Snohomish County Public Works Department
Surface Water Management Division

Precision and Repeatability of Wadable Stream Habitat Survey Methods

Project No. 236

November 2002

Snohomish County Public Works Department
Surface Water Management Division
2731 Wetmore Avenue, Suite 300
Everett, WA 98201

Contact: Frank Leonetti (425) 388-3464 Ext. 4659

Working together for people, streams & fish.

Snohomish County
PUBLIC WORKS
SURFACE WATER MANAGEMENT
(425) 388-3464

Acknowledgements

Snohomish County Council

John Koster	Council District 1
Kirke Sievers	Council District 2
Gary Nelson, Chair	Council District 3
Dave Gossett	Council District 4
Jeff Sax	Council District 5

Snohomish County Executive

Robert Drewel

Department of Public Works

Peter Hahn, Director

Surface Water Management (SWM) Division

Joan Lee, P.E., Director
Meg Moorehead, Salmon Recovery Program Manager

Project Manager

Michael Purser

Project Staff

Andrew Haas, Frank Leonetti, L. Ted Parker, Michael Rustay, Rob Simmonds

Contributing Staff

The following people have made significant contributions to this project and the completion of this report: SWM staffers Robert Aldrich, Suzanne Brunzell, Christine Coats, Kathy Hagerty, Deborah Haynes, Dave Ward and the Contracted Staff. Jeff Collins and Jennifer Slavick were project interns in 2001. Comments and suggestions provided by Ashley Steel (NWFSC, NOAA-Fisheries), Kurt Nelson (Tulalip Tribes) and Dr. Fran Solomon (King County) were extremely valuable in the completion of this document.

Cover photograph Large Woody Debris in Squire Creek a tributary to the North Fork Stillaguamish River, 2000. Photographer L. Ted Parker.

printed on recycled paper

Executive Summary

Snohomish County developed an Early Action Salmon Conservation Program in anticipation of the Federal listings under the Endangered Species Act of Puget Sound salmonids (i. e. chinook salmon (*Oncorhynchus tshawytscha*) and bull trout char (*Salvelinus confluentus*)) for the purpose of promoting the recovery of these species. . The Early Action plan included a commitment to inventory and monitor stream physical habitat on a subbasin scale.

To meet this commitment, Snohomish County Public Works Surface Water Management (SWM) initiated a multi-year stream habitat inventory of fish-bearing streams in Snohomish County. The methodology is County described in *Physical Habitat Survey and Monitoring Protocol for Wadable Streams* versions 5.0 and 6.0 (Snohomish County 2000, Snohomish County 2001b). These data are intended to provide a better understanding of current conditions, trends, restoration opportunities and a technical foundation for Water Resource Inventory Area (WRIA) planning, as well as to support stormwater, capital improvement projects and road maintenance programs.

In 2000 and 2001 SWM habitat staff and contractors surveyed streams in 10 subbasins throughout Snohomish County. In total, approximately 96.5 km (60 miles) of stream, with 247 unit reaches were surveyed, 23 unit reaches were resurveyed for quality control in the same year to estimate the precision and repeatability of the survey method.

This document explores the suitability and effectiveness of the stream habitat survey method developed in 2000 and refined in 2001. It addresses questions posed at the beginning of the project such as:

- How repeatable is the method?
- Are there gaps in the protocol that prevent SWM from answering critical management questions?
- Was any data collected that is not useful in answering management questions?

This document will also examine the hypothesis formed early in the development of the survey protocol that would allow surveyors to monitor just one bank for instability.

Summary of Analysis Results

Bankfull Width (BFW)- Analysis of the year 2000 replicate surveys for BFW measures resulted in minor changes in the 2001 protocol. Analysis of 2001 data showed bankfull width measurements were precise and repeatable between replicate surveys.
Riffle Wetted Width (RWW)- In general, measurements of RWW were highly repeatable.
Wood-Analysis of repeatability of Large Woody Debris (LWD) and Small Woody Debris (SWD) classifications were marginally precise and similar. Additional tests

should be conducted to estimate the repeatability of Stumps and Rootwads (SSR) counts and assess the more qualitative measurements of decay class, wood type, rootwad presence and logjam association.

Pools- Pool frequency (pools/km) and percent pool area, had relatively high repeatability between replicate samples. At this level of analysis, pool measurements were generally found to be repeatable between replicate surveys.

