozygotes that occur when the null allele is part of a heterozygous pair. The most common reason for a null allele is when a mutation occurs at the site that the microsatellite primer anneals to during PCR reactions. Thus, the annealing site is “unrecognizable” to the primer and replication of that fragment cannot occur. Such mutations are most likely in populations that have diverged from the population that was used to develop the microsatellite markers. In our case, because microsatellite development used individuals from only one region for each species, we may encounter null alleles if the populations in the other regions have been genetically isolated for some time. We estimated the occurrence of null alleles at each locus for each site using the software FreeNA (Chapuis and Estoup 2007).

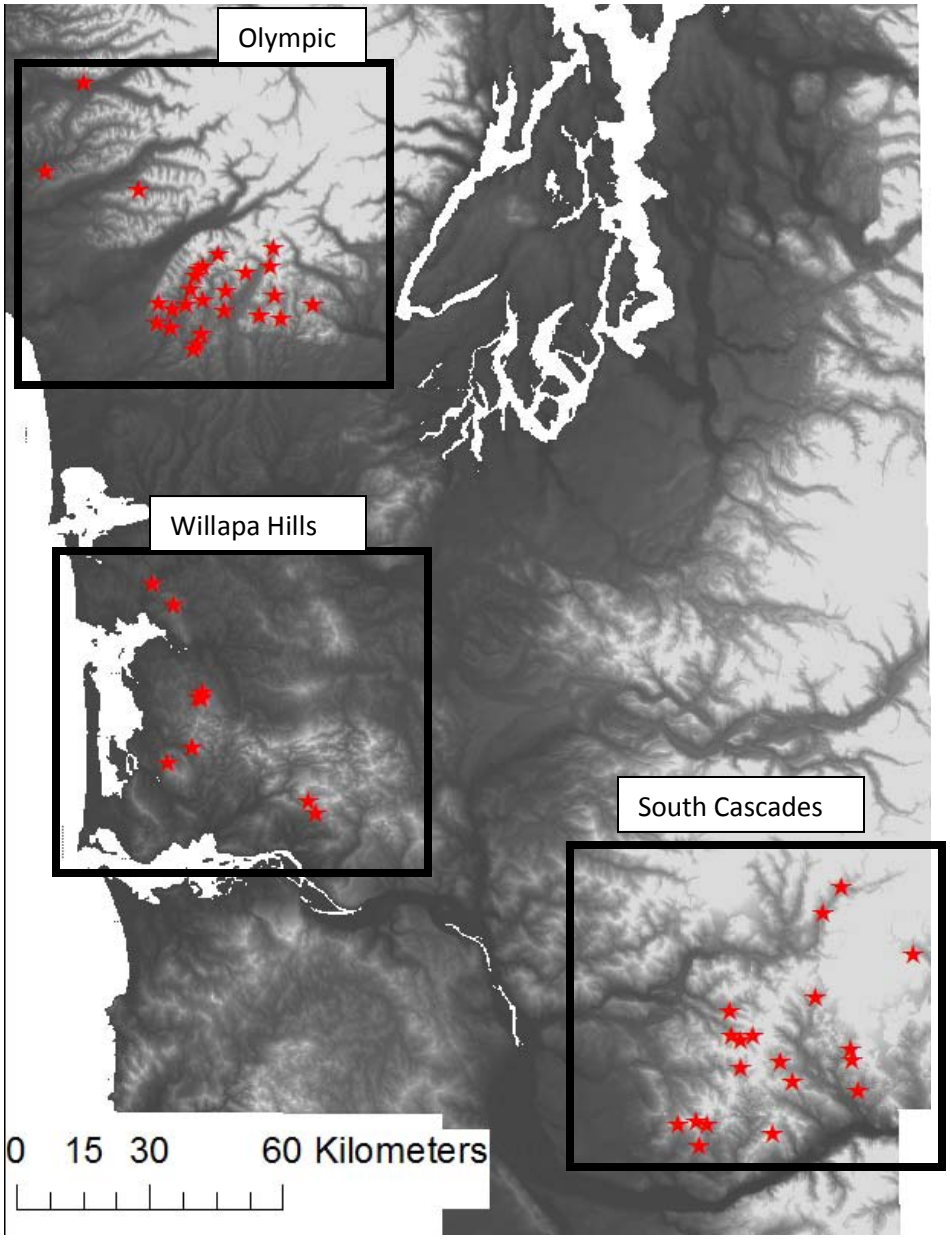


Figure 1. Overview map of Type N study area with each region designated by a black box. Background is a USGS Digital Elevation Model (DEM) with lighter colors as higher elevations. White patches indicate background areas with zero elevation.

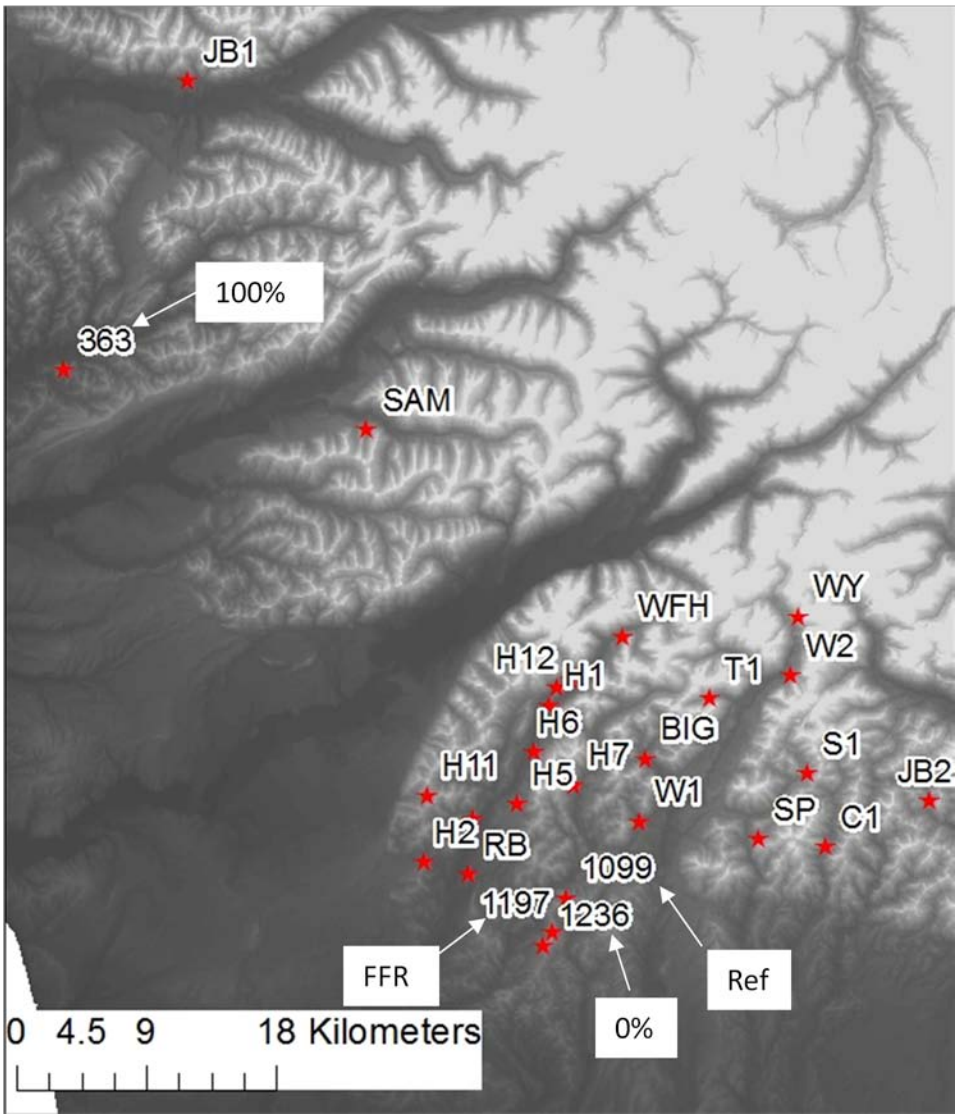


Figure 2. Map of both Type N basin (sites 363, 1099, 1197 and 1236) and cluster sites across the Olympic region. Text box and arrow represents assigned treatment for each basin.

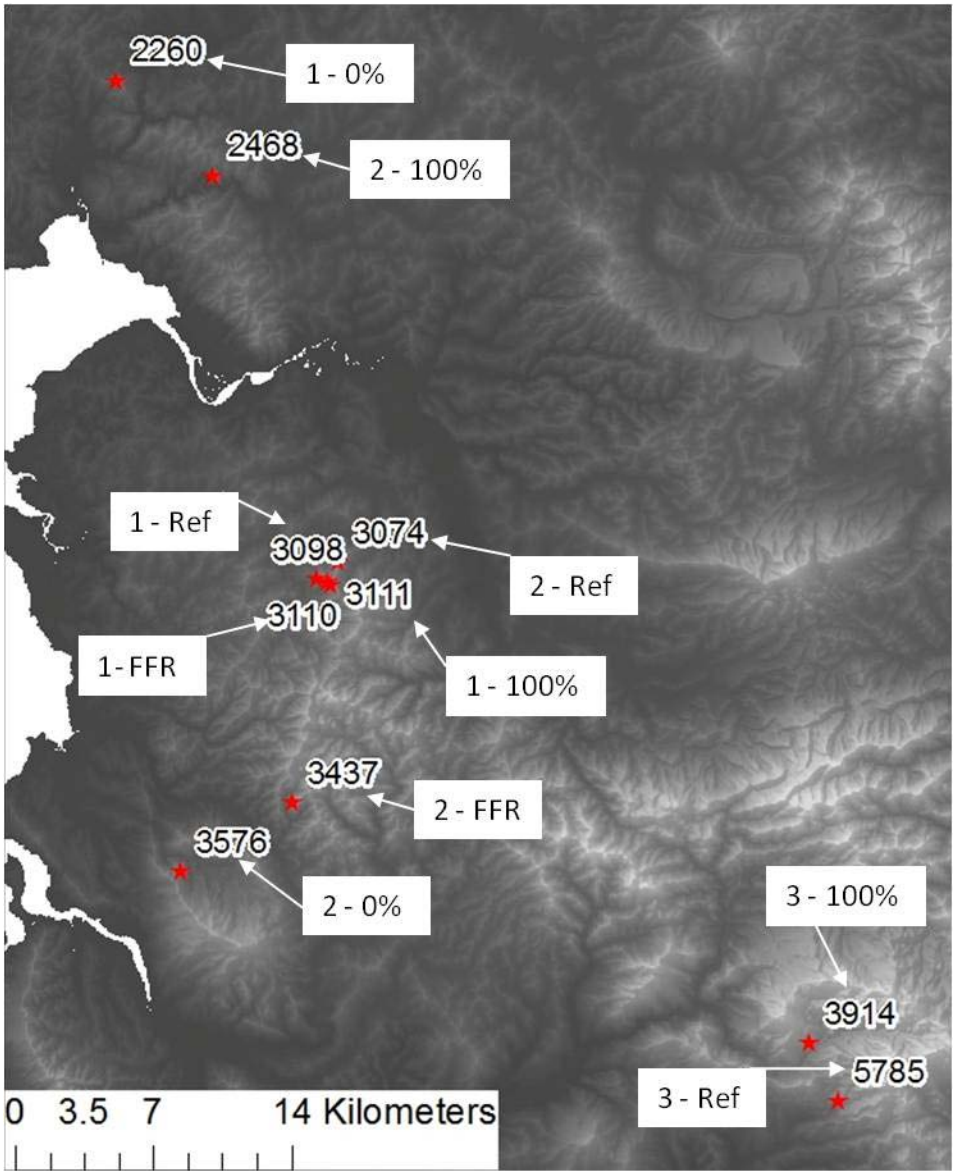


Figure 3. Map of Type N basin sites across the Willapa Hills region. Text boxes with areas indicate block and treatment types (number before dash is block number, followed by treatment type).

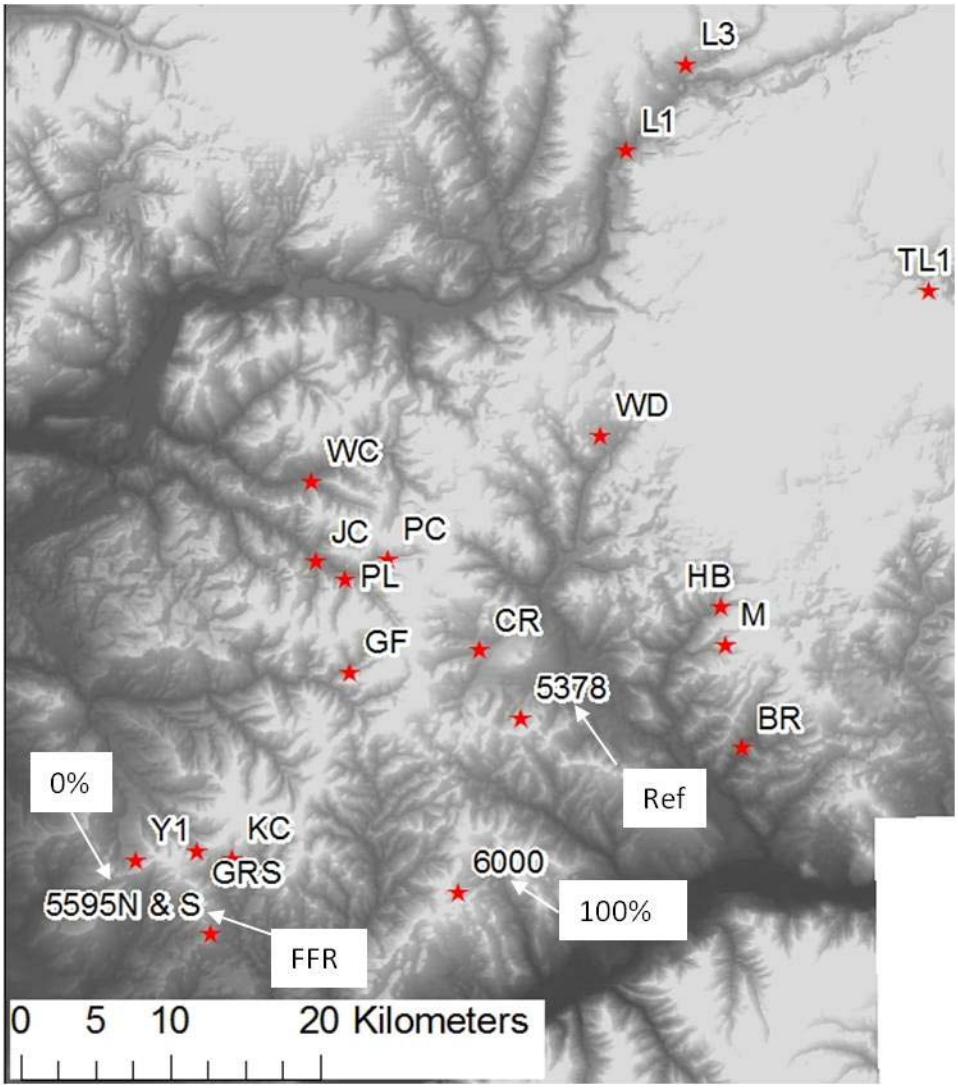


Figure 4. Map of Type N basins (sites 5378, 5595 N, 5595S, and 6000) and cluster sites across the South Cascades region. Text box and arrow represents assigned treatment for each basin.

For both tailed frogs and giant salamanders, most sampled individuals were larvae; this increases the likelihood that full siblings were sampled. If a sample has a proportionally large number of full siblings by chance, then sites may appear as inbred or as subdivided from other populations, even if the population as a whole is randomly mating and/or not subdivided. We identified full sibling pairs using a maximum likelihood approach implemented in the software COLONY (Wang 2004). This method was demonstrated to be appropriate for larval amphibians by Goldberg and Waits (2009). We identified the number of full sibling pairs per site, and in such cases we only included one family member for our analyses. Therefore, this method might create unequal

sample sizes among basins, and as a result, we account for unequal sample sizes in our estimates of genetic diversity (see Objective 2 methods).

2.2. Objective 1 Methods: Distinguishing giant salamander species

The first objective, determining the species identity of giant salamander samples, used a method that did not require us to identify *a priori* "pure" individuals belonging to each species. The program NEWHYBRIDS (Anderson and Thompson 2002) can accomplish this using a Bayesian method that assigns a posterior probability of an individual belonging to each species. With 10 or more microsatellites, probabilities can usually be assigned with very high probability (typically above 99%) if individuals are pure bred. Therefore, this program allowed us to confidently distinguish Cope's giant salamander individuals from the Coastal giant salamander individuals. In addition, if hybridization has occurred between the two species, the method can determine the probability of different types of hybrids (F_1, F_2, backcross Cope's [BC_{Dc}], or backcross Coastal [BC_{Dt}]). If any hybrids were observed, they were removed from the dataset and were not analyzed further.

Table 1. List of Type N basins by region, experimental block and buffer treatment. UTM coordinates are in NAD 83, Zone 10 and represent field verified fish end points for each basin. Numbers are used to identify sites in figures of clustering analysis results presented later in this report. FFR represents the current Forest Practices prescription buffer.

#	Basin/Site	Region	Block	Treatment	UTME	UTMN
1	363	Olympic	Olympic	100%	409832	5277868
2	1099	Olympic	Olympic	Reference	444753	5241629
3	1197	Olympic	Olympic	FFR	443819	5238402
4	1236	Olympic	Olympic	0%	442850	5237549
14	2260	Willapa Hills	Willapa 1	0%	433824	5184969
15	2468	Willapa Hills	Willapa 2	100%	438413	5179481
16	3074	Willapa Hills	Willapa 2	Reference	445225	5160174
17	3098	Willapa Hills	Willapa 1	Reference	443721	5159250
18	3110	Willapa Hills	Willapa 1	FFR	443750	5159054
19	3111	Willapa Hills	Willapa 1	100%	444223	5158712
20	3437	Willapa Hills	Willapa 2	FFR	442298	5148062
21	3576	Willapa Hills	Willapa 2	0%	442297	5148060
22	3914	Willapa Hills	Willapa 3	100%	469149	5135507
23	5785	Willapa Hills	Willapa 3	Reference	470128	5132348
30	5378	South Cascades	South Cascades	Reference	577925	5073227
25	5595N	South Cascades	South Cascades	0%	557451	5058323
26	5595S	South Cascades	South Cascades	FFR	557388	5058295
24	6000	South Cascades	South Cascades	100%	574450	5061390

2.3. Objective 2 Methods: Providing baseline levels of genetic diversity

2.3.1. Summary indices of genetic diversity

We estimated summary measures of genetic diversity for each sampled site for each species. Specifically, we estimated allelic richness, observed and expected heterozygosity and the inbreeding coefficient (F_{IS}). Allelic richness is a measure of the number of alleles present per locus per site corrected for the variance in sample sizes among sites. Observed and expected heterozygosity are additional measures of diversity. Expected heterozygosity is the percentage of heterozygotes that are expected to occur based on Hardy-Weinberg equilibrium and the number of alleles present. Observed heterozygosity is the percentage of heterozygotes that are actually found at the sampling site. F_{IS} is estimated based on the difference between observed and expected heterozygosity in a population, and value significantly different from 0 suggests a population is out of equilibrium. This metric ranges from -1 to 1, with positive values indicative of some degree of inbreeding and negative values suggestive of outbreeding. Values not different from zero suggest random mating. Finally, we estimated F_{ST} as a measure of genetic diversity among sites, and assessed significance of pairwise F_{ST} values. F_{ST} (range: 0-1) is positively correlated with differentiation, and thus low values indicate sites that are similar genetically. All genetic diversity parameters were estimated using either the software GDA (Lewis and Zaykin 2001) or FSTAT (Goudet 2001).

2.3.2. Estimate of effective population size

Effective population size (N_e) is a more meaningful metric of population viability than census population size. Effective population size is a measure of the number of breeding individuals in a population. Thus, a measure of effective population size can provide insight into its long-term viability, as well as predict the rate of inbreeding and loss of genetic diversity. We used a Brownian motion microsatellite model in MIGRATE 3.0.3 (Beerli 2008), which uses a coalescent approach (essentially following lineages back in time) to estimate the parameter θ, which is equal to $4N_e\mu$, with μ representing the microsatellite mutation rate. While the exact microsatellite mutation rate for each locus is unknown, a generally accepted estimate for microsatellites is 5×10^{-3} (Busch et al. 2007). However, it should be noted that mutation rate is only an estimate and likely subject to error. This would influence the exact value of N_e estimated, but would not change the relative differences among sites. Therefore, from θ we can estimate the differences in effective population size among sites, as well as a 95% confidence interval around the maximum likelihood estimate. Due to time limitations and memory limitations we could not estimate migration rates among basins, so we ran each Type N basin individually for each species to estimate effective population sizes. Each run consisted of 10 short chains of 10,000 generations each (sampled every 20 generations) and three long chains of 100,000 generations each (sampled every 20 generations). For each chain, a burn-in length of 10,000 generations was discarded.

The coalescent model used above represents historic effective population size. To provide more recent estimate of effective population size, we used a technique based on approximate Bayesian computation summarizing eight different metrics known as OneSamp (Tallmon 2004, 2008). This method generates 50,000 simulations with the same number of individuals and loci as the real population and with an initial effective size chosen from a range specified by the user. Effective population sizes from simulated populations that have similar summary statistics as the real population are accepted and the final estimated N_e is inferred from a weighted local regression of simulated values. We always chose 2 as the lower limit of N_e for simulations and the upper limit was chosen based on the upper confidence limit of MIGRATE results.