Surface Fine Sediment- Percent surface fine sediment measurements have low precision and repeatability compared to the other measured parameters. In reaches with high variability of fines distribution throughout the reach, there was a corresponding high variation between survey teams. Streams with either very low or very high levels of fines seem to have higher repeatability than other streams.

Instability- Bank instability measurements have high precision and repeatability suggesting that survey methods are appropriate.

The variability observed in a majority of the parameters reflects the variability which is seen in habitat conditions, and is not reflective of the error variance associated with data collected from different surveying teams. These results are consistant with previous studies examining the variance in and error associated with stream habitats (e.g. Kaufman et al. 1999). Where error was attributed to differences between survey teams (surface fine sediments, wood), recommendations to improve the precision of surveys are offered below.

Presently, there are no gaps in the protocol that will prevent us from answering critical management questions as they apply to wadable streams.

Recommendations for 2002

Bankfull Width (BFW)- Changes made in 2001 protocol are sufficient.

Riffle Wetted Width (RWW)- Changes made in 2001 protocol are sufficient.

Large Woody Debris (LWD) - Provide additional clarity in description of channel dimensions and the channel location from which LWD is sampled. Emphasize location in field training.

Small Woody Debris (SWD)- Provide additional clarity in description of channel dimensions and the bankfull channel from which SWD is sampled. Emphasize location in field training.

Pools- Continue with current protocol. Perform additional training of SWM and contracted survey personnel.

Fines- Revise, the protocol to ensure a minimum number of riffles are sampled per unit reach Increase the number of riffles sampled for unit reaches with high surface fine variability. Perform additional training of survey crews.

Instability - Continue with current protocol. Perform additional field training.

SUMMARY

This report represents an analysis of the effectiveness, suitability, and repeatability of selected habitat parameters from Snohomish County's wadable stream physical habitat sampling protocol (Snohomish County 2001b). Habitat parameters from the protocol were evaluated by conducting replicate surveys, which comprised approximately 10% of survey length in 2000 and 2001. This approach was also used to evaluate precision of measurement between survey teams. This effort was undertaken principally to address the first of five key management questions, but results gained here will contribute substantially to future analysis to answer additional management questions:

1. How effective and suitable is the survey methodology developed for this analysis?
2. How do existing habitat conditions compare to local targets?
3. How does land use/land cover impact instream physical habitat conditions?
4. How does the data inform a restoration strategy?
5. What are the physical habitat trends over time?

Precision of wadable stream habitat variables ranged from very high to low. In general, streambank or wetted channel associated measurements (i.e. bankfull width, riffle wetted width, bank instability, hydromodifications) and pools were most precisely measured followed by wood and fine sediments. Whereas, wood is clearly defined and the measurement protocol described, it is often difficult to access and large accumulations hinder observations. Fine sediment criteria and measurement position within riffles are most difficult to define or interpret in the field. Thus, precision and repeatability are reflective of the variability in habitat conditions and the degree to which field interpretation is required (sometimes for many situations) in order to implement methods in the field for all parameters of interest. These results are consistent with previous efforts examining the variance in and error associated with wadable stream habitats (e.g. Kaufmann et al. 1999). Where error was attributable to differences between survey teams, steps are recommended to improve the precision of surveys based on additional rule making for implementing the protocol.

Additional survey team training is warranted for each methodology used, as is adopting an approach for replicate surveying (original vs. quality control) that is reliant on one day replicate survey sampling of known pool, wood and fine sediment conditions. This will control for stream discharge, start and end points, pool selection, and riffle transect placement, in particular. This will allow greater scrutiny of the manner in which the protocol and field methodologies are applied and will contribute to the reduction in error among survey teams, especially for measurement of fine sediments. A reduction in error variance will help to both minimize future habitat monitoring effort, and increase likelihood that trends in habitat condition over time will be detectable. This increases confidence that not only are observed trends real, but that management actions are being directed to address accurately defined problems.

INTRODUCTION

In May 1999, the National Marine Fisheries Service (NMFS) listed chinook salmon (*Oncorhynchus tshawytscha*) in the Puget Sound Ecologically Significant Unit (ESU) as a threatened species under the Endangered Species Act. Several months later the U.S. Fish & Wildlife Service (USFWS) listed bull trout char (*Salvelinus confluentus*) throughout the Coastal/Puget Sound Distinct Population Segment (DPS) as a threatened species. To promote species recovery, in anticipation of these listings, Snohomish County developed a Salmon Conservation program. The program includes the baseline inventory of instream physical habitat described in this report. These data are intended to provide a better understanding of current conditions, trends, restoration opportunities and a technical foundation for Water Resource Inventory Area (WRIA) salmon conservation planning, as well as to support drainage, capital improvement and road maintenance programs. The inventory covers fish-bearing streams in Snohomish County subbasins and an interpretation of land cover (e.g. impervious surface, mature forest, etc.) using satellite imagery.