2.3.3. Evidence for recent population declines

We used three tests to assess whether any sites had undergone recent declines in effective population size. Specifically, we tested for: 1) significant heterozygosity excess relative to equilibrium expectations (Cornuet and Luikart 1996); 2) shifted distribution of allele frequencies (Luikart et al. 1998); and 3) M-ratios lower than those expected under random mating (Garza and Williamson 2001). Heterozygosity excess is an ephemeral pattern that can only be detected for the first several generations after a decline. While a heterozygosity excess may seem counterintuitive, it initially occurs because allelic diversity is lost faster than heterozygosity; this results in the appearance of greater heterozygosity than expected based on the number of alleles in the population (Cornuet and Luikart 1996). This method typically has low power unless declines are 90% or greater (Cornuet and Luikart 1996) and thus detected declines are likely to be severe. We assumed a two-phase mutation model with 10% multistep mutations. The choice of mutation model can influence the result of heterozygosity excess test in a predictable fashion (Cornuet and Luikart 1996). The two-phase model we use is the most conservative test that is realistic for our microsatellite loci. We used a Wilcoxon sign-rank test to detect significant excess in the software BOTTLENECK (Piry et al. 1999).

We also used BOTTLENECK to detect shifts in allelic distributions. Specifically, we tested for deviations from the expected distribution of microsatellite allele frequency of many rare alleles and few common alleles. As rare alleles are more likely to be lost during a bottleneck event, a population that has been through a recent decline will have an allele frequency distribution that is shifted so that common alleles occur at an even greater frequency.

Finally, the M-ratio is the ratio of number of alleles (k) to allelic size range (r). Allelic size range is simply the range in size of the microsatellite alleles that are determined by an automated DNA sequencer. To illustrate, if the microsatellite repeat is GATA and an individual has 8 repeats, their allele size would be 32. The allele size range (r) thus reflects the variation in number of repeats across all individuals in a population. The loss of alleles during a decline should be independent of allelic size, and thus k should decline faster than r. M-ratios are best suited for detecting more severe

declines over a longer period of time. Garza and Williamson (2001) suggested an overall critical M-ratio of 0.68, below which populations can be considered bottlenecked. However, sample size and effective population size can influence the critical value of the M-ratio. Therefore we used the Critical M software to determine the correct critical value for populations associated with Type N Study sites.

2.4. Objective 3 Methods: Testing for differences in genetic diversity among sites or blocks

To identify any pre-existing significant differences in genetic parameters, we conducted three comparisons of allelic richness, observed heterozygosity, F_{IS}, and effective population size (estimated through MIGRATE). We compared basins at three levels: among regions, blocks and treatment types. We used FSTAT v 2.9.3 (Goudet 2001) to run comparisons of allelic richness, heterozygosity, and F_{IS} using a randomization procedure with 10,000 permutations to assess significance. If we detected a significant difference by group for a variable, we tested all pairwise comparisons using the same randomization procedure to identify the specific comparisons that were significant after Bonferroni correction. We compared values for both estimates of effective population size using an ANOVA. To meet ANOVA assumptions, we conducted tests for normality using the R package normtest. If normality assumptions were violated, we log-transformed the variables and tested whether this corrected deviations from normality. For this test, if we detected a significant effect with the ANOVA, we used a t-test with a Bonferroni correction to assess which pairwise comparisons were different from each other. Second, we used simulations and population genetic expectations to predict the magnitude of decline necessary to detect changes in genetic diversity. To model changes in allelic richness, we used a population genetic simulator based on the current allelic diversity and number of loci of our data for each species. The simulation consisted of one population for one generation, with the number of loci and maximum number of alleles set to the specific values generated for each species in this study. We assumed equal sex ratio and ran 999 simulations for a number of population sizes. All simulations were run in EASYPOP 1.7 (Balloux 2001). To predict change in heterozygosity, we used Wright's equation for genetic drift (Wright 1969):

$$H_t = H_0\left[1 - 1/\left(2N_e\right)\right]^t$$

Where H_t is the final heterozygosity, H_o is the average heterozygosity of the study species from this study, N_e is the effective population size, and *t* is the generation time (in this case, 1). We used this equation for a range of possible effective population sizes. For both allelic richness and heterozygosity, we estimated effect size and identified the population size necessary to get a statistically significant result at the 0.05 level.

2.5. Objective 4 Methods: Estimating genetic clusters

We collected samples at an additional number of sites proximal to treatment basins to estimate the number and size of genetic clusters for each species. The genetic cluster size represents the extent of gene flow across the study areas and elucidates how dispersal may influence experimental results, thereby providing insight for spatial scale of management. Based on preliminary data from the Olympics and South Cascades, we predicted a maximum cluster size of 10-20 km for tailed frogs (Spear and Storfer 2008) and less than 5 km for Cope's giant salamander (Steele et al. 2009). The most important variable for the cluster analysis was spatial proximity to previously selected treatment basins. Therefore, we attempted to sample all accessible streams that were within 20 km of all treatment basins, although this was exceeded if suitable sites were difficult to find. We identified sites primarily by driving roads within the designated radius of Type N treatment sites and stopping at all streams that appeared permanent and contained suitable amphibian habitat (flowing water, riffles, pools, etc.). To maximize efficiency of sampling, we attempted to obtain sufficient sample size in one survey, and thus sites were only chosen if the focal species were found within the first 10-15 minutes. As our primary goal for this objective was obtaining sufficient number of samples, we did not standardize survey times or stream length, although most surveys lasted 1-2 hours and covered 50-100 meters of stream length. We only conducted cluster site sampling for the Olympics and South Cascades; however, the number of treatment basins sampled in the Willapa Hills allowed for a cluster analysis of this region as well.

We used a Bayesian population clustering algorithm, STRUCTURE 2.3.1 (Pritchard et al. 2000) as our primary tool to determine genetic cluster size. This analysis allows us to deduce which sites are exchanging migrants and thus estimate the spatial scale of gene flow. Additionally, conducting this analysis for each study region will allow us to infer whether the extent of gene flow varies across the entire Type N Study area. In STRUCTURE, log likelihood values for each run at a hypothesized K (number of clusters) are used to determine the most likely number of population groupings. However, estimating K by raw likelihood values can be problematic because as K increases, likelihood values tend to plateau after the true K is reached with slight improvements in value (Evanno et al. 2005). Therefore, choosing a clustering group based solely on the highest likelihood value may not make biological sense. To address this, Evanno et al. (2005) developed a method that estimated ΔK, the second order rate of change, to identify the value of K that had the greatest increase in likelihood with the lowest variance. The one problem with this method is that it does not allow for the evaluation of $K = 1$, since ΔK cannot be estimated for this value. However, if K truly equals one, then iterations assuming $K = 2$ will assign individuals to each cluster equally.

Finally, STRUCTURE often creates hierarchical subdivisions; clusters identified in one run of the program can often be further subdivided in a subsequent run that only includes sites that belong to the initial cluster. As a result, we ran additional iterations of STRUCTURE for all identified clusters until the program indicated that a group of sites

equaled 1 cluster, or when *K* equaled the number of sites in the run. All STRUCTURE results will be presented as bar plots by sampling site. Different clusters are denoted in the bar plots by different colors and the probability that individuals at a site assign to a particular cluster is determined by the amount of that color at each site in the bar plot. For each value of *K* we ran five replicates. Each replicate consisted of 100,000 iterations with 10,000 burn-in runs, which was sufficient for convergence in our dataset. For each hierarchical analysis, we used a range of *K* ranging from 1-10. We used an Analysis of Molecular Variance (AMOVA; Excoffier et al. 1992) to verify the hierarchical clusters indicated by STRUCTURE. An AMOVA is simply an ANOVA based on genetic data. In particular, the analysis partitions the amount of genetic variation due to between groups, individual sites within groups, and individuals within sites. Additionally, the analysis tests the significance of each component. If we have clustered sites into the proper genetic groups, then variation among groups should be a greater percentage than variation among sites within groups. AMOVA analyses were implemented in the software ARLEQUIN 3.1.1 (Excoffier et al. 2005).

3. Results and Conclusions

3.1. Objective 1 – *Dicamptodon* species identification and hybrids

We are able to assign individuals to either pure species group with high probability (98% of individuals that were assigned to either species had a probability greater than 0.90). As expected, we identified both species of giant salamanders at sites in the South Cascades and Willapa Hills (Table 2). We did not detect Coastal giant salamanders in the Olympics, which is consistent with previous observations from that region and the documented range limit for Coastal giant salamanders (Welsh 2005). Interestingly, the northern Willapa Hills site numbers 2260 and 2468 (Fig. 3) contained only Cope's giant salamanders, which suggests the species boundary of Coastal giant salamanders is at the Willapa River. A geographic trend in species composition also seems to occur in the Willapas (Fig. 3). Sites centrally located (3098/3110/3111) and in the southeast (3914/5785) were predominantly Coastal giant salamanders, whereas the southwest pair of sites (3437/3576) were dominated by Cope's giant salamanders.

We also detected 31 hybrid individuals, almost exclusively in the Willapa Hills. Four hybrid categories were possible: F_1 (offspring of pure Cope's and Coastal), F_2 (offspring of two hybrid individuals), backcross to Cope's (offspring of hybrid individual and pure Cope's) and backcross to Coastal (offspring of hybrid individual and pure Coastal). We detected 31 hybrid individuals belonging to only two of these hybrid types: F_2 (26) and backcross to Cope's (5), with the majority in the F_2 type. This means that we have no evidence for hybridization between the two species in the current generation. The posterior probabilities of the assigned F_2 hybrids ranged between 0.48-0.99. Eighteen of the 26 putative hybrids had posterior probabilities 0.75 or greater. Therefore, some of our assigned hybrids may be misidentified, but the majority had high posterior

probabilities for the F_2 classification. Although we cannot say when in the past this hybridization occurred, it is possible that it resulted from recent droughts that may have forced increased contact between the two species in the Willapa Hills. Consistent with this hypothesis is the fact that WDFW personnel have observed over 50 post-metamorphic *D. copei* during this time period. Nonetheless, this analysis clearly suggests that hybrid offspring are viable with a possible differential in mating success among hybrid types.

The presence of the two species in several treatment sites may complicate our ability to interpret post-treatment results, but it also provides opportunities for additional insight. Potentially the biggest challenge is the disparity in numbers between the two species, and how some of the Willapa sites are arranged into blocks. The Willipa 2 block is completely composed of sites that are dominated by Cope's giant salamanders and thus this block should provide inference for this species. However, the Willapa 1 block is composed of one site (2260) with exclusively Cope's giant salamanders, while the other 3 basins in this block have a greater abundance of Coastal giant salamanders. This could cause a problem after experimental treatment, particularly if a significant change occurs in 2260 but not the other 3 sites; that is, treatment effects may be confounded if the two giant salamander species respond differently to disturbance. We have little reason to believe that hybridization will greatly affect experimental results, as low numbers of hybrids were present, and introgression appears to be historic and not ongoing. It will be interesting, however, to see whether the relative abundance of species or number and type of hybrids changes from the observed baseline levels in post-treatment samples.

Table 2. Numbers of giant salamander individuals belonging to four classes: pure Cope's giant salamanders, pure Coastal giant salamanders, F_2 hybrids, and backcross to Cope's hybrids (Bc_{cope}).

Region	Site	Cope's	Coastal	F_2	Bc_{cope}	Total
Olympics	363	31	0	0	0	31
	1099	62	0	0	0	62
	1197	28	0	0	0	28
	1236	44	0	0	0	44
	Total	**165**	**0**	**0**	**0**	**165**
Willapa Hills	2260	96	0	0	0	96
	2468	59	0	0	0	59
	3074	21	4	0	0	25
	3098	12	41	0	0	53
	3110	22	78	2	2	104
	3111	22	137	5	0	164
	3437	67	5	1	3	76

Region	Site	Cope's	Coastal	F_2	Bc_{cope}	Total
	3576	37	16	6	0	59
	3914	6	71	6	0	83
	5785	28	40	5	0	73
	Total	**370**	**392**	**25**	**5**	**792**
South Cascades	5595N	19	66	0	0	85
	5595S	69	102	1	0	172
	5378	52	31	0	0	83
	6000	100	107	0	0	207
	Total	**240**	**306**	**1**	**0**	**547**
	Grand Total	**775**	**698**	**26**	**5**	**1504**

3.2. Equilibrium tests, presence of null alleles and sibling estimation

While these results do not clearly fit one of the study objectives, the analyses are necessary to ensure the data do not violate assumptions of analyses that are contained in the following sections.

3.2.1. Coastal tailed frogs

Tailed frogs were collected at 15 of the 18 total Type N basins, with no frogs found at 363, 5595N and 5595S despite multiple collecting visits. We found that despite a large number of larval individuals, relatively few full siblings were detected (Table 3). Eighty to 95% of all sampled individuals were from different families, which likely partly reflects the WDFW approach of non-proximate sampling of sequential individuals. Little evidence exists of a relationship between sample size and number of siblings, as the total number of individuals explained only 13% of variation in the number of sibling individuals (Fig. 5). This suggests that the difference in the number of samples collected at sites is representative of the actual differences in census size. In other words, large sample sizes at some sites (such as 5785, 3110, etc.) are not simply due to sampling many related individuals.

We detected deviations from Hardy-Weinberg equilibrium at several sites for 4 loci: A14, A2, A29 and A3. Sites out of Hardy-Weinberg equilibrium at these loci displayed a clear geographic pattern, with A14 and A29 out of equilibrium at South Cascades sites and A2 and A3 out of equilibrium at the Willapa Hills sites. For each of these loci, the Hardy-Weinberg deviations are most likely due to the presence of null alleles at each locus from individuals in the respective regions. Specifically, across the South Cascades, A14 had an estimated 18% of alleles characterized as null, with 27% for locus A29. Similarly, within the Willapa Hills region, 18% of alleles were null for A2

and 10% for A3. Other than these locus/region pairs, we found no other evidence of null alleles. Because the presence of null alleles can provide misleading estimates of genetic diversity parameters, we eliminated A14 and A29 for South Cascades sites and A2 and A3 for Willapa Hills sites from further analyses. In addition, because we conducted statistical tests among all 3 regions for genetic diversity parameters, we only used the 9 loci the three regions had in common for estimation of allelic richness, heterozygosity and inbreeding. However, this reduced number of loci should still be sufficient for detecting significant changes in genetic diversity.

The fact that no evidence existed for null alleles in the Olympic region (the area used to sample individuals for marker development) strongly suggests that null alleles are due to mutation as described in the methods. This result also implies that tailed frogs on the Olympic Peninsula are genetically divergent from other tailed frog populations and have not regularly exchanged migrants in the recent past. This is consistent with findings from Nielson et al. (2006) who examined genetic variation in tailed frogs with allozyme and mitochondrial DNA markers (less variable genetic markers best used for historical inference) and found that Olympic sites were significantly divergent from Cascade sites. In fact, Nielson et al. (2006) argue that the Olympic population should be considered an Evolutionarily Significant Unit (ESU), which designates a lineage that contains genetic diversity unique to that population. In particular, they demonstrated that Olympic tailed frogs formed a monophyletic mitochondrial lineage (i.e. Olympic sites share a common ancestor that is not shared with any other site) and had high allozyme differences from other regions, including a unique allozyme allele only found in the Olympics.

Table 3. Sibling results and genetic diversity measures for Coastal tailed frogs at Type N Study basins. N is total sample size, Family # represents the number of unique families sampled at the site, A_r is allelic richness, H_e is expected heterozygosity, H_o is observed heterozygosity, F_{IS} is the inbreeding coefficient (* indicates significant after Bonferroni correction; adjusted *p*-value = 0.003), and N_e is effective population size (derived from θ, assuming a mutation rate of 0.005), along with the 95% confidence interval. N_e was estimated through either a coalescent approach (MIGRATE) or approximate Bayesian computation (ONESAMP).

Basin	N	Family #	A_r	H_e	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
1099	42	41	11.21	0.9	0.86	0.05	4750 (4300-5050)	24 (16-36)
1197	24	19	9.96	0.88	0.92	-0.04	3550 (3250-3900)	214 (122-552)
1236	15	12	9.89	0.9	0.91	-0.01	3500 (3100-4000)	44 (31-110)
2260	50	46	10.44	0.86	0.84	0.02	2700 (2500-2850)	28 (21-38)
2468	33	27	10.52	0.87	0.85	0.03	4600 (4250-4950)	7572 (2182-76162)
3074	26	23	11.99	0.9	0.86	0.04	7500 (6850-8250)	29676 (4653-1794483)
3098	26	22	11.79	0.91	0.86	0.06	5700 (5100-6350)	3729 (1209-31533)
3110	54	51	12.5	0.91	0.84	0.08*	6750 (6350-7150)	5175 (1256-38934)
3111	54	52	11.76	0.9	0.86	0.06*	6350 (6000-6750)	29292 (5318-1220356)
3437	39	34	12.22	0.9	0.88	0.03	9450 (8750-10150)	4721 (1195-35933)
3576	48	42	11.91	0.9	0.89	0.01	7300 (6850-7750)	57 (34-94)
3914	25	24	11.95	0.9	0.89	0.01	8450 (7800-9150)	15971 (2849-345476)
5785	61	52	12.37	0.91	0.89	0.03	7700 (7250-8100)	0.19 (0.18-0.20)
5378	28	24	11.98	0.85	0.82	0.02	6100 (5600-6700)	1254285 (0-infinity)
6000	48	41	11.47	0.85	0.84	0.02	4300 (4000-4600)	0.38 (0.36-0.4)

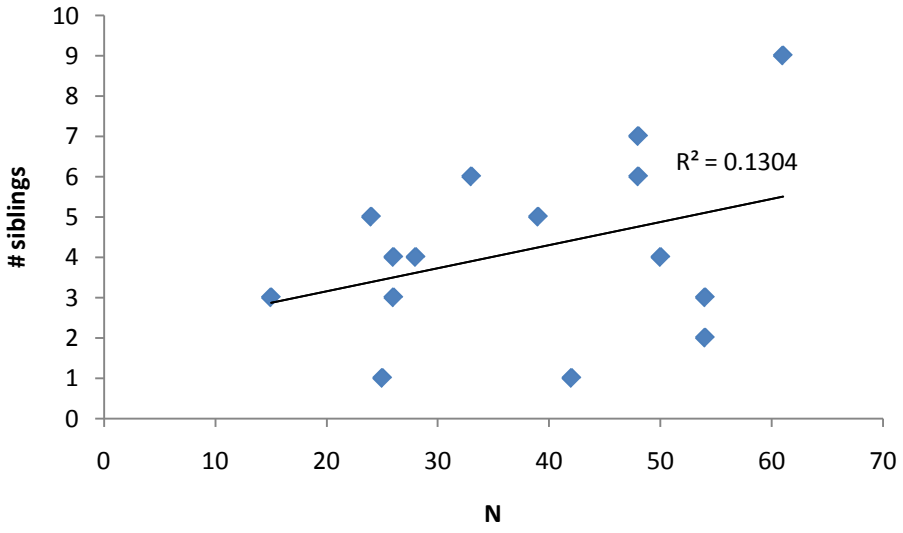


Figure 5. Relationship between number of individuals eliminated from data set as siblings and total sample size for Coastal tailed frogs.

Tests for linkage disequilibrium revealed that loci can be considered independent. No locus pair was significantly out of equilibrium at more than 2 or 3 sites (out of 47 total sampled sites). If two loci were truly statistically linked, we would expect to see disequilibrium for that pair at a majority of the 47 sites. As a result, once we accounted for null alleles, all loci used for tailed frogs are consistent with Hardy-Weinberg expectations and can be considered statistically independent, and thus are useful for further analyses.

3.2.2. Cope's giant salamanders

Cope's giant salamanders were collected at all Type N Study basins. In contrast to tailed frogs, full siblings often comprised a large portion of the total individuals of Cope's giant salamanders sampled (Table 4). Though broad variation existed in the number of siblings per site, the number of unique families at a basin averaged only 76% (compared to 88% for tailed frogs). The higher number of siblings may reflect fundamentally different microhabitat use than what was observed in tailed frog; we often found giant salamanders in low-flow areas including pools and under large debris, and their occurrence in such areas may reflect more sedentary behavior. Additionally, the number of full sibling groups is dependent on sample size for Cope's giant salamander. Extra siblings range from 0-34 per site, and total sample size explains 67% of the variation in sibling individuals (Fig. 6). Thus, the more total individuals that are sampled, the greater the likelihood that additional siblings will be sampled.

In general, all 11 loci conform to Hardy-Weinberg expectations, as each locus was in equilibrium at most sampling sites. However, 5 loci at basin 6000 violated equilibrium assumptions. This suggests that non-random mating is occurring, though the specific mechanism is unknown and could range from populations substructure to assortative mating Similarly, evidence of linkage disequilibrium among several locus pairs was found only at site 6000, again likely due to non-random mating. Finally, we found no evidence of null alleles at any loci for any region, and as such, all loci are included for further analyses.

Table 4. Sibling results and genetic diversity measures for Cope's giant salamanders at Type N Study basins. N is total sample size, Family # represents the number of unique families at the site, A_r is allelic richness, H_e is expected heterozygosity, H_o is observed heterozygosity, F_{IS} is the inbreeding coefficient (* indicates significant after Bonferroni correction; adjusted *p*-value = 0.003), and N_e is effective population size (derived from θ, assuming a mutation rate of 0.005), along with the 95% confidence intervals. N_e was estimated through either a coalescent approach (MIGRATE) or approximate Bayesian computation (ONESAMP).

Basin	N	Family #	A_r	H_e	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
363	31	17	3.03	0.54	0.48	0.12	150 (120-155)	48 (30-159)
1099	62	46	4.01	0.65	0.65	0	450 (440-500)	157 (104-425)
1197	28	17	3.92	0.66	0.65	0.01	300 (275-350)	17 (14-32)
1236	44	29	3.92	0.64	0.65	-0.01	300 (265-315)	26 (18-55)
2260	96	62	4.75	0.74	0.71	0.05	600 (585-650)	188 (118-529)
2468	59	47	4.94	0.75	0.66	0.12*	600 (575-650)	363 (207-1304)
3074	21	17	6.14	0.85	0.8	0.06	2250 (2000-2550)	44 (29-93)
3098	12	11	6.13	0.83	0.77	0.08	3150 (2750-3600)	26 (18-61)
3110	22	14	5.93	0.82	0.81	0.01	2400 (2100-2700)	36 (26-69)
3111	22	18	6.39	0.86	0.78	0.09*	3200 (2900-3600)	67 (40-155)
3437	67	41	6.56	0.88	0.84	0.04	1800 (1700-1950)	592 (276-2381)
3576	37	33	6.18	0.85	0.75	0.12*	1250 (1150-1350)	126 (71-355)
3914	6	6	5.84	0.8	0.74	0.09	1150 (950-1450)	10 (7-18)
5785	28	23	6.23	0.85	0.75	0.11*	2750 (2500-3000)	742 (365-2947)
5595N	19	16	5.58	0.83	0.75	0.09*	1950 (1750-2200)	36 (25-80)
5595S	69	62	5.66	0.83	0.76	0.09*	700 (650-750)	170 (88-524)
5378	52	37	4.14	0.69	0.66	0.04	400 (350-415)	122 (60-370)
6000	100	81	5.44	0.8	0.74	0.09*	1050 (1000-1100)	2798 (883-24298)

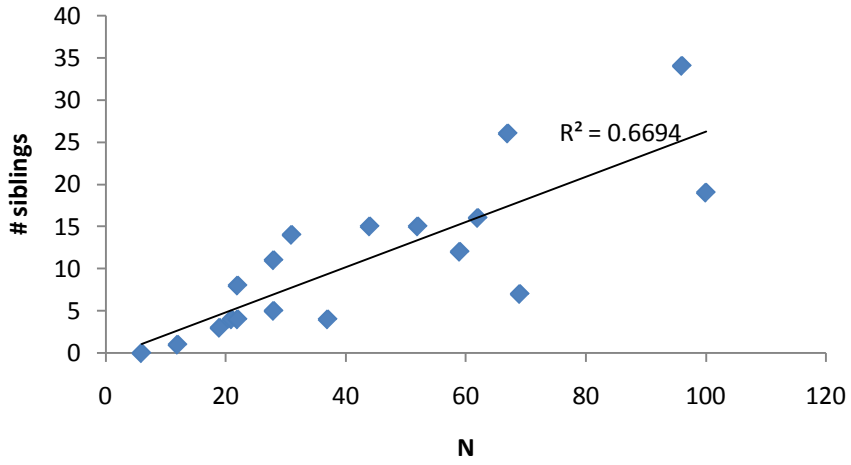


Figure 6. Relationship between number of individuals eliminated from data set as siblings and total sample size for Cope's giant salamanders.

3.2.3. Coastal giant salamanders

We detected high proportions of siblings across most sites for Coastal giant salamanders (Table 5). The average proportion of unique families was only 55%, and site 3110 had only 25% unrelated individuals. Furthermore, the number of siblings was strongly related to sample size, as $r^2 = 0.91$ (Figure 7), suggesting that high individual sample sizes will include many sibling pairs. The high number of sibling individuals may be related to moderately large clutch sizes, up to 200 eggs (Welsh 2005) which are larger than those laid by Cope's giant salamanders (average of 60; Jones and Bury 2005).

Three loci were consistently out of Hardy-Weinberg equilibrium at multiple sites. These loci were D05, D07, and D23. Additionally, basins 3914 and 6000 were out of Hardy-Weinberg equilibrium at several other loci besides the three above. As with basin 6000 for the Cope's giant salamander, the likely reason for equilibrium deviations is non-random mating, although the mechanism is unclear. Estimation of null alleles indicated that loci D05, D07 and D23 had high (>0.10) proportions of null alleles, likely resulting in deviations from equilibrium of those loci. Thus, these three loci were excluded from further analyses, leading to the use of 6 total loci for Coastal giant salamanders. Therefore, we probably have less power to detect changes in Coastal giant salamanders as compared to the other two species. Finally, no evidence existed of linkage disequilibrium among locus pairs.

Table 5. Sibling results and genetic diversity measures for Coastal giant salamanders at Type N Study basins. N is total sample size, Family # represents the number of unique families at the site, A_r is allelic richness, H_e is expected heterozygosity, H_o is observed heterozygosity, F_{IS} is the inbreeding coefficient (* indicates significant after Bonferroni correction; adjusted *p*-value = 0.004), and N_e is effective population size (derived from θ, assuming a mutation rate of 0.005), with the 95% confidence intervals. N_e was estimated through either a coalescent approach (MIGRATE) or approximate Bayesian computation (ONESAMP).

Basin	N	Family #	A_r	H_e	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
3074	4	4	3.83	0.67	0.79	-0.21	600 (400-1050)	5 (4-7)
3098	41	14	3.82	0.68	0.67	0.01	550 (450-650)	31 (23-72)
3110	78	20	3.38	0.61	0.62	-0.02	350 (300-400)	56 (35-156)
3111	137	53	3.96	0.69	0.65	0.07	500 (450-550)	347 (154-1085)
3437	5	4	2.33	0.4	0.54	-0.42	50 (40-75)	7 (5-13)
3576	16	10	3.68	0.7	0.56	0.19	550 (450-650)	18 (12-39)
3914	71	43	3.79	0.67	0.51	0.25*	650 (600-700)	455 (188-1996)
5785	40	23	3.75	0.67	0.61	0.1	400 (350-450)	106 (59-388)
5595N	66	35	3.26	0.57	0.49	0.14*	350 (300-400)	120 (72-572)
5595S	102	53	3.42	0.63	0.63	0	450 (400-465)	99 (62-332)
5378	31	14	2.67	0.43	0.42	0.02	200 (190-255)	17 (12-44)
6000	107	56	4.43	0.74	0.69	0.07	750 (700-800)	2260 (739-1582)

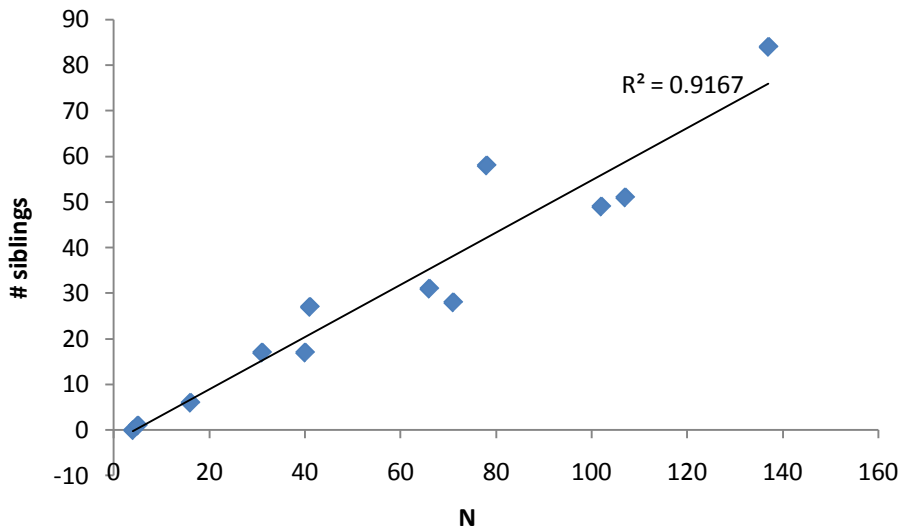


Figure 7. Relationship between number of individuals eliminated from data set as siblings and total sample size for the Coastal giant salamander.

3.3. Objective 2: Measures of genetic diversity and population size

3.3.1. Coastal tailed frogs

A. Summary measures of genetic diversity

Levels of genetic diversity in all Coastal tailed frog populations sampled were high (Table 3). The mean number of alleles per locus per site (once corrected for sample size) was greater than 11, and average expected and observed heterozygosity of 0.89 and 0.87, respectively. Since the observed heterozygosity was lower than the expected value, a positive, but low F_{IS}, averaging 0.03, was estimated. However, only two basins had F_{IS} values significantly greater than zero: 3110 and 3111. These two basins are adjacent, suggesting an unknown common cause. These high levels of F_{IS} must be considered in analysis of post-treatment genetic data, such that if significant levels of inbreeding are observed post-harvest at 3110 and 3111, it may be due to a prexisting condition. Significant F_{ST} values primarily corresponded to regional differences (Appendix 3), although there were some significant differences between sites within the Olympics and Willapa Hills, suggesting that genetic diversity is primarily apportioned by region.

Tailed frogs are an ancient lineage that has no closely related extant frog lineage to which to compare as to whether these levels of genetic diversity are typical. Similar high levels of heterozygosity and allelic richness are seen in the Rocky Mountain tailed frog (Spear 2008), so our results are at least consistent between the two tailed frog species. However, comparison with microsatellite studies among other frog species

suggests that these observed levels of diversity are much higher than expected for frogs in general. Ficetola et al. (2007) summarized 16 studies of frog microsatellite genetic diversity, and only one had levels of genetic diversity similar to that documented here. Most studies reported a mean allelic richness of 4-7 and heterozygosity of 0.4-0.6 (although differing sample sizes could lead to some differences in allelic richness). The high diversity detected in tailed frogs may reflect their long evolutionary history (a long time has been available to accumulate new alleles).

B. Effective population size

Based on coalescent MIGRATE analyses, tailed frogs appear to have very high effective population sizes across the entire study area, numbering between 3000-9000 (Table 3). Comparison with recent estimates of N_e (based on ONESAMP) demonstrates that many sites have similar or greater values as compared to the coalescent analysis. This is particularly true in the Willapa Hills region, where most sites have very high sizes. The Olympic region has significantly lower recent population sizes (based on non-overlap of confidence intervals), as well as sites 2260 and 3576 in the Willapas. Finally, sites 5785 and 6000 had extremely low sizes (less than 1). However, despite the differences in the two methods, it is clear that tailed frogs generally have a high effective population sizes Effective population size is often much smaller than the census population size, so the census population sizes associated with these N_e estimates are probably much larger. High effective population sizes undoubtedly play a large role in the high genetic diversity observed in these populations. High effective population sizes serve to buffer populations against loss of alleles due to inbreeding and genetic drift. Therefore, because such populations are less likely to undergo genetic changes due to drift, even small detected changes in post-treatment estimates of genetic diversity could reflect a strong treatment effect. However, smaller sizes in the Olympics suggest this area might respond to treatment effects of smaller magnitude, and sites 5785 and 6000 should be evaluated carefully due to the apparent current loss of N_e.

Table 6. Population bottleneck results for Coastal tailed frogs. H_e is the actual expected heterozygosity, H_{eq} is the expected heterozygosity under the assumed two-phase mutation model, and p-value is for the significance test of heterzygosity excess. Italics represent significant values.

Basin	**H_e**	**H_{eq}**	**p-value**	**Shifted?**	**M-ratio**	**M-crit**
1099	0.91	0.89	*0.0026*	normal	0.841	0.612
1197	0.90	0.88	0.0199	normal	0.712	0.520
1236	0.91	0.89	*0.0003*	normal	0.709	0.441
2260	0.88	0.89	0.4155	normal	0.738	0.656
2468	0.89	0.89	0.7935	normal	0.676	0.547
3074	0.91	0.90	0.3501	normal	0.679	0.473
3098	0.91	0.91	0.2886	normal	0.728	0.493

Basin	H_e	H_{eq}	p-value	Shifted?	M-ratio	M-crit
3110	0.92	0.91	*0.0034*	normal	0.812	0.615
3111	0.92	0.91	*0.0024*	normal	0.821	0.621
3437	0.91	0.91	0.4492	normal	0.769	0.522
3576	0.91	0.91	0.1602	normal	0.774	0.580
3914	0.91	0.90	0.0615	normal	0.714	0.469
5785	0.92	0.92	0.0415	normal	0.840	0.610
5378	0.86	0.90	0.1602	normal	0.689	0.502
6000	0.86	0.87	0.2597	normal	0.879	0.617

C. Declines in effective population size

We detected evidence of significant heterozygosity excess at 4 basins, but no evidence of declines due to shifted allele distributions or M-ratios (Table 6). The four basins with significant heterozygosity excess were 1099, 1236, 3110 and 3111. However, it is puzzling why these sites should not also have shifted allele distributions or low M-ratios. A likely confounding factor is migration. Bottleneck analyses assume closed populations, which is clearly not the case with tailed frogs, which have high levels of gene flow. Busch et al. (2007) conducted a study of kangaroo rats that had recently undergone a known demographic decline, but did not display signatures of bottlenecks using these three tests. They attributed this result to migration among their study sites, which would replenish rare alleles and obscure a bottleneck signature. Pope et al. (2000) used a simulation, based on a stepping stone migration model (which assumes that individuals only move to the closest suitable patch), to show that high migration could lead to heterozygosity excess, and lead to a false positive of population bottleneck. However, this explanation is somewhat unsatisfactory. This is because all basins have high migration rates and if high migration alone was contributing to false positives, we would expect all sites to display heterozygosity excess. Furthermore, the sites with significant heterozygosity excess also had some independent evidence of reduced size. Specifically, the significant inbreeding coefficient detected at 3110 and 3111 could be consistent with a very recent decline, and 1099 and 1236 have lower effective population sizes than the average. Obviously, none of these possibilities are compelling, and as such we cannot conclusively determine whether recent declines have occurred. Including the post-treatment samples in the future will be much more insightful, as temporal comparisons should be more sensitive to demographic changes (Luikart et al. 1998).

3.3.2. Cope's giant salamanders

A. Summary measures of genetic diversity

Overall, levels of genetic diversity in Cope's giant salamanders were lower than those documented for tailed frogs. Average allelic richness across all basins was 5. Expected heterozygosity and observed heterozygosity were 0.77 and 0.72, respectively.

Due to the deficiency of heterozygotes, a positive F_{IS} with an average value of 0.07 was found. Several individual basins had significantly high F_{IS} values, including: 2468, 3111, 3576, 5785, 5595N, 5595S and 6000. No strong geographic pattern appears to exist among sites with significant inbreeding, with the potential exception of 3 sites in the South Cascades. However, Steele et al. (2009) sampled 11 sites in the South Cascades in an area between sites 5595N, 5595S and 6000, and these sites displayed a low F_{IS}, so inbreeding does not seem uniformly distributed across the South Cascades region. Based on available data, Cope's giant salamanders appear to be primarily restricted to streams even as adults (Jones and Corn 1989), and the higher numbers of siblings detected suggests that individuals may not move frequently. Such limited mobility and observed proximity of full siblings could lead to an intrinsically higher prevalence of inbreeding than a species with a life history that promotes greater mobility. Pairwise measures of genetic differentiation (Appendix 4) indicate that almost all Cope's giant salamanders sampled at Type N basins are differentiated from one another. This suggests that if population sizes become small, Cope's giant salamanders are especially vulnerable to erosion of genetic diversity.

Overall, the genetic diversity of Cope's giant salamander is similar to that found in a number of other population genetic studies done primarily with ambystomatid salamanders (the salamander family thought to be most closely related to *Dicamptodon*). Such studies have found levels of heterozygosity ranging from 0.3-0.7 and a pattern of significant levels of F_{IS} in some, but not all, sites sampled (Spear et al. 2006, Eastman et al. 2007, Giordano et al. 2007, Chandler and Zamudio 2008, Purrenhage et al. 2009). The level of heterozygosity is much higher than observed values (0-0.244) in a population of Coastal giant salamanders in British Columbia (Curtis and Taylor 2003), although this population likely underwent a historic population bottleneck due to recent colonization after glaciations (Steele et al. 2006). Thus, Cope's giant salamander seems to have maintained a higher level of diversity relative to founder populations of *D. tenebrosus*.

B. Effective population sizes

The estimated effective population sizes based on coalescent analyses for Cope's giant salamander per site are much lower than those documented for tailed frogs (on average about 1/5 of tailed frog population sizes) and highly variable across sites (150-3150). In general, current effective sizes are significantly lower than the historic sizes, and most sites have N_e less than 200. There is also much more consistency among sites for Cope's giant salamander as compared to tailed frogs. Interestingly, sites 5785 and 6000 have the highest population size for Cope's giant salamanders, but extremely low for tailed frogs. Overall, we should expect smaller effective population sizes for Cope's giant salamander than for tailed frogs due to the former's more restricted habitat requirements (i.e. the post-metamorphic stages of tailed frogs can use terrestrial habitats, whereas most Cope's giant salamander adults are restricted to the stream environment). A geographic pattern seems to exist to the magnitude of effective population size.

Cope's giant salamanders reach their highest population size in the Willapa Hills and South Cascades and the lowest in the Olympics. This is contrary to expectations because Cope's giant salamanders are sympatric with Coastal giant salamanders in the Willapa Hills and South Cascades and may experience competition. This result also suggests that Olympic populations are more susceptible to genetic and demographic changes than in other areas, although the Olympic sites have a lower average F_{IS}, which could be due to behavioral avoidance of inbreeding due to prolonged small population sizes, a hypothesis in need of further testing. Finally, in comparison to tailed frogs, it appears that Cope's giant salamanders have been more susceptible to recent declines across their range, although it is unclear whether this is due to population subdivision or declines at each site.

Table 7. Population bottleneck results for Cope's giant salamanders. H_e is the actual expected heterozygosity, H_{eq} is the expected heterozygosity under the assumed two-phase mutation model, and p-value is for the significance test of heterzygosity excess. Italics represent significant values.

Basin	**H_e**	**H_{eq}**	**p-value**	**Shifted?**	**M-ratio**	**M-crit**
363	0.543	0.573	0.6812	normal	0.955	0.718
1099	0.66	0.71	0.9966	normal	0.833	0.706
1197	0.66	0.69	0.5508	normal	0.846	0.669
1236	0.64	0.68	0.5508	normal	0.870	0.700
2260	0.74	0.74	0.1826	normal	0.868	0.714
2468	0.75	0.75	0.2065	normal	0.817	0.701
3074	0.85	0.84	0.1602	normal	0.723	0.536
3098	0.83	0.83	0.1826	normal	0.692	0.432
3110	0.82	0.82	0.4155	normal	0.699	0.501
3111	0.86	0.87	0.6177	normal	0.764	0.514
3437	0.88	0.86	*0.0002*	normal	0.894	0.659
3576	0.85	0.83	0.2065	normal	0.807	0.657
3914	0.88	0.88	0.2783	normal	0.583	0.448
5785	0.85	0.87	0.3823	normal	0.718	0.568
5595N	0.83	0.83	0.2324	normal	0.718	0.538
5595S	0.83	0.84	0.6499	normal	0.876	0.712
5378	0.69	0.76	0.9939	normal	*0.656*	0.699
6000	0.80	0.87	0.9919	normal	0.724	0.723

C. Declines in effective population size

No widespread evidence exists for recent population declines across the Type N basins for Cope's giant salamanders (Table 7), in contrast to expectations based on effective population size estimates. Only two sites had a significant result for any of the

three bottleneck tests: 3437 had significant heterozygosity excess and 5378 had a low M-ratio. Basin 3437 had a high M-ratio and a normal allelic distribution, in contrast to the heterozygosity excess results. Due to their reliance on the stream environment, Cope's giant salamanders are more likely to follow a stepping-stone dispersal strategy, and thus stepping-stone migration to 3437 could potentially have led to the observed false positive (Pope et al. 2000). However, the explanation for the low M-ratio at 5378 is much more straight-forward. This site does not have a heterozygosity excess; in fact it has a strong heterozygote deficit. Low M-ratios without a corresponding heterozygosity excess is evidence of a more historic decline (Garza and Williamson 2001). Basin 6000 has a M-ratio that is barely above the critical level, and also has a heterozygote deficit. Finally, the sites of the Olympic peninsula all have strong heterozygote deficits (although higher M-ratios), suggesting historically low population sizes in this area as well. This is also consistent with the lower effective population sizes estimated across the Olympics and South Cascades.

3.3.3. Coastal giant salamanders

A. Summary measures of genetic diversity

Average values of genetic diversity in Coastal giant salamanders were lower than Cope's giant salamanders (average allelic richness = 3.5, average expected heterozygosity = 0.62 and average observed heterozygosity = 0.60) (Table 5). However, the levels of heterozygosity described here are much higher than a previous study of Coastal giant salamanders in British Columbia, which ranged from 0-0.244 (Curtis and Taylor 2003), although the BC population is a recent colonization, as described previously. Levels of F_{IS} were variable among sites, with 3 sites actually showing negative values (which indicates outbreeding, or mating among individuals less related than random), but two sites had significantly positive inbreeding (3914 and 5595N). Based on F_{ST} values, sites 3110 and 3111 were among the most differentiated, but there were several site comparisons between the Willapa Hills and South Cascades that have low F_{ST} values and thus low differentiation (Appendix 5). Note that sites 3074 and 3437 often had high F_{ST}, but values were not significant due to very low sample sizes and thus low power. It is somewhat surprising that Coastal giant salamanders would have lower levels of diversity than Cope's giant salamanders due to the greater dispersal capability of Coastal giant salamanders, which theoretically should reduce genetic drift and corresponding loss of alleles. Also surprising is the variability in F_{IS}, especially among sites that are very close together. Basin pairs 3110/3111 and 5595N/5595S are adjacent to each other, but have quite different inbreeding coefficients from each other, although the site pairs are not differentiated from each other. In addition, these results suggest that this variability must be considered in the post-treatment analysis.

B. Effective population sizes

The effective population sizes of Coastal giant salamanders associated with the respective treatment sites are much smaller than those found for either tailed frogs or Cope's giant salamanders, averaging only 450 based on the coalescent analyses, or roughly one third the mean value for Cope's giant salamanders. Current effective sizes are generally lower, with most sites less than 150. As with Cope's giant salamander, site 6000 has the highest population size. The overall low N_e is consistent with the high number of siblings detected for Coastal giant salamanders, which already demonstrates that census size is misleading for this species. The small population sizes may be largely due to late Pleistocene glaciations, as Steele and Storfer (2006) concluded that the area around the Columbia River valley was a refugium for this species, and populations expanded from this small refugium when glaciers receded. We would expect that population size would be small in a refugial area, and that perhaps local population sizes have not grown considerably since this historic bottleneck. However, presumably the glacial history should have affected Cope's giant salamander and tailed frogs similarly, but low population sizes are not seen for these species. Currently, these lower estimates of effective population size suggest that if Type N treatments do have a strong impact on stream amphibian populations, we might expect to see the greatest genetic change in Coastal giant salamanders.

C. Declines in effective population size

Widespread evidence of population declines based on M-ratios as well as two sites that had shifted allelic distributions was observed (Table 8). However, no significant heterozygosity excess was found at any sites, and only site 3437 had even a trend of heterozygosity excess. Interestingly, basin 3437 was the only site that had a significant heterozygosity excess for Cope's giant salamanders as well, suggesting that some recent event led to declines in giant salamanders in general at that site. Overall, however, strong evidence exists that Coastal giant salamanders in southern Washington have undergone a decline in the past, and that populations have not recovered, as indicated by low allelic diversity, low effective population sizes, heterozygote deficiencies and low M-ratios. These results underscore the possibility that this species may be particularly vulnerable to negative impacts due to chronically low population sizes in this region.

Table 8. Population bottleneck results for the Coastal giant salamander. H_e is the actual expected heterozygosity, H_{eq} is the expected heterozygosity under the assumed two-phase mutation model, and p-value is for the significance test of heterzygosity excess. Italics represent significant values.

Basin	**H_e**	**H_{eq}**	**p-value**	**Shifted?**	**M-ratio**	**M-crit**
3074	0.81	0.81	0.5000	shifted	0.474	0.436
3098	0.68	0.72	0.6563	normal	0.688	0.588
3110	0.61	0.69	0.9766	normal	*0.423*	0.637
3111	0.69	0.84	1.0000	normal	*0.660*	0.677
3437	0.61	0.59	0.1563	shifted	*0.621*	0.708
3576	0.70	0.70	0.5781	normal	*0.474*	0.551
3914	0.67	0.79	0.9219	normal	0.753	0.667
5785	0.67	0.72	0.9609	normal	*0.569*	0.642
5595N	0.57	0.69	0.7813	normal	*0.575*	0.667
5595S	0.63	0.68	0.9219	normal	*0.560*	0.679
5378	0.64	0.72	0.9375	normal	*0.590*	0.642
6000	0.74	0.85	0.9766	normal	*0.610*	0.679

3.4. Objective 3: Comparison of genetic diversity measures among groups

3.4.1. Coastal tailed frogs

A. Comparison among physiographic regions

Little evidence was found for differences among the three physiographic regions in measures of genetic diversity (Table 9). No significant differences were found in allelic richness, inbreeding coefficient or effective population size. The only significant difference is due to observed heterozygosity, although the magnitude of difference among the three regions was not large. Post-hoc pairwise analyses showed that the significant result was due to greater heterozygosity in the Olympics compared to the South Cascades (p = 0.037), and that the Willapa Hills was not significantly different from either the Olympics or South Cascades. We have evidence from null alleles and previous work (Nielson et al. 2006) that the three physiographic regions are represented by distinct evolutionary populations; thus it is interesting that whereas allelic frequencies differ between the regions, genetic diversity does not. In all likelihood, this stems from the high effective population sizes observed, which limit changes in genetic diversity caused by drift. Therefore, the different physiographic regions can be considered as replicates with respect to measures of genetic diversity for Coastal tailed frogs.

Table 9. Comparison of genetic diversity measures among regions for Coastal tailed frogs. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Region	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
Olympics	10.35 (0.74)	0.885 (0.03)	0.016 (0.05)	3933 (708)	94 (104)
Willapa Hills	11.745 (0.71)	0.865 (0.02)	0.037 (0.02)	6650 (1948)	9622 (11490)
South Cascades	11.724 (0.36)	0.834 (0.01)	0.021 (0)	5200 (1273)	627143 (886913)
p-value	0.058	*0.027*	0.555	0.089	0.638

B. Comparison among blocks

Not surprisingly, the results for the block comparison were very similar to the regional comparison, as two of the blocks were the same as the regions (Olympics, South Cascades) (Table 10). Again, the only measure that was significantly different was observed heterozygosity. The significant result was based not only on a lower value in the South Cascades than the Olympics (*p*-value = 0.037), but also on a lower heterozygosity in the South Cascades than the Willapa 3 pair (*p*-value = 0.035). No other pairwise comparisons were significantly different for observed heterozygosity. Thus, we can primarily consider the blocks to be replicates, and more importantly, the only two complete blocks in which tailed frogs were found in all four treatments (Willapa 1 and Willapa 2) are not significantly different for any parameter.

Table 10. Comparison of genetic diversity measures among blocks for Coastal tailed frogs. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Block	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
Olympic	10.35 (0.74)	0.885 (0.03)	0.016 (0.05)	3933 (708)	94 (104)
Willapa 1	11.624 (0.86)	0.848 (0.01)	0.053 (0.03)	5375 (1835)	9556 (13335)
Willapa 2	11.66 (0.77)	0.873 (0.02)	0.026 (0.01)	7212 (1993)	10506 (13150)
Willapa 3	12.157 (0.30)	0.89 (0)	0.022 (0.01)	8075 (530)	7986 (11293)
South Cascades	11.724 (0.36)	0.834 (0.01)	0.021 (0)	5200 (1273)	627143 (886913)

Block	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
p-value	0.134	*0.024*	0.537	0.072	0.761

C. Comparison among anticipated treatments

Among anticipated treatment types, no differences were found in any genetic diversity parameter (Table 11), including observed heterozygosity. This important result indicates that the sites chosen for different treatments are statistically equivalent with respect to levels of genetic diversity. Thus, we can be reasonably confident that any significant post-treatment effects in genetic diversity are due to that treatment and not due to confounding site-specific factor(s). Variance was low for measures of allelic richness and heterozygosity, but higher for F_{IS} and N_e. Thus, allelic richness and heterozygosity might be the most likely variables to detect statistical treatment effects. However, simulations suggest that a large reduction in effective population size is necessary to detect changes in allelic richness and heterozygosity. A reduction to an effective population size of 34 or lower is necessary to significantly reduce the number of alleles in one generation, and the reduction would have to be to seven individuals for a statistically significant change in heterozygosity in one generation. Thus, allelic richness is more likely to respond to a population change than heterozygosity, but there would still have to be a severe bottleneck given the large effective population sizes in tailed frogs. Thus, more subtle genetic measures, such as increasing proportion of full siblings, might be a better immediate measure of treatment effects. If a change in genetic diversity is detected, it would be a rather unequivocal signature of a strong change in the population.

Table 11. Comparison of genetic diversity measures among anticipated treatments for Coastal tailed frogs. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Anticipated Treatment	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
reference	11.867 (0.42)	0.865 (0.02)	0.039 (0.02)	6350 (1244)	257543 (557334)
100%	11.515 (0.63)	0.863 (0.02)	0.017 (0.01)	5925 (1911)	13209 (12550)
FFR	11.428 (1.39)	0.867 (0.04)	0.039 (0.05)	6583 (2954)	3370 (2743)
0%	10.745 (1.04)	0.872 (0.04)	0.014 (0.02)	4500 (2458)	43 (15)
p-value	0.384	0.981	0.466	0.598	0.750

3.4.2. Cope's giant salamanders

A. Comparison among regions

Significant differences existed in all genetic diversity measures among regions for Cope's giant salamanders (Table 12), with the exception of recent effective population size. These differences were all due to lower levels of all measures for basins on the Olympic Peninsula. Significant differences existed for all variables in pairwise comparisons between the Olympics and Willapa Hills (allelic richness *p*-value = 0.001, heterozygosity *p*-value = 0.001, F_{IS} *p*-value = 0.02, historic N_e *p*-value = 0.0004), but not for any comparisons between the Olympics and South Cascades or between the Willapa Hills and South Cascades. Interestingly, the Olympics had a lower level of inbreeding, despite having uniformly lower allelic richness, lower heterozygosity and lower effective population sizes. Generally, low levels of these three parameters would be associated with higher inbreeding coefficients. This discrepancy may be explained by evolved behavioral avoidance of inbreeding due to prolonged small population sizes, as mentioned previously. Nonetheless, it appears that the Olympic Peninsula cannot be considered a replicate for genetic diversity measures in the Type N Study, and highlights the importance of the paired sampling approach of the study to disentangle regional effects.

Table 12. Comparison of genetic diversity measures among regions for Cope's giant salamanders. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Region	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
Olympics	3.722 (0.46)	0.625 (0.08)	0.018 (0.06)	300 (122)	62 (65)
Willapa Hills	5.909 (0.60)	0.748 (0.05)	0.077 (0.04)	1915 (984)	219 (261)
South Cascades	5.204 (0.72)	0.732 (0.05)	0.079 (0.02)	1025 (671)	781 (1345)
p-value	*0.0018*	*0.0109*	*0.0429*	*0.012535*	0.315

B. Comparison among blocks

Results of among block comparisons demonstrated that allelic richness and historic effective population size were significantly different among blocks (Table 13). Pairwise comparison of the Olympic block and Willapa 1 block was significant (p = 0.002), as well as between Olympic block and Willapa 2 (p = 0.001). There were no significant pairwise comparisons after Bonferroni correction (adjusted *p*-value = 0.005)

for N_e, although Olympic v. Willapa 1 (*p*-value = 0.02) and Olympic v. Willapa 2 (*p*-value = 0.02) were significantly different if multiple comparisons were not accounted for. No differences existed between the South Cascades block and any of the Willapa Hills blocks.

Table 13. Comparison of genetic diversity measures among blocks for Cope's giant salamanders. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Block	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
Olympic	3.722 (0.46)	0.625 (0.08)	0.018 (0.06)	300 (122)	62 (65)
Willapa 1	5.801 (0.73)	0.741 (0.04)	0.053 (0.04)	2338 (1215)	79 (75)
Willapa 2	5.954 (0.70)	0.754 (0.08)	0.087 (0.04)	1475 (712)	136 (811)
Willapa 3	6.034 (0.28)	0.75 (0.01)	0.109 (0.01)	1950 (1131)	376 (518)
South Cascades	5.204 (0.72)	0.732 (0.05)	0.079 (0.02)	1025 (671)	781 (1345)
p-value	*0.0142*	0.1602	0.0699	*0.03516*	0.491

C. Comparison among anticipated treatments

As with tailed frogs, no significant differences were found in genetic diversity among anticipated treatment types (Table 14). Therefore, no pre-existing conditions appear to exist that might bias overall comparisons among pre- and post-treatment. Once again, allelic richness and heterozygosity have the lowest variance within treatments, although the variance is slightly higher than seen in tailed frogs. This could mean that treatment effects will be easier to detect in tailed frogs. In fact, simulations indicate that effective size would need to decrease to 12 for a detectable change in allelic richness and only two individuals to detect a change in heterozygosity. However, on average, effective sizes are smaller in giant salamanders, and thus a strong treatment effect may be more likely to be detected in giant salamanders as the proportion of population reduction would be smaller than in tailed frogs.

Table 14. Comparison of genetic diversity measures among treatments for Cope's giant salamanders. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Anticipated Treatment	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
reference	5.33 (1.15)	0.701 (0.07)	0.049 (0.04)	1800 (1295)	218 (298)
100%	5.129 (1.29)	0.695 (0.12)	0.099 (0.02)	1230 (1171)	657 (1205)
FFR	5.516 (1.13)	0.779 (0.08)	0.055 (0.04)	1300 (970)	204 (268)
0%	5.108 (0.99)	0.711 (0.05)	0.063 (0.06)	1025 (733)	94 (77)
p-value	0.9433	0.3162	0.3575	0.425461	0.944

3.4.3. Coastal giant salamanders

A. Comparison between regions

As no Coastal giant salamanders were detected in the Olympics in this study, our regional comparison was restricted to the Willapa Hills and South Cascades. The only significant difference between the two was in historic N_e, which, though low in both regions, was higher in the Willapa Hills (Table 15).

Table 15. Comparison of genetic diversity measures between regions for the Coastal giant salamander. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Region	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
Willapa Hills	3.57 (0.53)	0.604 (0.09)	0.097 (0.22)	456 (192)	128 (174)
South Cascades	3.444 (0.73)	0.602 (0.12)	0.056 (0.06)	438 (232)	624 (1092)
p-value	0.760	0.983	0.596	*0.022*	0.311

B. Comparison among blocks

No significant differences were found for any genetic diversity parameter for Coastal giant salamanders, despite some large differences in F_{IS} (Table 16). Thus, this

statistical technique may have low power to detect differences in the inbreeding coefficient; low power does not seem to be a problem for the other measures of genetic diversity across the three species. Lower power might be expected for this species because of the lower number of loci used. Overall, though, it appears that the blocks can be considered replicates for genetic diversity for Coastal giant salamanders.

Table 16. Comparison of genetic diversity measures among blocks for the Coastal giant salamander. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Block	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
Willapa 1	3.722 (0.3)	0.645 (0.03)	0.04 (0.05)	467 (104)	145 (176)
Willapa 2	3.283 (0.83)	0.637 (0.14)	-0.017 (0.31)	400 (304)	10 (7)
Willapa 3	3.773 (0.03)	0.54 (0.07)	0.198 (0.11)	525 (177)	280 (247)
South Cascades	3.444 (0.73)	0.602 (0.12)	0.056 (0.06)	438 (232)	624 (1092)
p-value	0.781	0.686	0.088	0.135	0.111

C. Comparison among anticipated treatments

We also found little significant difference by anticipated treatment, which is consistent with the finding for the other two species (Table 17). The only exception is recent effective population size, which is due to much larger population sizes in the 100% buffer treatment sites. Therefore, the process for choosing basins, and treatments assigned to those basins, was generally successful in creating an initial baseline without significant genetic differences among the experimental treatments for all taxa involved. However, inferences related to population size in Coastal giant salamanders should consider the baseline higher levels in the 100% treatment. Levels of variance are similar to Cope's giant salamander, except for lower variation in allelic richness. Due to the lower genetic diversity found in Coastal giant salamanders as compared to the other two species, a reduction to an effective size of six would be necessary to detect differences in allelic richness, and an effective size of a single individual would be required to detect changes in heterozygosity. Thus, an extreme bottleneck in giant salamanders would likely need to occur. This further supports the argument that genetic measures such as proportion of siblings detected and potentially direct estimates of effective population size may be more useful in detecting changes.

Table 17. Comparison of genetic diversity measures among anticipated treatments for the Coastal giant salamander. Abbreviations are as in Table 3. Values under each measure represent the average for that region, with parentheses indicating the standard deviation. P-value represents the result of the permutation comparison test (for A_r, H_o, and F_{IS}) or ANOVA (for N_e). Italics represent a significant value. Note that due to deviations from normality, tests for differences among ONESAMP N_e values were log-transformed for ANOVA analysis; averages of raw values are presented in the table.

Anticipated Treatment	A_r	H_o	F_{IS}	MIGRATE N_e	ONESAMP N_e
reference	3.519 (0.57)	0.589 (0.15)	0.036 (0.13)	438 (180)	40 (45)
100%	4.062 (0.33)	0.624 (0.09)	0.112 (0.1)	633 (126)	1021 (1075)
FFR	3.045 (0.62)	0.625 (0.05)	-0.021 (0.24)	283 (208)	54 (46)
0%	3.469 (0.3)	0.51 (0.05)	0.138 (0.03)	450 (141)	69 (72)
p-value	0.191	0.543	0.330	0.173	*0.033*

3.5. Objective 4: Genetic cluster size

3.5.1. Cluster site selection

Altogether, we sampled individuals from 38 streams to use as genetic cluster sites in the Olympics and South Cascades (Table 18). We did not detect both tailed frogs and giant salamanders at all sites, and overall we were more successful capturing tailed frogs than giant salamanders. We sampled tailed frogs from 32 sites (84%) and Cope's giant salamanders from 12 sites (32%). We also collected Coastal giant salamanders at 3 sites in small numbers. While we had relatively uniform success in sampling tailed frogs from both the Olympics and South Cascades, we had much greater sampling success in the Olympics for giant salamanders. This was primarily due to sampling effort; we had difficulty collecting giant salamanders in both regions, but spent much more time in the Olympics. This was because Steele et al. (2009) previously sampled 11 streams in a small region in the South Cascades that was located near basin 6000. They found high subdivision in Cope's giant salamander, as most individual streams clustered separately. In contrast, Coastal giant salamanders formed a single cluster over all 11 sites. Therefore, collecting giant salamanders from additional sites in the South Cascades was a lower priority than the Olympics because we already had a good idea of the genetic cluster size in the South Cascades from Steele et al. (2009).

Table 18. Sites used in genetic cluster analysis. Number is the site number in figures of clustering analysis. The species column represents which species were collected at each site (ASTR=tailed frog, DICO=Cope's giant salamander, DITE=Coastal giant salamander). UTM coordinates are in NAD83, Zone 10.

#	Site	Region	Species	UTME	UTMN
5	H1	Olympic	ASTR/DICO	443429	5254134
6	H11	Olympic	DICO	434932	5247868
7	H12	Olympic	DICO	443962	5255347
8	H3	Olympic	ASTR/DICO	438186	5246262
9	JB1	Olympic	DICO	418302	5297451
10	JB2	Olympic	DICO	469843	5247483
11	S1	Olympic	ASTR/DICO	461328	5249413
12	SAM	Olympic	DICO	430771	5273240
13	W1	Olympic	ASTR/DICO	449711	5246092
27	KC	South Cascades	DICO/DITE	558530	5063399
28	PC	South Cascades	ASTR/DICO/DITE	568972	5083368
29	Y1	South Cascades	ASTR/DICO/DITE	552106	5063310
31	BIG	Olympic	ASTR	450086	5250444
32	C1	Olympic	ASTR	462657	5244349
33	H2	Olympic	ASTR	434725	5243275
34	H5	Olympic	ASTR	441202	5247319
35	H6	Olympic	ASTR	442440	5250902
36	H7	Olympic	ASTR	445133	5248623
37	PETE	Olympic	ASTR	444970	5255907
38	RB	Olympic	ASTR	437876	5242411
39	SP	Olympic	ASTR	457945	5244876
40	T1	Olympic	ASTR	454565	5254620
41	W2	Olympic	ASTR	460196	5256203
42	WFH	Olympic	ASTR	448541	5258877
43	WY	Olympic	ASTR	460734	5260262
44	BR	South Cascades	ASTR	592624	5070774
45	CR	South Cascades	ASTR	575040	5077317
46	GF	South Cascades	ASTR	566425	5075823
47	GRS	South Cascades	ASTR	556202	5063905
48	HB	South Cascades	ASTR	591216	5080242
49	JC	South Cascades	ASTR	564180	5083242
50	L1	South Cascades	ASTR	584835	5110676
51	L3	South Cascades	ASTR	588866	5116446
52	M	South Cascades	ASTR	591480	5077681

#	Site	Region	Species	UTME	UTMN
53	PL	South Cascades	ASTR	566096	5082021
54	TL1	South Cascades	ASTR	605038	5101361
55	WC	South Cascades	ASTR	563891	5088589
56	WD	South Cascades	ASTR	583136	5091665

3.5.2. Coastal tailed frogs

Consistent with the patterns of null alleles, we found Coastal tailed frog sites to be strongly structured by region (Fig. 8). Values of ΔK clearly indicate that 3 clusters is the primary subdivision among sampled tailed frogs. The three clusters corresponded exactly to the three regions of the Olympics, Willapa Hills, and South Cascades (Fig. 8). There does appear to be some limited migration between the Willapa Hills and South Cascades, as evidenced by individuals assigned to one region being found in the other region. Further subdivision was only detected in the Willapa Hills (Fig. 9). Both the Olympics and South Cascades had all individuals assigned evenly to two clusters, strongly suggesting that on the basis of sampled alleles, no genetic structure exists in either region. The Willapa Hills has the greatest ΔK and log likelihood at $K = 3$, although visualizing the assignments revealed only two clear clusters (Fig. 9). Additionally, it appeared that basin 2468 (15) may have individuals belonging to the cluster represented by basin 2260 (14), as well as to the larger cluster containing the remaining Willapa sites. Therefore, we included 2468 (15) in a run with 2260 (14), and also included 2468 (15) in a separate run with the other Willapa sites. The results demonstrated that 2260 (14) and 2468 (15) clustered separately (Fig. 10), but with a number of migrants from 2468 (15) into 2260 (14) (but not the other direction; see relative locations in Fig. 3). There was no subdivision detected between 2468 (15) and all remaining Willapa sites. The AMOVA results supported the presence of only 4 clusters (Olympics, 2260, remaining Willapa Hills, and South Cascades) as only 0.43% of the variation was between sites within groups, whereas 4.8% of variation was between groups (the vast majority of variation was within sites, which is a common result in AMOVA analyses).

Tailed frogs have very high levels of gene flow even over broad geographic distances, indicating a very large genetic cluster size for a small amphibian at a regional scale, extending for 20-30 km (Spear and Storfer 2008). Indeed, it appears that genetic clusters are best defined by large tracts of open lowlands that form the divisions between the Olympics, Willapa Hills and South Cascades. The only apparent exception to this is basin 2260 (14), which is subdivided from the other Willapa Hills sites, and while receiving migrants from 2468 (15), does not seem to produce migrants to 2468 (15). These results are evident in the bar plot (Fig. 10) because 2260 (14) has several predominantly green bars, which indicate genetic assignment to basin 2468 (15). However, 2468 (15) has a very small proportion of individuals belonging to the red cluster suggesting unidirectional migration from 2468 (15) to 2260 (14). Therefore, basin

2260 (14) may represent a sink for tailed frogs (i.e., a site that relies on outside immigration to maintain a population; Pulliam 1988). Additional support for this comes from the fact that site 2260 (14) has the lowest N_e of all basins. The high migration of tailed frog across sites may complicate conclusions of the Type N Study because movement from individuals at other sites into treatment basins may obscure treatment effects. It is difficult to assess whether this connectivity is due to consistent high migration between sites every generation, or whether the high effective population size prevents rapid differentiation even if there is not constant migration between sites, or a combination of both. However, if statistically significant post-treatment differences in genetic diversity or genetic cluster size are detected for tailed frogs, it likely represents a very strong treatment effect that disrupted migration into the experimental basin or influenced the surrounding landscape. Thus, we would consider tailed frogs to serve as a conservative test of experimental buffer effects.

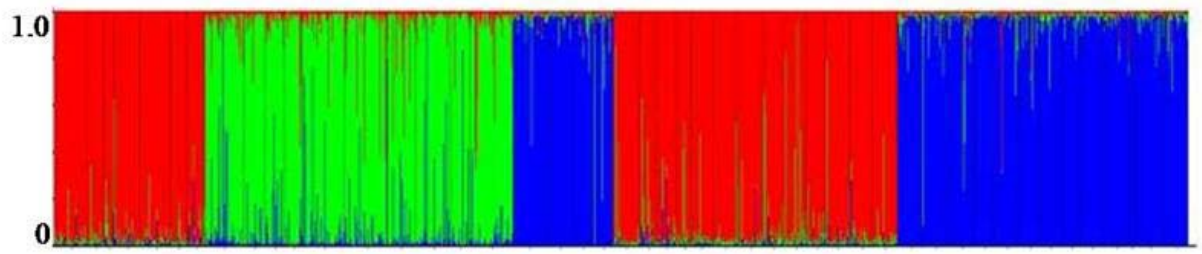


Figure 8. STRUCTURE bar plots for all tailed frog sites (suggesting *K*=3). Each color represents a population cluster. Red represents Olympic sites, green are Willapa Hills sites and blue are South Cascades sites. The y-axis represents the proportion of membership of each individual to each cluster. Each individual is represented by a different vertical line. That is, if an individual (represented by a vertical line) is primarily green, then its genotype indicates that it is most likely from the Willapa Hills. Similarly, if a line representing an individual's genotype is half green and half red, then it is equivocal whether that individual is from the Olympics or Willapa Hills.

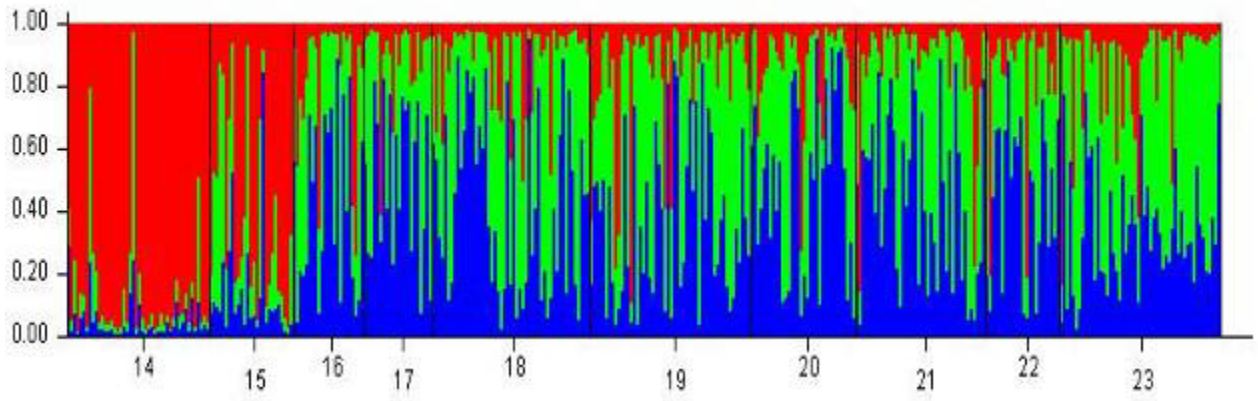


Figure 9. STRUCTURE bar plots for Willapa Hills tailed frog sites. Site numbers are as in Tables 1 and 18. This plot suggests two genetic clusters, sites 14 and 15 seem distinct from 16-23.

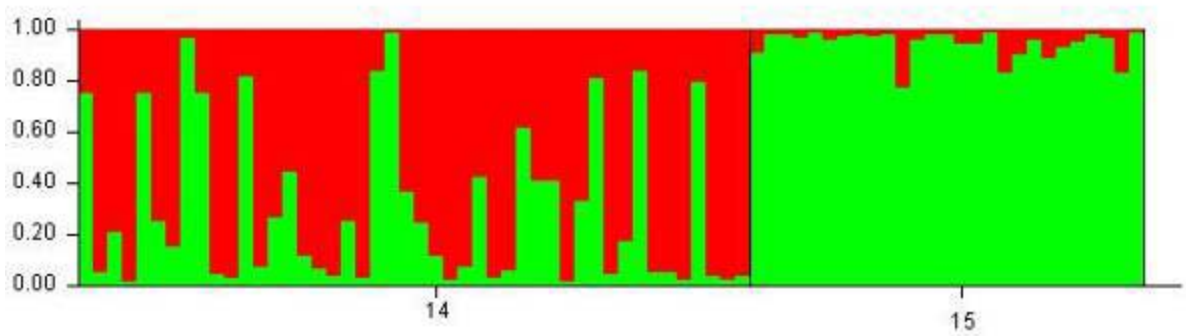


Figure 10. STRUCTURE bar plots for tailed frogs at basins 2260 (14) and 2468 (15). This plot suggests site 14 is distinct from 15, but there is a reasonable amount of gene flow from 15 into 14, but not vice versa (as indicated by the small amount of red for each individual in site 15).

3.5.3. Cope's giant salamanders

The first main population subdivision (which probably describes historic differentiation) in Cope's giant salamander sites was a split into two clusters, roughly corresponding to the Olympics and Willapa Hills/S. Cascades (Fig. 11). However, the two northern Willapa Hills basins, 2260 (14) and 2468 (15), cluster with the Olympic sites, and not with other Willapa Hills sites. This result differs from that for tailed frogs in two major ways. First, in tailed frogs, 2260 (14) and 2468 (15) cluster with the

Willapa Hills. Second, tailed frogs first partition into three major regions instead of the initial two clusters for Cope's giant salamander.

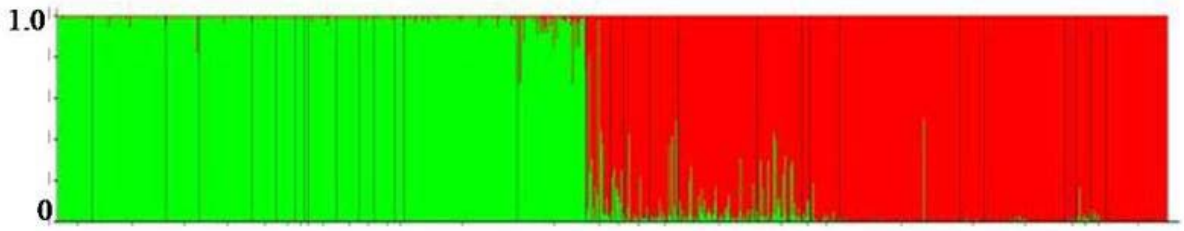


Figure 11. STRUCTURE bar plots for all Cope's giant salamander sites. Each color represents a population cluster. Green represents the Olympic sites as well as 2260 (14) and 2468 (15), whereas red is the remaining Willapa Hills sites and all South Cascade sites.

The AMOVA analysis based on the two groups suggested that there was further subdivision, as there was a greater percentage of variation among sites within groups (10%) than among groups (9%). Further subdividing the Olympic/2260/2468 cluster revealed an intuitive split, with the Olympic sites in one cluster and the 2260/2468 basins in the other (Fig. 12). While this division is expected, what is somewhat surprising is that there appears to be some migration from 2260/2468 into the Olympic sites, but less in the other direction. The other main cluster (Willapa Hills and South Cascades) breaks out into three different clusters (Fig. 13). One of these includes all the Willapa Hills (except 2260/2468) and the other two divide the South Cascades sites. One cluster includes basins 5378 (30) and 6000 (24), and the other cluster is comprised of 5595N (25), 5595S (26) and two of the cluster sites. One cluster site (28) appears to be highly admixed with all three clusters (Fig. 13). This leads to a total of five clusters; the AMOVA analysis did indicate more variation explained among groups (10%), but still that 6% of variation was due to differences between sites within groups, suggesting that further subdivision may be appropriate.

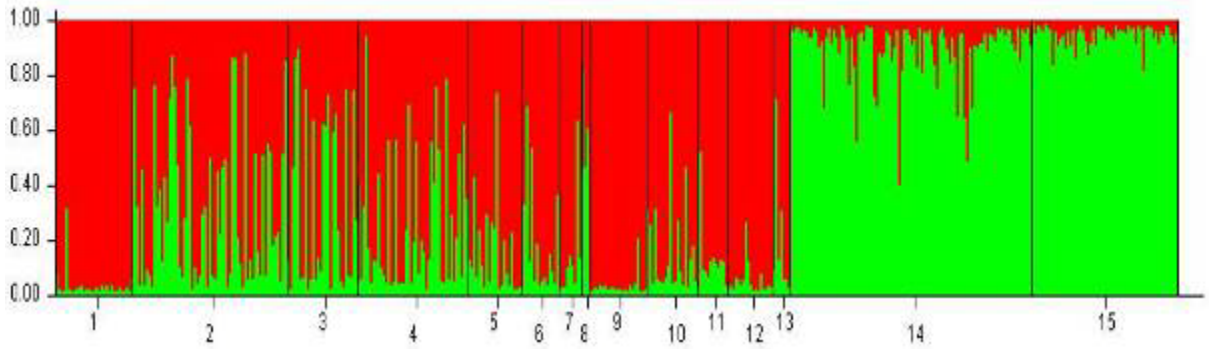


Figure 12. STRUCTURE bar plots for all Olympic Cope's giant salamander sites as well as basins 2260/2468. Site numbers are as in Tables 1 and 18.

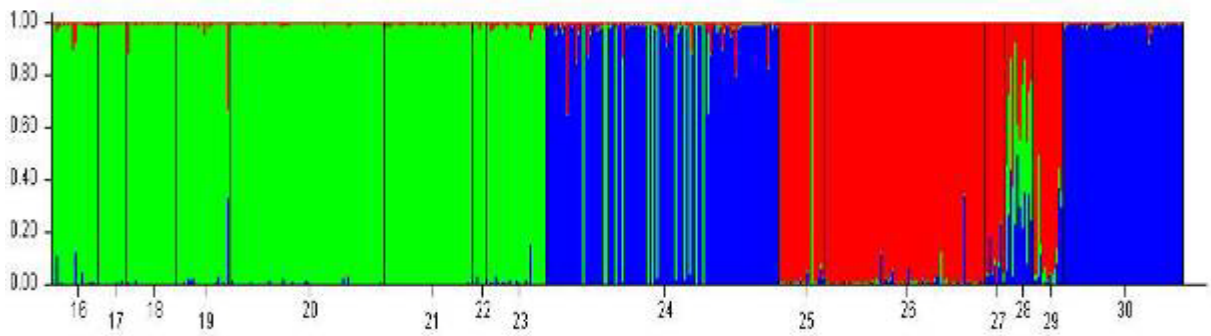


Figure 13. STRUCTURE bar plots for all Willapa Hills and South Cascades Cope's giant salamander sites. Site numbers are as in Tables 1 and 18.

A. Subdivision within the Olympic region

After 2260 (14) and 2468 (15) are split off from the Olympic sites, the first main subdivision within the Olympics is into two clusters: one cluster containing basin 363 (1) and cluster sites JB1 (9), JB2 (10), S1 (11) and SAM (12) and the other containing basins 1099 (2), 1197 (3), 1236 (4) and the remaining cluster sites (Fig. 14). The grouping of this split is interesting because 363 (1), JB1 (9) and SAM (12) form one geographical group in the northwest portion of the Olympic sampling area, whereas JB2 (10) and S1 (11) are a geographically disjunct group on the east side of the study area (Fig. 2). The fact that this disjunct group is genetically distinct from most of the central Olympic sites suggests that occasional dispersal might occur through the southern part of Olympic National Park to connect these two areas. However, the next level of subdivision divides 363/JB1/SAM from JB2/S1 (Fig. 15), so these two regions must only occasionally exchange migrants. Finally, the next round of hierarchical subdivision reveals that these two groups can be further split, so that 363 (1) is its own cluster and JB1 (9) and SAM

(12) are connected (Fig. 19), and JB2 (10) and S1 (11) are each split into unique clusters (Fig. 17). Therefore, the green cluster in Fig. 14 can eventually be almost completely reduced to each stream as a separate population, with the exception of JB1 (9) and SAM (12), which form one population despite a distance of 27 km between them.

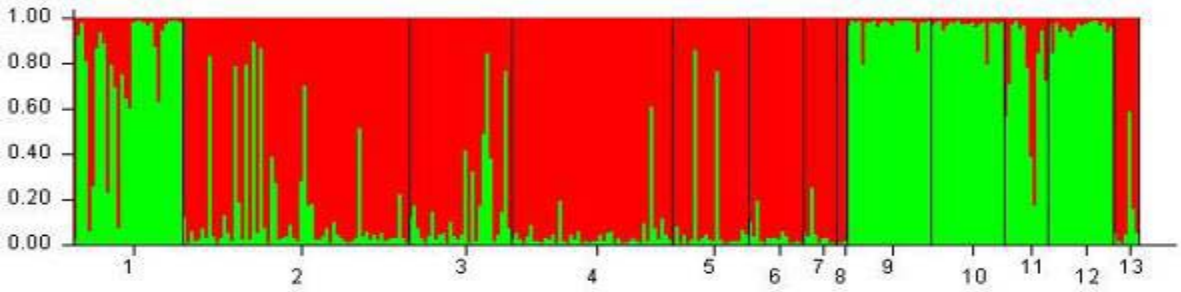


Figure 14. STRUCTURE bar plots for all Olympic Cope's giant salamander sites. Site numbers are as in Tables 1 and 18.

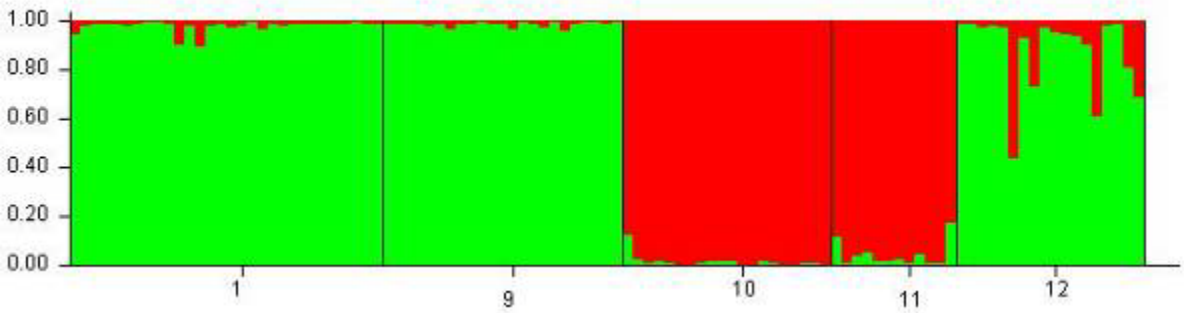


Figure 15. STRUCTURE bar plots for all Olympic sites in the green cluster (Fig. 17) for Cope's giant salamander. Site numbers are as in Tables 1 and 18.

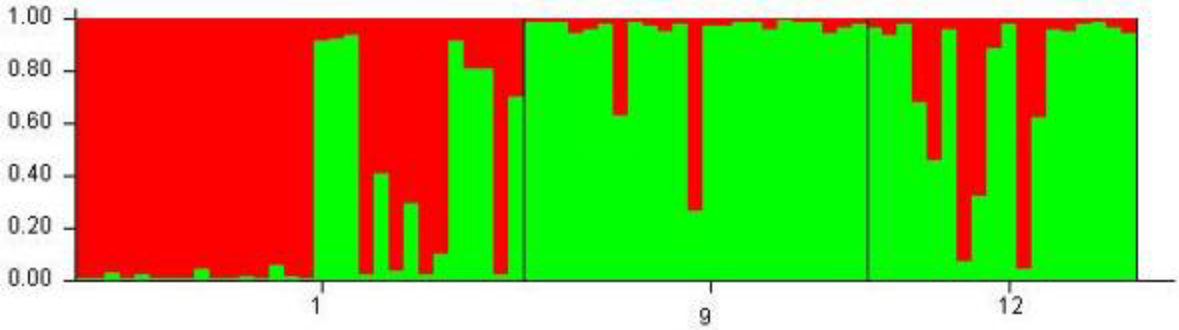


Figure 16. STRUCTURE bar plots for sites 363 (1), JB1 (9) and SAM (12) for Cope's giant salamander.

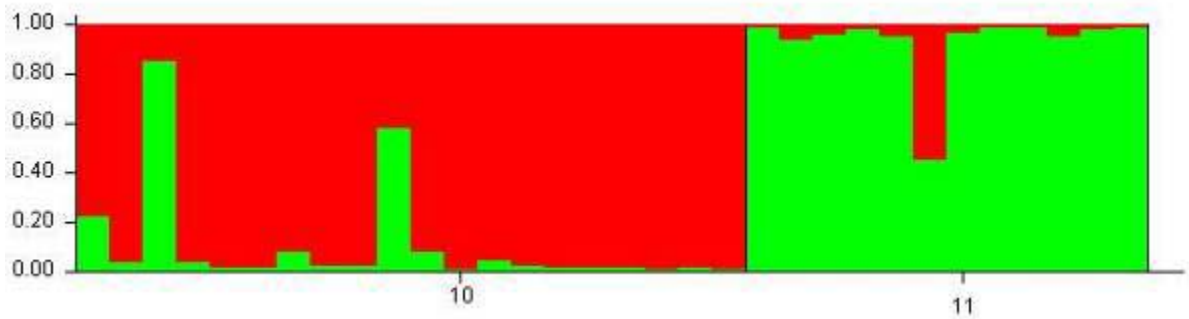


Figure 17. STRUCTURE bar plots for sites JB2 (10) and S1 (11) for Cope's giant salamander.

The other cluster of Olympic sites (the red cluster in Fig. 14) is similarly reduced into more structured populations. The first genetic break occurs into two clusters and results in one site splitting off (basin 1236 (4)), and basin 1197 (3) appears to be admixed between the two clusters (Fig. 18). Further analysis of the green cluster in Fig.18 revealed one last division in which 1099 (2), 1197 (3) and W1 (13) form one group, and the remaining Olympic cluster sites form the second group (Fig. 19). No further subdivision was found within this group of sites, and Fig. 19 indicates only weak genetic structure as mixing exists between the two groups. The maximum genetic cluster size observed in this portion is between 10-15 km. However, it is interesting that basin 1236 (4) is genetically separate from all other sites, despite the fact that it is very geographically close to 1099 (2) and 1197 (3). To summarize genetic patterns across the Olympic region, we detect a high degree of population structuring, as predicted by Steele et al. (2009), but we also observe occasional evidence of long-distance dispersal (Fig. 20). Two of the Type N Study basins (363 (1) and 1236 (4)) are isolated from all other sampled sites, whereas 1099 (2) and 1197 (3) are genetically connected, along with one other cluster site 13 km away.

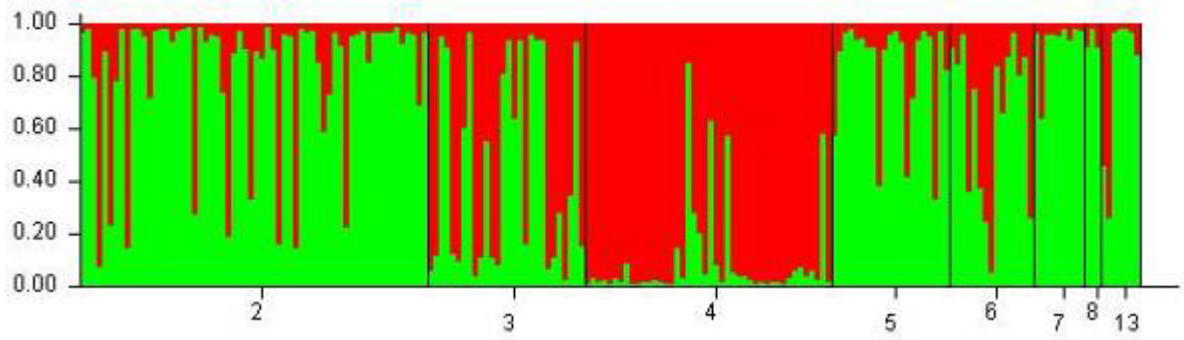


Figure 18. STRUCTURE bar plots further dividing the red cluster from Fig. 14 in the Olympic region for Cope's giant salamander. Site numbers are as in Tables 1 and 18.

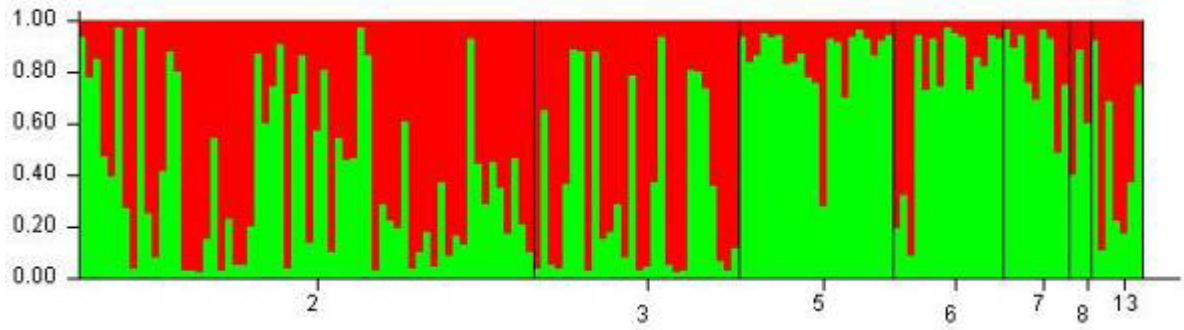


Figure 19. STRUCTURE bar plots further dividing the green cluster from Fig. 18 in the Olympic region for Cope's giant salamander. Site numbers are as in Tables 1 and 18.

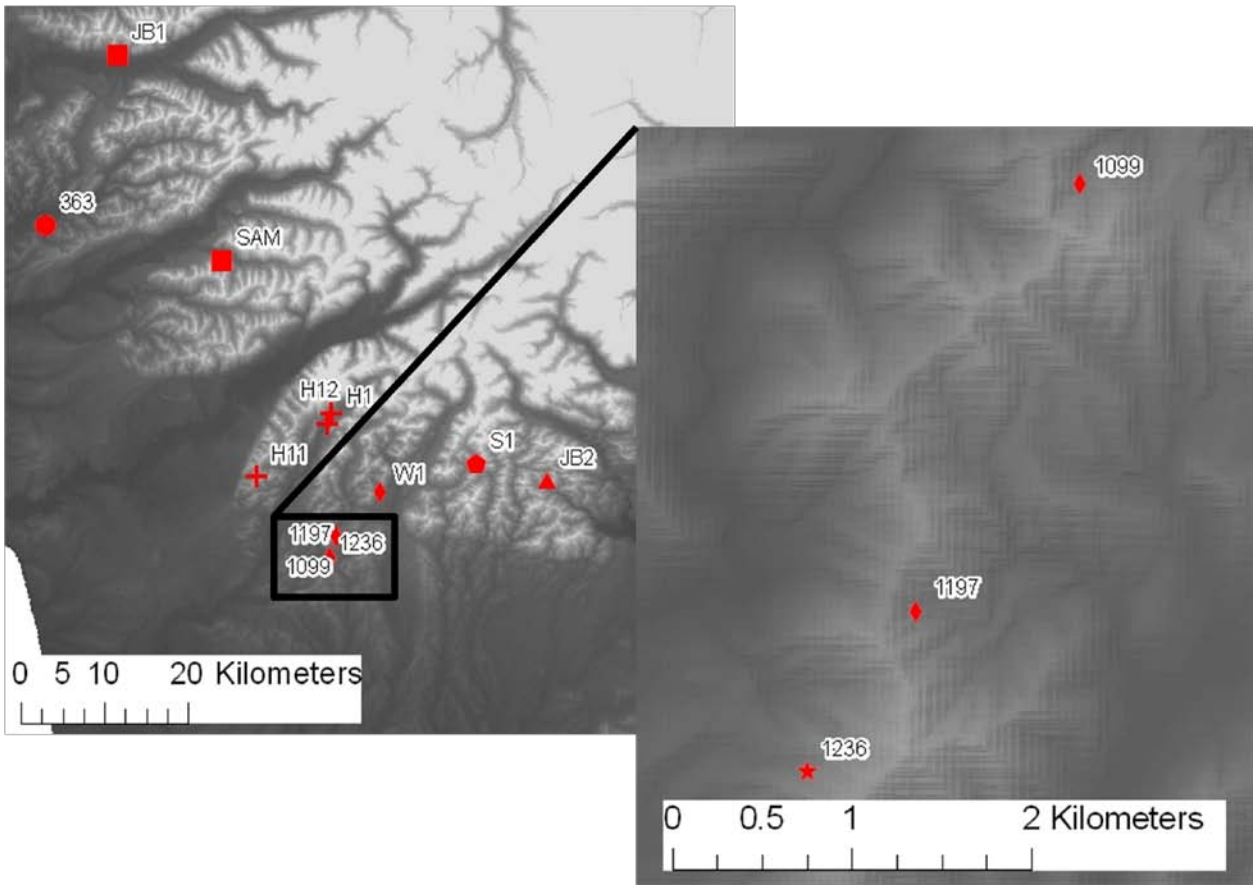


Figure 20. Map of *D. copei* genetic clusters in the Olympic region. Each shape represents a unique cluster.

B. Subdivision within the Willapa Hills region

We have already identified that basins 2260 (14) and 2468 (15) are separate from the other Willapa Hills sites (and are more closely related to the Olympics) (Fig. 12), and that the remaining Willapa Hills basins can be separated as a whole from the South Cascades (Fig. 13). We detected genetic differentiation between 2260 (14) and 2468 (15) when they were analyzed separately (Fig. 21) and similar to the pattern in tailed frogs, there appears to be more migration into 2260 (14) from 2468 (15) than vice versa. The remaining Willapa Hill sites break into three clusters that align well with geography (Fig. 3; Fig. 22). Basins 3074 (16), 3098 (17), 3110 (18) and 3111 (19) all cluster together as green, 3437 (20) and 3576 (21) form a group as red, and the third cluster is 3914 (22) and 5785 (23) in blue. The pair of 3914 (22) and 5785 (23) is most distinct, while there is a fair degree of genetic mixing between the other two groups (Fig. 22). The only further divisions identified within these subgroups are the differentiation of many individuals in basin 3110 (18) from the remaining group (Fig. 23) and the division of 3437 (20) from 3576 (21) (Fig. 24). However gene flow occurs from 3576 (21) into 3437 (20), but very

little evidence of flow exists in the opposite direction (Fig. 24). Therefore 3437 (20) may be located in a sink area for Cope's giant salamanders, and this was the only basin to show evidence of a significant heterozygosity excess in the bottleneck tests. Overall the genetic cluster size of the Willapa Hills is quite restricted, with a maximum detected cluster extent of only 2-6 kilometers (Fig. 25).

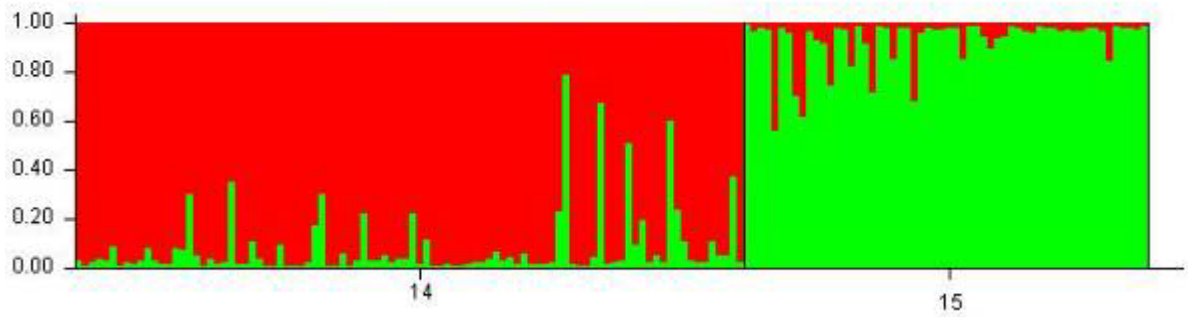


Figure 21. STRUCTURE bar plots demonstrating the division of site 2260 (14) and 2468 (15) for Cope's giant salamander.

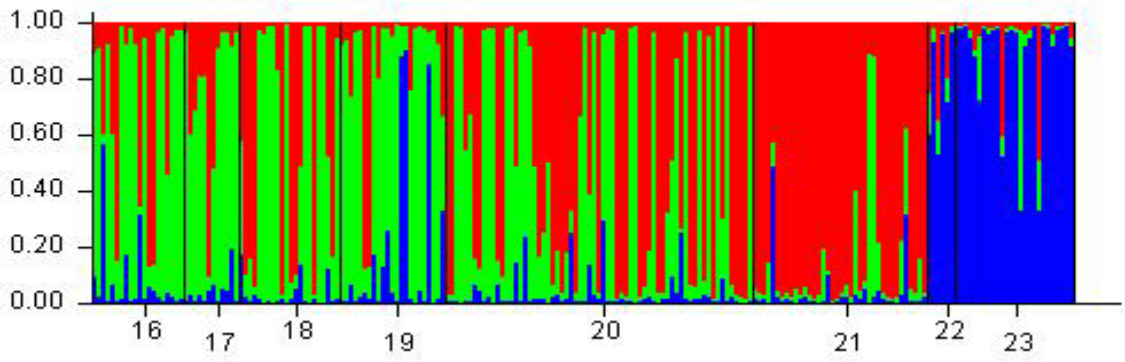


Figure 22. STRUCTURE bar plots for Willapa Hill sites for Cope's giant salamander. Site numbers are as in Tables 1 and 18.

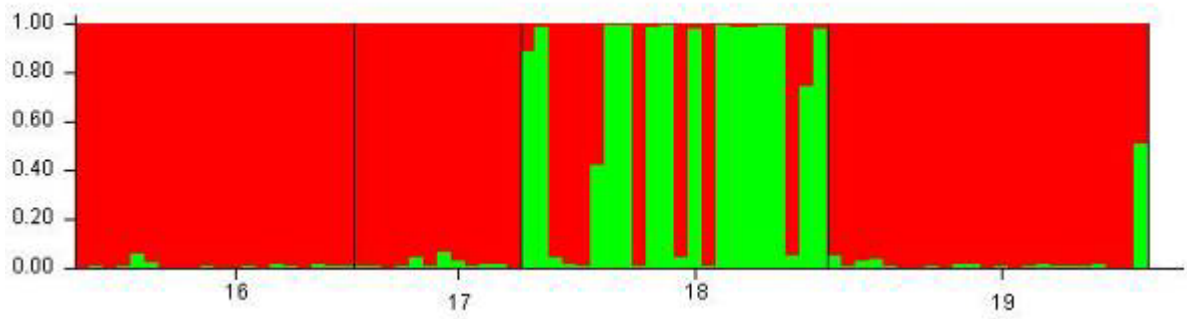


Figure 23. STRUCTURE bar plots for 3074 (16), 3098 (17), 3110 (18) and 3111 (19) for Cope's giant salamander.

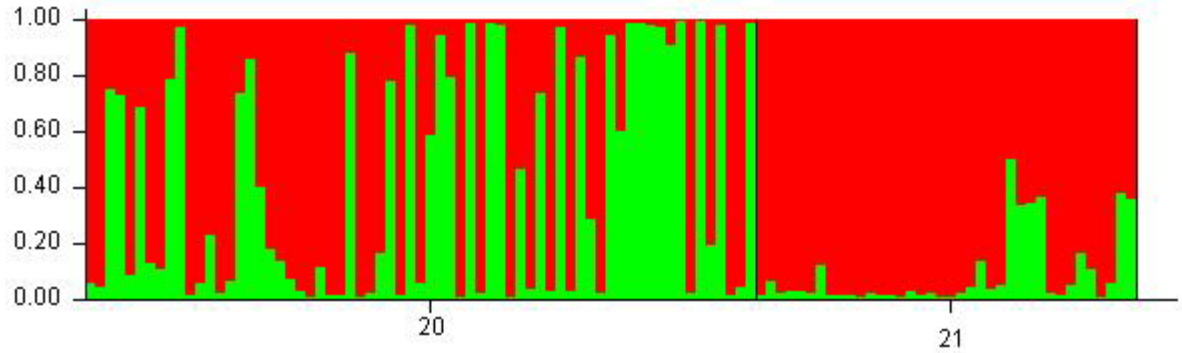


Figure 24. STRUCTURE bar plots for 3437 (20) and 3576 (21) for Cope's giant salamander.

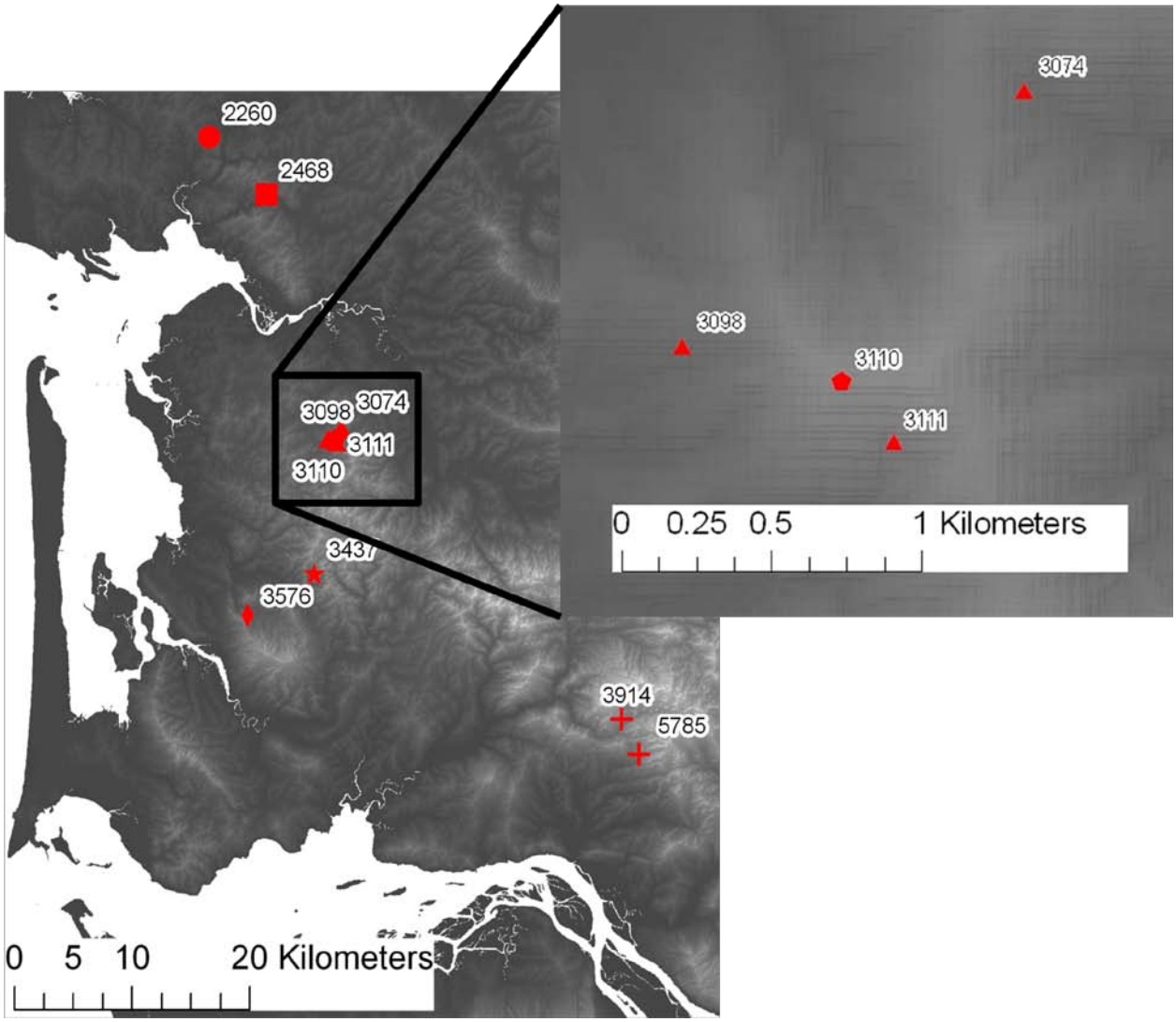


Figure 25. Map of *D. copei* genetic clusters in the Willapa Hills region. Each shape represents a unique cluster.

C. Subdivision within the South Cascades region

We have already identified two genetic clusters in the South Cascades in Fig. 13. The two basins comprising the first cluster, 5378 (30) and 6000 (24), are clearly differentiated from each other (Fig. 26). The second South Cascades cluster in Fig. 13 included basins 5595N (25), 5595S (26) and the three cluster sites. When this cluster is analyzed alone, two of the cluster sites (PC (28) and Y1 (29)) split off, while there is connectivity among 5595N (25), 5595S (26) and the cluster site KC (27) (Fig 27), which

is 5 km north of the 5595 basins. No further subdivision was found among 5595N (25), 5595S (26) and KC (27), but PC (28) and Y1 (29) are genetically differentiated (Fig. 28). Thus, our findings with these sampled sites agree well with the results of Steele et al. (2009) that show very low gene flow among Cope's giant salamander in the South Cascades (Fig. 29).

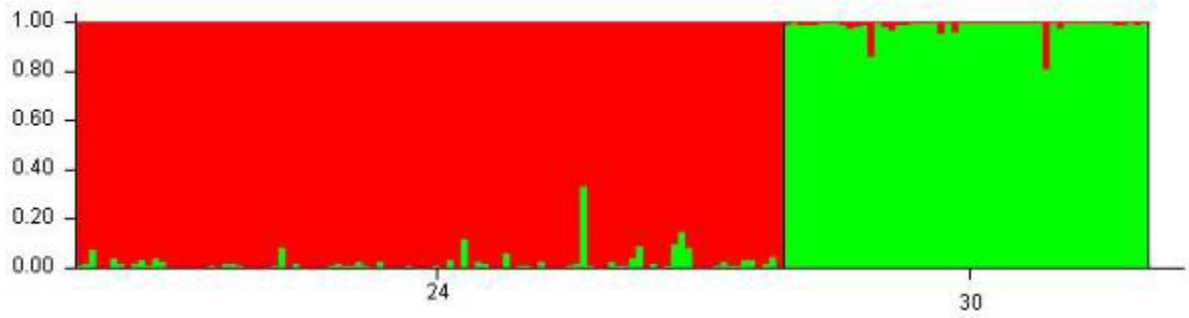


Figure 26. STRUCTURE bar plots for 6000 (24) and 5378 (30) for Cope's giant salamanders in the South Cascades.

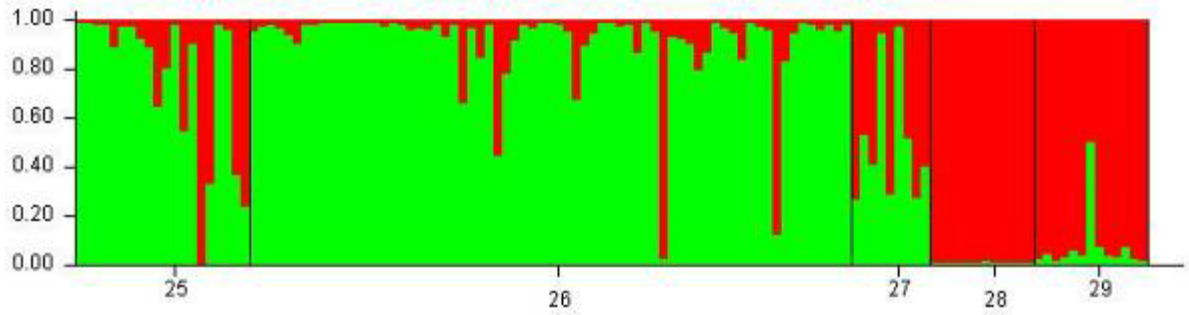


Figure 27. STRUCTURE bar plots for 5595N (25) and 5595S (26), as well as three cluster sites for Cope's giant salamanders in the South Cascades.

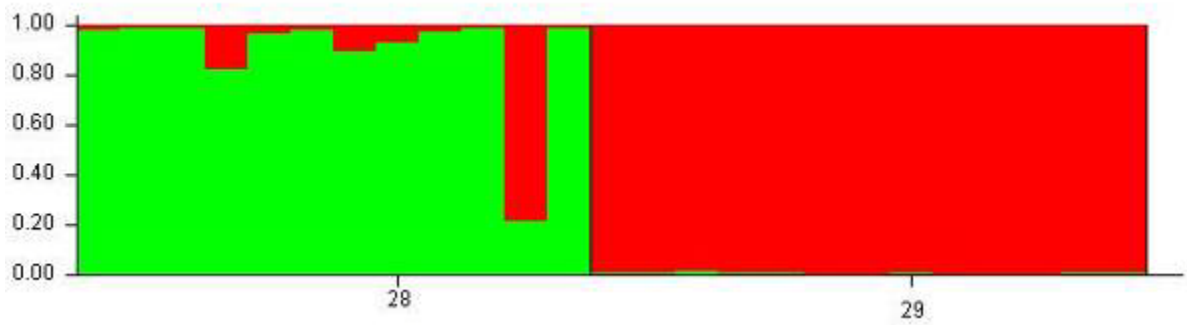


Figure 28. STRUCTURE bar plots for two cluster sites, PC (28) and Y1 (29), for Cope's giant salamanders in the South Cascades.

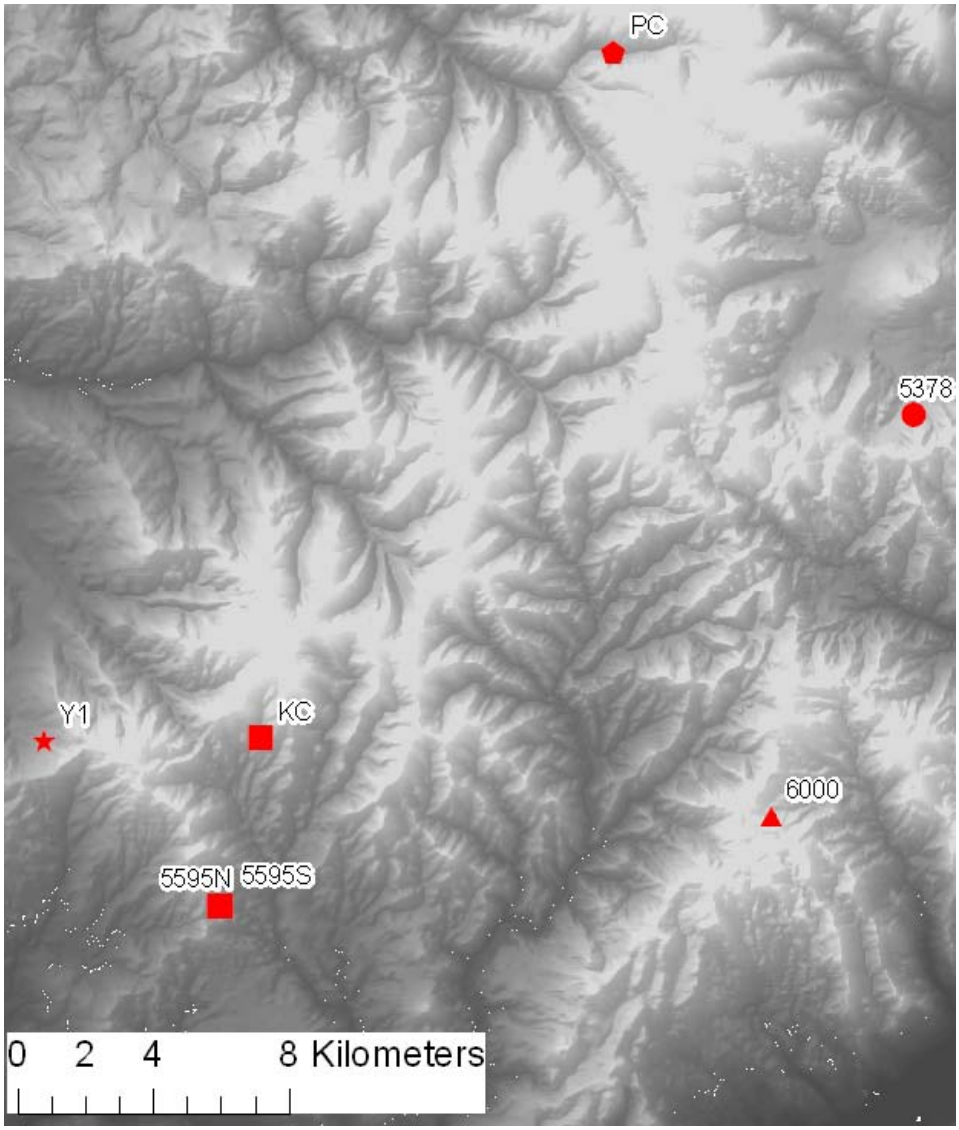


Figure 29. Map of *D. copei* genetic clusters in the South Cascades region. Each shape represents a unique cluster. 5595 N and S are both represented by the square cluster.

The AMOVA supports the full hierarchical clustering across all three regions, as described above, which leads to 18 different genetic groups for Cope's giant salamander. This highly partitioned genetic grouping has only 1.9% of variation due to among sites within groups, versus 13% of variation among groups. Across all regions, a gradient of increased genetic connectivity in Cope's giant salamanders appears to exist as one moves from the South Cascades to the Olympics. The genetic cluster size in the South Cascades is 5 km at most, increases to around 10-15 km in the Willapa Hills, and finally reaches a maximum close to 30 km in the Olympics. This broadly correlates with the precipitation gradient that increases to the Olympics, which may allow periodic overland dispersal, but

explaining this pattern remains a topic for further study. However, even in the Olympics, some proximate sites are differentiated, so a constant genetic cluster size cannot be assumed across an entire region. One consequence of the increased differentiation seen in Cope's giant salamanders is that migration is less likely to influence Type N Study results. Therefore, we predict that if treatments do have an effect, it will likely be most easily detected for Cope's giant salamanders.

3.5.4. Coastal giant salamanders

In contrast to tailed frogs and Cope's giant salamanders, the Coastal giant salamander does not have complete separation between the Willapa Hills and South Cascades (Fig. 30). In fact, basin 3437 (20) grouped with South Cascades sites and South Cascade genetic cluster sites KC (27) and PC (28) grouped with Willapa Hills. Notably, all three of these sites had a low sample size, which may have biased the analysis. Analysis of the predominantly Willapa Hills cluster produced two additional groups, of which one of the clusters contained 3111 (19) and 3576 (21), counter to geographical proximity (Fig. 31). We removed 3111 (19) and 3576 (21) and found two more clusters with a high degree of admixture that grouped 3074 (16), 3098 (17), 3110 (18) in one cluster and 3914 (22), 5785 (23) and the two South Cascades cluster sites in the other cluster (Fig. 32). No further subdivision was detected in the Willapa Hills group. A highly uneven genetic cluster size exists in the Willapa Hills (Fig. 33). Adjacent basins 3110 (18) and 3111 (19) were divergent, but 3111 (19) and 3576 (21) were genetically continuous despite being 10 km apart, and even more surprising, 3914 (22) and 5785 (23) are more than 100 km away from the South Cascades cluster sites. Across the South Cascades, no evidence was found for genetic differentiation by site (Fig. 34) and basin 3437 (20) continued to be included with South Cascades sites. This is also consistent with the findings of Steele et al. (2009), as they found no genetic structure for Coastal giant salamanders. Based on the four clusters indicated by the clustering analysis, the AMOVA that the same percentage of variation (3%) was explained by both among groups and among sites within groups. In concert, with the nonintuitive geographical clusters, this is consistent with the hypothesis that recent colonization by the Coastal giant salamander from a common refugium has not allowed enough time for genetic structure to become evident.

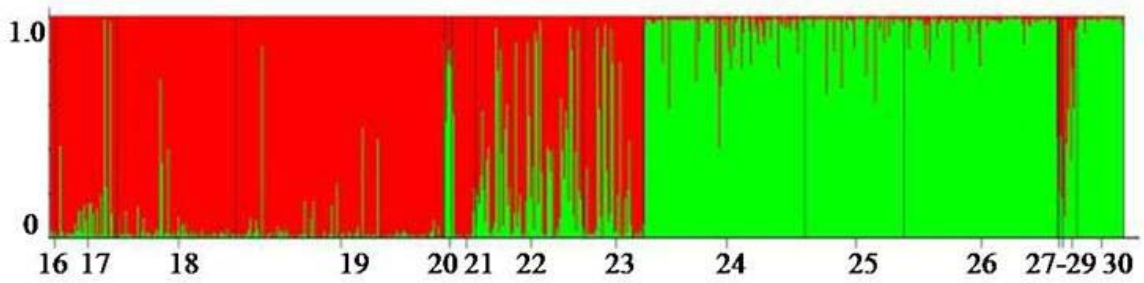


Figure 30. STRUCTURE bar plots for all Coastal giant salamander sites. Site numbers are as in Tables 1 and 18.

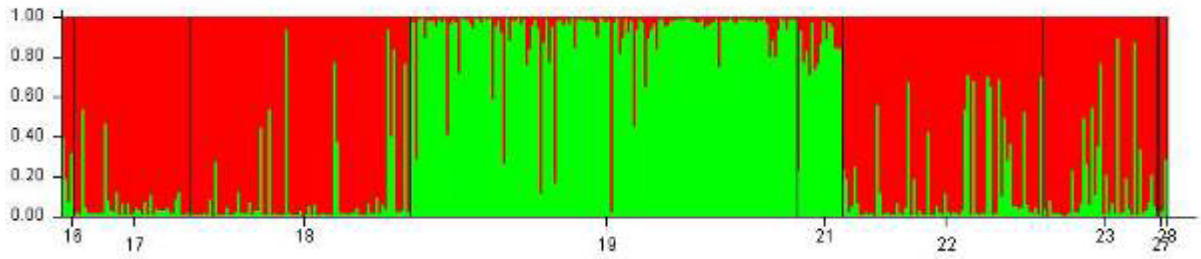


Figure 31. STRUCTURE bar plots for Coastal giant salamander sites in the predominantly Willapa Hills cluster. Site numbers are as in Tables 1 and 18.

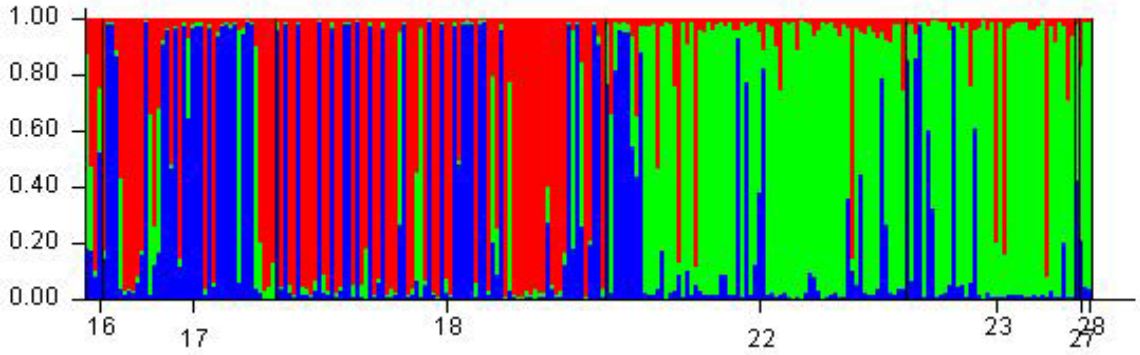


Figure 32. STRUCTURE bar plots for the Coastal giant salamander red cluster depicted in Fig. 32. Site numbers are as in Tables 1 and 18.

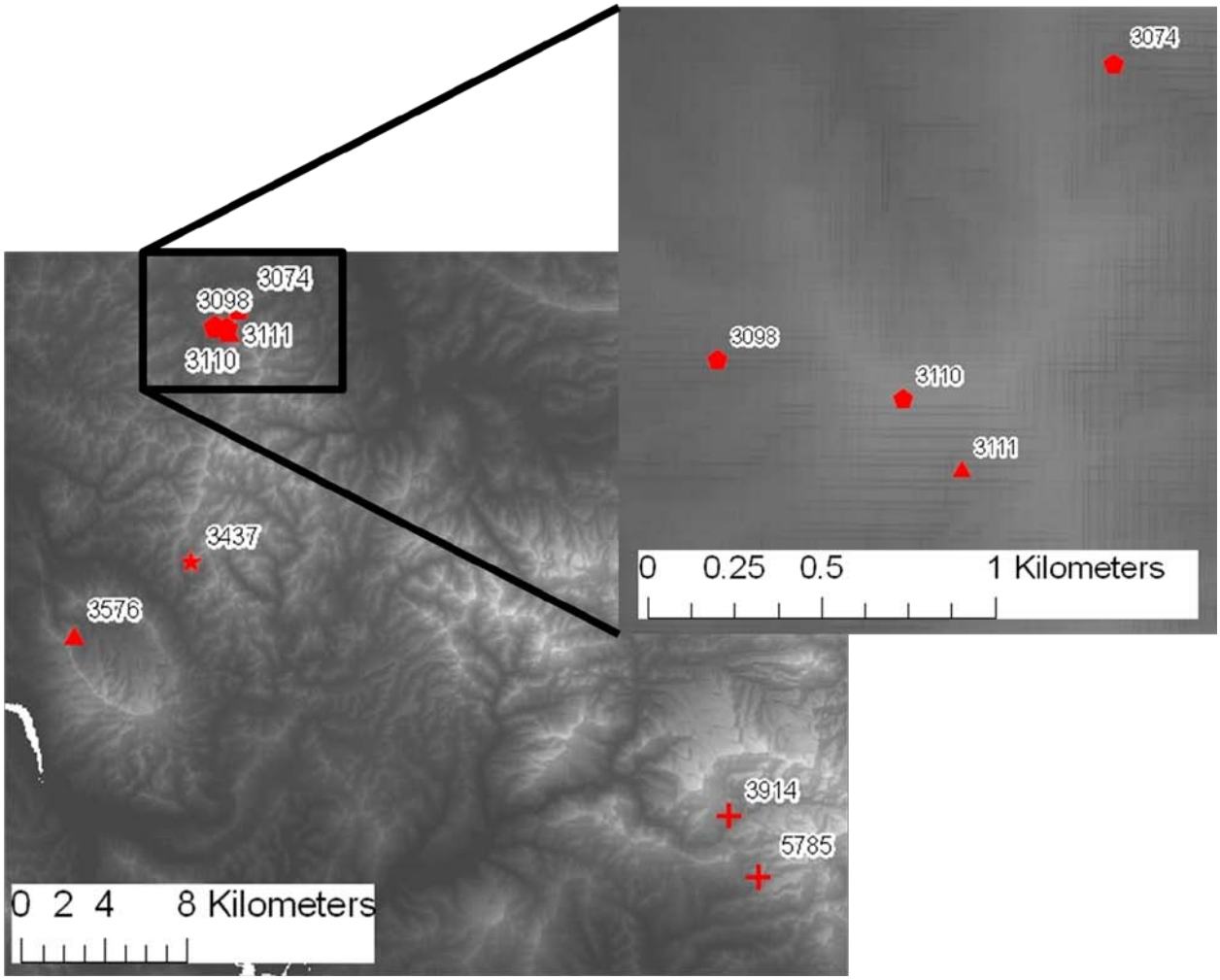


Figure 33. Map of *D. tenebrosus* genetic clusters in the Willapa Hills region. Each shape represents a unique cluster.

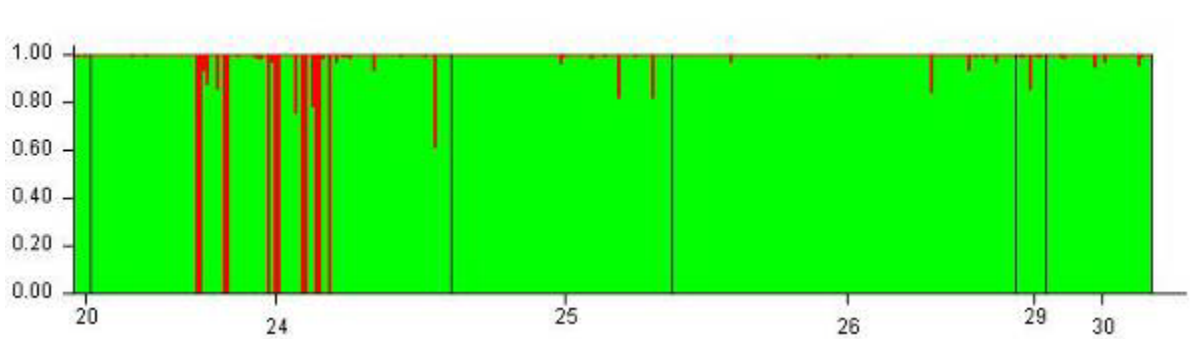


Figure 34. STRUCTURE bar plots for the predominantly South Cascades cluster. Site numbers are as in Tables 1 and 18.

4. Final Conclusions and Synthesis

4.1. Coastal tailed frogs synthesis

Collectively, Coastal tailed frogs are a genetically diverse group of populations, among which we found limited evidence for recent declines. Tailed frogs have among the highest genetic diversity seen in frogs, relatively large effective population sizes and very high rates of gene flow. These characteristics suggest that tailed frogs have been stable for a long time, are resistant to the negative effects of drift and should be buffered from large declines. However, large patches of open habitat are barriers to population connectivity of tailed frogs, as evidenced at a broad scale by the observed regional differentiation. Another study supports this result, as populations in Olympic National Park were separated by alpine meadows (Spear and Storfer 2008). Second, evidence of recent declines exists in the Olympics and two sites in the Willapa Hills and pairwise genetic distance values were negatively correlated with timber harvest in the Olympics (Spear and Storfer 2008), suggesting possible genetic effects of timber management. Whether forest management will have serious long-term effects on tailed frog persistence and genetic diversity is unclear. Results of the Type N Study will be extremely useful in this regard. If detectable changes resulting from a particular buffer treatment occur, then effects are likely to be quite strong due to such high baseline effective population sizes and connectivity in this species. Finally, our genetic analyses suggest that despite a large larval sample, abundance measures are probably representative of population size for this species. We detected few siblings, so most individuals belonged to unique families.

4.2. Cope's giant salamanders

Cope's giant salamander populations are characterized by levels of heterozygosity intermediate between tailed frogs and Coastal giant salamanders (but high in comparison with other salamander studies), large effective population sizes, and little evidence of recent declines, but very restricted genetic connectivity. Significant differences existed in genetic diversity by region, and this is likely due to effects of genetic drift resulting from isolation. However, we observed a consistent pattern of lower diversity and population size on the Olympic Peninsula, which is counterintuitive as there is a larger genetic cluster size in the Olympics, as well as a lack of Coastal giant salamanders as potential competitors. In total, we predict that Cope's giant salamanders are the most likely of the three study species to show a response to harvest treatments. Although our simulations indicated that a larger reduction in absolute numbers of individuals would be required as compared to tailed frogs, it is a lower proportion of average current effective population size than for tailed frogs. Most individuals will be restricted to their home basin and thus are less likely to evade any negative effects. Furthermore, "genetic rescue" from individuals emigrating from other streams is less likely due to low overall genetic cluster sizes and neoteny in most populations. Lastly, genetic data are especially important for estimating Cope's giant salamander population response in this study

because there were a number of full siblings detected. As a result, abundance counts will likely not accurately estimate adult population sizes.

4.3. Coastal giant salamanders

Genetic results for Coastal giant salamanders were perhaps the most surprising. Coastal giant salamanders have been characterized as highly abundant and the most resilient to disturbance of stream-associated amphibians (Welsh 2005). However, we detected low levels of genetic diversity and small effective population sizes, as well as widespread evidence for declines. High gene flow was found, as expected, but some of the geographical patterns of genetic clustering confounded effective interpretation based on available information. Simple abundance surveys for Coastal giant salamanders are also likely to be unreliable because of the great number of siblings detected, at least for our sample basins. Low diversity and population sizes may simply represent the residual effect of contraction resulting from Pleistocene glaciations and recent re-expansion from refugia across Washington landscapes. If this is true, it suggests that populations have not greatly increased in size since expanding across their range, as bottleneck effects are still observable. Therefore, in one aspect, this species could be considered highly vulnerable to habitat disturbance due to low effective population sizes. Nonetheless, the high dispersal capacity of this species might also keep the species from becoming extirpated locally in any streams. Clearly further studies are needed to disentangle these possibilities.

4.4. Final conclusions

An important aspect of all the analyses for the Type N Study is that the three stream amphibian species are not interchangeable: all have different genetic characteristics. Therefore, to truly understand how forest management affects the entire stream amphibian community, each species needs to be examined individually. This is clearly one of the strengths of the Type N study, as it allows for the inclusion of multiple species. Second, the baseline analyses presented here are absolutely necessary to obtain meaningful conclusions from post-treatment data. Genetic diversity was unequal across blocks (especially for Cope's giant salamanders), providing strong support for the importance of blocking; thus, baseline data for each block must be accounted for when analyzing post-treatment effects. Fortunately, however, no consistent differences were detected among treatment types for any species, so treatment assignments should not lead to any significant bias in the results. However, there would need to be strong treatment effects for genetic parameters to significantly vary in only one generation. Despite this, we would like to reiterate the importance of including genetic data after harvest for a thorough evaluation of post-treatment effects for each species. Tailed frog census surveys may be an accurate assessment of the breeding population, but without knowledge of the genetic connectivity of the focal basin with other sites, sampled individuals may simply represent immigrants from nearby sites not exposed to the treatment, although it is unclear the average number of immigrants per generation into a

site. On the other hand, abundance counts do not seem wholly accurate for both species of giant salamanders due to high numbers of sampled full siblings and a high degree of difficulty in identifying specimens to species in the field. Thus, genetic data are needed to determine effective population sizes, as well as for species identification, respectively.

5. Acknowledgments

This work was funded by and prepared for the Cooperative Monitoring, Evaluation and Research Committee under the Washington Department of Natural Resources Forest Practices Adaptive Management Program IAA 04-047 to WSU. We would like to thank Aimee McIntyre, Marc Hayes and Eric Lund for help with logistics in the field and advice on site selection, as well as insightful discussions regarding stream amphibians. Craig Steele provided much advice and assistance regarding giant salamander analyses. Sean Anderson, April Barreca, Karen Chojnacki, Nick Chojnacki, Jen Cotter, Tierra Curry, Jennifer Dhundale, Cristina Dressel, Robert Dyer, Jon Eastman, Andy Giordano, Marc Hayes, Tiffany Hicks, Denim Jochimsen, Scott Jones, Eric Lund, Doré Mangan, Jeffrey Marsten, Aimee McIntyre, Teresa Miskovic, Casey Richart, Maureen Thompson, Charissa Waters, Frithiof Waterstat, Nicholas Wenzel, Anna Yost, Kevin Young and Kayla Zaret all assisted with sample collection. Jon Eastman and Melanie Murphy assisted with laboratory work. Marc Hayes, Aimee McIntyre, Ivan Phillipsen, Teresa Miskovic, Chris Mendoza, Adrian Spidle, and Terry Jackson provided suggestions on the draft report.

6. Literature Cited

Anderson, E.C., and E.A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160:1217–1229.

Balloux, F. 2001. A computer program for the simulation of population genetics. *Journal of Heredity* 92:301-302.

Beebee, T.J.C. 2005. Conservation genetics of amphibians. *Heredity* 95:423-427.

Beerli, P. 2008. MIGRATE Version 3.0.3: Documentation and program, available at: http://popgen.csit.fsu.edu/Migrate-n.html

Bury, R.B., P.S. Corn, K.B. Aubry, F.F. Gilbert, and L.L.C. Jones. 1991. Aquatic amphibian communities in Oregon. Pp. 353-362 *In*: L.F. Ruggiero, K.B. Aubry, and M.H. Huff (technical coordinators), *Wildlife and vegetation of unmanaged Douglas-fir forests*. US Department of Agriculture, Forest Service, General Technical Report, PNW-GTR-285.

Busch, J.D., P.M. Waser, and J.A. DeWoody. 2007. Recent demographic bottlenecks are not accompanied by a genetic signature in banner-tailed kangaroo rats (*Dipodomys spectabilis*). *Molecular Ecology* 16:2450-2462.

Chandler, C.H., and K.R. Zamudio. 2008. Reproductive success by large, closely related males facilitated by sperm storage in an aggregate breeding amphibian. *Molecular Ecology* 17:1564-1576.

Chapuis, M.P., and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621-631.

Corn, P.S., and R.B. Bury. 1989. Logging in western Oregon: Responses of headwater habitats and stream amphibians. *Forest Ecology and Management* 29: 39-57.

Cornuet, J.M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001-2014.

Curtis, J.M., and E.B. Taylor. 2003. The genetic structure of coastal Pacific giant salamanders (*Dicamptodon tenebrosus*) in a managed forest. *Biological Conservation* 115:45-54.

Diller, L.V., and R.L. Wallace. 1999. Distribution and habitat of *Ascaphus truei* in streams on managed, young growth forests in north coastal California. *Journal of Herpetology* 33: 71-79.

Eastman, J.M., T.A. Spradling, and J.W. Demastes. 2007. Conservation assessment of the blue-spotted salamander. *American Midland Naturalist* 158:233-239.

Evanno G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-2620.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.

Excoffier, L., P.E. Smouse, and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction sites. *Genetics* 131: 479-491.

Ficetola, G.F., T.W.J. Garner, and F. de Bernardi. 2007. Genetic diversity, but not hatching success, is jointly affected by postglacial colonization and isolation in the threatened frog, *Rana latastei*. *Molecular Ecology* 16:1787-1797.

Garza, J.C., and E.G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* 10:305-318.

Giordano, A.R., B.J. Ridenhour, and A. Storfer. 2007. The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactylum*). *Molecular Ecology* 16:1625-1637.

Glaubitz, J.C. 2004. CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes* 4:309-310.

Goldberg, C.S., M.E. Kaplan, and C.R. Schwable. 2003. From the frog's mouth: Buccal swabs for collection of DNA from amphibians. *Herpetological Review* 34:220-221**.**

Goldberg, C.S. and L.P. Waits. 2009. Quantification and reduction of bias from sampling larvae to infer population and landscape genetic structure. *Molecular Ecology Resources*. doi: 10.1111/j.1755-0998.2009.02755.x

Good, D.A. 1989. Hybridization and cryptic species in *Dicamptodon* (Caudata: Dicamptodontidae). *Evolution* 43:728-744.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Updated from Goudet (1995). Available from http://www.unil.ch/izea/softwares/fstat.html. (accessed June 2009)

Jones, L.L.C., and P.S. Corn. 1989. Third specimen of a metamorphosed Cope's Giant Salamander (*Dicamptodon copei*). *Northwestern Naturalist* 70(2):37-38.

Jones, L.L.C., and R.B. Bury. 2005. Cope's giant salamander. Pp. 46-49. *In*: L.L.C. Jones, W.P. Leonard, and D.H. Olson (editors), *Amphibians of the Pacific Northwest.* Seattle Audubon Society, Seattle.

Kroll, A.J., J.G. Maccracken, T.C. Mebride, J. Bakke, J. Light, P. Peterson, and J. Bach. 2010. Basin-scale surveys of stream-associated amphibians in intensively managed forests. *Journal of Wildlife Management* 74:1580-1587.

Lewis, P.O., and D. Zaykin. 2001. Genetic Data Analysis: Computer Program for the Analysis of Allelic Data (2000) Version 1.0. http://hydrodictyon.eeb.uconn.edu/people/plewis/research.php.

Luikart, G., F.W. Allendorf, J.-M. Cornuet, and W.B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89:238-247.

McIntyre, A.P., M.P. Hayes, and T. Quinn. 2009. Type N Feasibility Study. A report submitted to the Landscape and Wildlife Advisory Group, Amphibian Research Consortium, and the Cooperative Monitoring, Evaluation, and Research Committee.

Nielson, M., K. Lohman, C.H. Daugherty, F.W. Allendorf, K.L. Knudsen, and J. Sullivan. 2006. Allozyme and mitochondrial DNA variation in the tailed frog (Anura: *Ascaphus*): The influence of geography and gene flow. *Herpetologica* 62:235-258.

Nussbaum, R.A. 1976. Geographic variation and systematic of salamanders of the genus *Dicamptodon* Strauch (Ambystomatidae). *Miscellaneous Publications of the Museum of Zoology, University of Michigan* 149:1-94.

Piry, S., G. Luikart, and J.-M. Cournet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* 90:502-503.

Pope, L.C., A. Estoup, and C. Moritz. 2000. Phylogeography and population structure of an ecotonal marsupial, *Bettongia tropica*, determined using mtDNA and microsatellites. *Molecular Ecology*, 9:2041-2053.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.

Pulliam, H.R. 1988. Sources, sinks and population regulation. *American Naturalist* 132**:**652-661.

Purrenhage, J.L., P.H. Niewiarowski, and F.B.G. Moore. 2009. Population structure of spotted salamanders (*Ambystoma maculatum*) in a fragmented landscape. *Molecular Ecology* 18:235-247.

Raymond, M., and F. Rousset 1995. Genepop (Version 1.2): Population genetics software for exact tests and ecumenicism *Journal of Heredity* 86:248-249.

Spear, S.F., C.R. Peterson, M. Matocq, and A. Storfer. 2006. Molecular evidence for historical and recent population reductions of tiger salamanders (*Ambystoma tigrinum*) in Yellowstone National Park. *Conservation Genetics* 7:605-611.

Spear, S.F., J. Baumsteiger, and A. Storfer. 2008. Newly developed polymorphic microsatellite markers for frogs of the genus *Ascaphus*. *Molecular Ecology Resources* 8:936-938.

Spear, S.F., and A. Storfer. 2008. Landscape genetic structure of coastal tailed frogs (*Ascaphus truei*) in protected vs. managed forests. *Molecular Ecology* 17:4642-4656.

Spear, S.F. 2009. Landscape genetics of habitat alteration across multiple spatial and temporal scales in the anuran genus *Ascaphus*. Ph.D. dissertation, Washington State University.

Steele, C.A., and A. Storfer. 2006. Coalescent-based hypothesis testing supports multiple Pleistocene refugia in the Pacific Northwest for the Pacific giant salamander (*Dicamptodon tenebrosus*). *Molecular Ecology* 15:2477-2487.

Steele, C.A., J. Baumsteiger, and A. Storfer. 2008. Polymorphic tetranucleotide microsatellites for Cope's giant salamander (*Dicamptodon copei*) and Pacific giant salamander (*Dicamptodon tenebrosus*). *Molecular Ecology Resources* 8:1071–1073.

Steele, C.A., J. Baumsteiger, and A. Storfer. 2009. Influence of life history variation on the genetic structure of two sympatric salamander taxa. *Molecular Ecology* 18:1629-1639.

Storfer, A., J.M. Eastman, S.F. Spear. 2009. Modern molecular methods for amphibian conservation. *Bioscience* 59:559-571.

Stuart, S.N., J.S. Chanson, N.A. Cox, B.E. Young, A.S.L. Rodrigues, D.L. Fischman, and R.W. Waller, 2004. Status and trends of amphibian declines and extinctions worldwide. *Science* 306:1783-1786.

Tallmon, D.A., G. Luikart and M.A. Beaumont. 2004. A comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. *Genetics* 167:977-988.

Tallmon, D.A., A. Koyuk, G. Luikart, and M.A. Beaumont. 2008. ONESAMP: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources* 8:299-301.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. *Genetics* 166: 1963-1979.

Welsh, H.H., 1990. Relictual amphibians and old-growth forests. *Conservation Biology* 4: 309-319.

Welsh, H.H. 2005. Coastal giant salamander. Pp. 54-57 *In*: L.L.C. Jones, W.P. Leonard, and D.H. Olson (editors), *Amphibians of the Pacific Northwest.* Seattle Audubon Society, Seattle.

Wilkins, R.N. and N.P. Peterson. 2000. Factors related to amphibian occurrence and abundance in headwater streams draining second-growth Douglas-fir forests in southwestern Washington. *Forest Ecology and Management* 139:79–91.

Wright, S. 1969. *Evolution and the genetics of populations. Vol 2 The theory of gene frequencies.* University of Chicago press, Chicago.

7. Glossary

Alleles – the different forms of a gene.

Allelic richness – The number of alleles in the sampled population, adjusted for unequal sample sizes among populations.

Effective population size – The population size that represents the number of breeding individuals in a population, and therefore is a more meaningful population size for predicting population response to change and long-term viability

Heterozygosity – For diploid organisms (two sets of chromosomes), heterozygous individuals have two different alleles at a single locus.

Heterozygosity excess test – A test for recent large population declines, in which an excess of heterozygotes (relative to equilibrium expectations) is expected because allelic diversity is lost faster than heterozygosity. However, this excess signature only exists for 5-10 generations after a decline, and historic declines would actually be expected to have a heterozygote deficiency.

Inbreeding coefficient (F_{IS}) – A metric based on the difference between observed and expected heterozygosity that is generally indicative of non-random mating within a population. Positive values of F_{IS} represent the results of mating between individuals more closely related than random, and negative values represent the results of mating between individuals more distantly related than random.

Linkage disequilibrium – A statistical association between two loci such that alleles at one locus are correlated with presence of particular alleles at another locus. If linkage disequilibrium exists, loci cannot be considered independent. Linkage disequlibrium could be due to several factors ranging from proximity on the same chromosome to population substructure.

Locus (plural: loci) – the location on a chromosome where a gene is located.

M-ratio – A test to detect population declines that is based on the ratio of number of alleles to allele size range. Allelic size range is the difference in size of the longest microsatellite allele from the size of the smallest allele. As a population bottleneck or decline occurs, alleles of all sizes are randomly lost, and thus the number of alleles is reduced faster relative to the total size range, and a lowered M-ratio results. A critical value based on effective population size can be calculated to determine if a M-ratio value signifies a decline.

Microsatellite DNA marker – A locus that is characterized by alleles with multiple repeats of a short sequence of nucleotides. These loci do not code for proteins (i.e. are selectively neutral) and mutate rapidly. Therefore, these loci have high genetic variability and are excellent for investigating fine-scale genetic differences.

Appendix 1**:**

Multiplex PCR conditions for *Ascaphus* (locus names are as in Spear et al. 2008)
All runs have 10 µl total volumes/sample
All multiplexes have: 5ul Qiagen Master Mix/sample
0.5 µl Q solution/sample
1 µl template DNA/sample
All primer volumes are equal for both forward and reverse primers, and are at 5 µM stock concentrations.

Multiplex 1- [96 deg 15 min, (94 deg 30 s, 55 deg 90 s, 72 deg 60 s)29X, 60 deg 30 min]
A2 – 0.3 µl
A4 – 0.5 µl
A12 – 0.05 µl
A31 – 0.05 µl
1.7 µl water

Multiplex 2 – [96 deg 15 min, (94 deg 30 s, 55 deg 90 s, 72 deg 60 s)28X, 60 deg 30 min]
A15 – 0.3 µl
A24 – 0.15 µl
A26 – 0.25 µl
A29 – 0.1 µl
1.9 µl water

Multiplex 3 – [96 deg 15 min, (94 deg 30 s, 60 deg 90 s, 72 deg 60 s)26X, 60 deg 30 min]
A1 – 0.075 µl
A3 – 0.45 µl
A13 – 0.1 µl
A14 – 0.15 µl
A17 – 0.15 µl
1.65 µl water

Dilute PCR product 1:15 (1 µl template PCR product, 0.25 µl LIZ500 size standard, 13.75 µl water)

Appendix 2:

Multiplex PCR conditions for *Dicamptodon* (locus names are as in Steele et al. 2008)
All runs have 10 µl total volumes/sample
All multiplexes have: 5ul Qiagen Master Mix/sample
0.5 µl Q solution/sample
1 µl template DNA/sample
1 µl primer mix (both forward and reverse mixed at a concentration of 2 µM each)
2.5 µl water

Multiplex 1- [95 deg 15 min, (94 deg 30 s, 53 deg 90 s, 72 deg 60 s)30X, 60 deg 30 min]
D18 – 6FAM
D13 – VIC
D04 – PET
D25 - NED

Multiplex 2 – [95 deg 15 min, (94 deg 30 s, 60 deg 90 s, 72 deg 60 s)30X, 60 deg 30 min]
D24 – 6FAM
D17 – PET
D07- NED
D20 - VIC

Multiplex 3 – [95 deg 15 min, (94 deg 30 s, 60 deg 90 s, 72 deg 60 s)30X, 60 deg 30 min]
D23 – VIC
D05 – NED
D14 – 6FAM
D06 – PET
D15 – 6FAM

Dilute PCR product 1:20 (1 µl template PCR product, 0.25 µl LIZ500 size standard, 18.75 µl water)

Appendix 3. F_{ST} pairwise values for Coastal tailed frogs across all Type N basins. Bold, italicized value indicates sites that are significantly genetically differentiated.

	1099	1197	1236	2260	2468	3074	3098	3110	3111	3437	3576	3914	5785	5378
1197	***0.01***													
1236	0.00	0.01												
2260	***0.05***	***0.06***	***0.05***											
2468	***0.05***	***0.06***	***0.04***	***0.03***										
3074	***0.04***	***0.05***	***0.04***	***0.03***	***0.02***									
3098	***0.03***	***0.04***	***0.03***	***0.03***	***0.02***	0.00								
3110	***0.03***	***0.04***	***0.03***	***0.02***	***0.02***	0.01	0.00							
3111	***0.04***	***0.04***	***0.03***	***0.03***	***0.02***	0.01	0.00	0.00						
3437	***0.03***	***0.04***	***0.03***	***0.03***	***0.02***	***0.01***	0.01	0.01	0.01					
3576	***0.04***	***0.05***	***0.04***	***0.04***	***0.02***	***0.01***	***0.01***	0.01	***0.01***	0.01				
3914	***0.04***	***0.04***	***0.03***	***0.02***	***0.01***	0.01	0.00	0.00	0.00	0.00	***0.00***			
5785	***0.03***	***0.04***	***0.03***	***0.03***	***0.02***	***0.01***	0.01	0.00	0.00	0.01	***0.01***	0.00		
5378	***0.08***	***0.09***	***0.08***	***0.08***	***0.07***	***0.04***	***0.06***	***0.04***	***0.06***	***0.05***	***0.05***	***0.05***	***0.05***	
6000	***0.07***	***0.09***	***0.07***	***0.07***	***0.06***	***0.05***	***0.06***	***0.04***	***0.05***	***0.05***	***0.05***	***0.05***	***0.04***	0.00

Appendix 4. F_{ST} pairwise values for Cope's giant salamander across all Type N basins. Bold, italicized value indicates sites that are significantly genetically differentiated.

	363	1099	1197	1236	2260	2468	3074	3098	3110	3111	3437	3576	3914	5378	5785	6000	5595N
1099	***0.0946***																
1197	***0.1271***	***0.0329***															
1236	***0.1048***	***0.0441***	***0.0522***														
2260	***0.1403***	***0.0569***	***0.055***	***0.0897***													
2468	***0.1651***	***0.074***	***0.0712***	***0.0916***	***0.0406***												
3074	***0.1781***	***0.1272***	***0.1129***	***0.142***	***0.065***	***0.0626***											
3098	***0.2048***	***0.1429***	***0.1289***	***0.1497***	***0.0757***	***0.0846***	0.0064										
3110	***0.2099***	***0.1562***	***0.1446***	***0.1541***	***0.0893***	***0.0927***	***0.0397***	0.0172									
3111	***0.196***	***0.1451***	***0.1283***	***0.1435***	***0.0894***	***0.0818***	***0.029***	0.0162	***0.0338***								
3437	***0.1679***	***0.1331***	***0.1155***	***0.1351***	***0.0802***	***0.0821***	***0.016***	***0.0196***	***0.0333***	***0.0225***							
3576	***0.2068***	***0.1526***	***0.1274***	***0.1421***	***0.0971***	***0.092***	***0.0314***	***0.0232***	***0.0467***	***0.0429***	***0.0214***						
3914	***0.2309***	***0.1602***	***0.1629***	***0.1662***	***0.0924***	***0.0911***	***0.0456***	***0.042***	***0.0592***	***0.0601***	***0.0481***	***0.0531***					
5378	***0.3397***	***0.2827***	***0.2816***	***0.2919***	***0.2432***	***0.2278***	***0.2041***	***0.2128***	***0.213***	***0.1929***	***0.1845***	***0.2055***	***0.2212***				
5785	***0.1891***	***0.1477***	***0.1382***	***0.1514***	***0.09***	***0.0945***	***0.0352***	***0.0385***	***0.0448***	***0.0488***	***0.0366***	***0.0512***	***0.0174***	***0.1894***			
6000	***0.2552***	***0.2197***	***0.2085***	***0.2206***	***0.1742***	***0.16***	***0.119***	***0.1441***	***0.1479***	***0.1311***	***0.1159***	***0.1365***	***0.1466***	***0.1152***	***0.1195***		
5595N	***0.27***	***0.2287***	***0.2087***	***0.2338***	***0.165***	***0.1521***	***0.1015***	***0.1158***	***0.1336***	***0.0977***	***0.0915***	***0.1134***	***0.1384***	***0.1688***	***0.1095***	***0.099***	
5595S	***0.2488***	***0.2087***	***0.197***	***0.215***	***0.1586***	***0.1486***	***0.1003***	***0.1208***	***0.1318***	***0.0976***	***0.0923***	***0.1076***	***0.1257***	***0.1586***	***0.1038***	***0.1051***	0.0101

Appendix 5. F_{ST} pairwise values for Coastal giant salamanders across all Type N basins. Bold, italicized value indicates sites that are significantly genetically differentiated.

	3074	3098	3110	3111	3437	3576	3914	5378	5785	6000	5595N
3098	-0.0029										
3110	0.0137	0.0168									
3111	0.005	0.0321	0.0654								
3437	0.1891	0.1228	0.1957	***0.1116***							
3576	0.0117	0.019	0.0708	0.0281	0.1226						
3914	0.0145	0.0174	***0.0389***	***0.0647***	0.1286	0.0354					
5378	0.1769	***0.1612***	***0.1973***	***0.1211***	***0.1581***	0.1732	***0.1415***				
5785	0.0279	0.0083	***0.0324***	***0.0454***	0.135	***0.0307***	0.0126	***0.1133***			
6000	0.0232	0.0337	***0.072***	***0.0315***	0.0912	0.0509	***0.0538***	0.0862	***0.0412***		
5595N	***0.0918***	***0.0939***	***0.1358***	***0.0843***	0.1315	***0.1143***	***0.096***	***0.0445***	***0.0743***	***0.0404***	
5595S	***0.0403***	***0.0407***	***0.0719***	***0.0466***	***0.1196***	***0.0637***	***0.0648***	***0.095***	***0.0402***	***0.0214***	0.0199

Questions Leading to a Forests & Fish Policy Adaptive Management Recommendation to the Forest Practices Board

Project title: **Type N Experimental Buffer Treatment Project in Basalt Lithologies**
Report title: **Type N Experimental Buffer Treatment Study: Baseline Measures of Genetic Diversity and Gene Flow of Three Stream-Associated Amphibians.**

1. **Does the study inform a rule, numeric target, performance target, or resource objective?**

   The Type N Experimental Buffer Treatment Study (hereafter Type N Study) is an effectiveness study that will inform whether Type N riparian prescriptions are meeting Forest Practices Habitat Conservation Plan (FP HCP) resource objectives. This document is specific to the amphibian genetics component of the study.

2. **Does the study inform the Forest Practices Rules, the Forest Practices Board Manual guidelines, or Schedules L-1 or L-2?**

   A focus is addressing Schedule L-1, Key Question 2, "Will the prescriptions produce forest conditions and processes that achieve resource objectives while taking into account the natural spatial and temporal variability inherent in forest ecosystems?"

   With regard to existing rules, the amphibian genetics portion of the study will inform understanding of the effectiveness of forest practices buffers in maintaining the viability of amphibian populations through the identification of potential changes in genetic structure. This study provides the pre-harvest baseline to enable that evaluation.

3. **Was the study carried out pursuant to CMER scientific protocols (i.e., study design, peer review)?**

   Yes. CMER has reviewed and approved the study plan and final report and they have both been through ISPR (Independent Scientific Peer Review).

4. **What does the study tell us? What does the study not tell us?**

   The amphibian genetics component of the Type N Study was originally designed to evaluate genetic diversity and genetic neighborhood for two species of stream-associated amphibian (Coastal Tailed Frog and Cope's Giant Salamander, the two focal study species that are known to occur throughout the range of all study basins). However, because two species of giant salamanders co-occur at some sites and because evidence of hybridization was present, assessment of Coastal Giant Salamander was also included.

   The amphibian genetics component of the Type N Study will estimate potential changes in genetic structure associated with applications of the FP HCP patch buffer compared to: no riparian buffer, a complete riparian buffer, and an unharvested reference. This report

provides pre-harvest estimates of population sizes and genetic structure. It also gives estimates of genetic neighborhood for the species for which genetic structure is assessed.

The genetic data will also help distinguish between giant salamander species, Cope's and Coastal Giant Salamanders (*Dicamptodon copei* and *D. tenebrosus*) during density sampling to ensure proper determination of species composition and proper calculation of species richness and density metrics for each species. Genetics are required to distinguish the two species because the smaller larval stages (< 55 mm SVL – snout to vent length) cannot be distinguished morphologically with any confidence (e.g., Wilkins and Peterson 2000). Genetic data will also allow identifying hybrid individuals, as well as their frequency of occurrence and distribution.

This report does not present how amphibian genetics are affected by implementation of the above mentioned harvest treatments. To inform this, post-harvest data will need to be collected 7-8 years (2017/2018) after harvest (allowing approximately one generational turnover). This portion of the study does not address any of the other questions that the Type N Study was designed to address.

5. What is the relationship between this study and any others that may be planned, underway, or recently completed?

Several studies in the Type N Amphibian Response Program (NARP), Type N Delineation Program (NDP), and Stream Typing Program (STP), and the sole study in the Sensitive Site Program (SSP), inform or directly support this study. Several other studies relate indirectly to this study; Stream-Associated Amphibian (SAA) Detection/ Relative Abundance Methodology (NARP), Dunn's and Van Dyke's Salamander Study (NARP), SAA Sensitive Site Characterization (NARP), and SAA Site Identification Methods (SSP) all provide important information regarding how best to sample amphibians and which Type N habitats to include. The Tailed Frog Pilot Meta-analysis (NARP), the Amphibian Recovery Project, the pilot and full projects for the Perennial Stream Survey (NDP), and the Last Fish/Habitat Prediction Model Development also provided guidance for the appropriate design of this study.

This study will inform the development of the Amphibians in Intermittent Streams and the Amphibians and Geology studies.

6. What is the scientific basis that underlies the rule, numeric target, performance target, or resource objective that the study informs? How much of an incremental gain in understanding do the study results represent?

Research in managed forests has left a high degree of uncertainty in how Type N streams function under the forest practice rules with regard to long-term amphibian occupancy. Data on amphibians from Type N systems suggest that stream-associated amphibians are ecologically important in headwater habitats, but how they respond to current forestry practices is essentially unknown. Results from this study will increase our basic knowledge

of amphibian viability (population sizes and genetic variability), management effects, design of future studies, and the value of the FP HCP Type N prescription in specific areas.

This study provides no direct gain, but it provides the critical pre-harvest baseline that can be compared to post-harvest genetic sampling, which will produce a substantial gain in the understanding of how Type N buffers effect amphibian population sizes and genetic variability.