In 2000 and 2001, Snohomish County Public Works Surface Water Management (SWM) staff and contractors surveyed approximately 96.5 km (60 miles) of stream in 10 subbasins throughout Snohomish County (Map 1) using methods described in *Physical Habitat Survey and Monitoring Protocol for Wadable Streams* versions 5.0 and 6.0 (Snohomish County 2000, Snohomish County 2001b). During these two years, habitat data from 247 survey reaches were collected, processed and reported. Of these 247 survey reaches, 23 reaches were resurveyed during the same year as the original survey to estimate the precision and repeatability of the survey method.

This report addresses the first of several key management questions (below). It focuses on questions such as: How repeatable is the method? Are there gaps in the protocol that prevent SWM from answering critical management questions? This document will also test a hypothesis that no difference exists between right and left streambank instability requiring only one streambank for monitoring. The results gained here will also contribute to the ability to answer the second key management question:

1. How effective and suitable is the survey methodology developed for this analysis?
2. How do existing habitat conditions compare to local habitat suitability criteria?
3. How does land use/land cover impact instream physical habitat conditions?
4. How does the data inform a restoration strategy?
5. What are the physical habitat trends over time?

METHODS

Data were collected during original (OR) and quality control (QC) surveys of 23 replicate survey reaches. This section describes data management and processing procedures as well as the analytical approach.

Data Collection and Quality Assurance

Survey data collected in the field were entered directly into Microsoft Windows CE®based field personal computers (PC), eliminating the office data entry step associated with traditional paper field forms. Field data management was initiated at the start of each individually surveyed reach or "unit reach." Each unit reach file was saved as a numeric identifier (generated for each reach before the field season using Geographic Information System (GIS) software), and reach and surveyor-identifying information were stored in the header sheet of the workbook. To minimize manipulation of data during the transfer of files to a desktop PC, individual entries for each of the measured and calculated habitat parameters were compiled into a single spreadsheet during the collection process. Once uploaded to an office PC, the data were assessed for completeness and quality and added to a master spreadsheet file in Microsoft Excel®. Identified errors were corrected at this time. Entry corrections were flagged with attached comments describing what values were changed and why. Other questionable entries were flagged for additional data quality assessment. Occasionally a value was encountered that was incomplete or unusable. These entries were flagged and values were removed from the master file. Comments (embedded within the master spreadsheet) document the reasons for any changes and contain the altered cell's original values. When data collection was complete for the field season, master files received another round of quality checks and were imported into Microsoft Access® for analysis. Raw and processed files were saved on regularly backed-up network drives as well as on CD-ROM.

Data Analysis

SWM's protocol was designed to rapidly, but quantitatively, assess the habitat characteristics of a stream reach. It relies primarily on quantitative measurements of specific parameters rather than qualitative interpretation. This method allows data to be processed and analyzed on many levels, from individual measurements of discrete features to aggregate values at unit reach, reach and subbasin scales. Station numbers, collected as part of the survey, make it possible to locate paired features in OR and QC surveys and compare individual measurements of these features. This type of comparison was used sparingly in the analysis because station numbers between replicate reaches appeared different enough to make pairing of specific features time consuming and subject to bias. For most parameters, this analysis will focus on aggregate values calculated for each replicate pair of reaches.

Statistica® software (v. 5.5, Statsoft 1999) was used to generate statistics for many of the survey parameters. Wilcoxon matched pairs tests and sign tests were used for each paired reach to assess differences between pairs or bias (Zar 1984). For all statistical tests, an

α=0.05 was chosen for evaluating the significance of the tests. Additionally, expressions of precision and repeatability were calculated as the;

- Root Mean Square Error (RMSE or σ_{rep}, Kaufmann *et al.* 1999)
- Signal to Noise ratio (S/N, Kaufmann *et al.* 1999), and
- Repeatability (R, Krebs 1989)

RMSE is defined as $\sigma_{rep,}$ where,

σ_{rep} =$\sum$ standard deviation of repeat measurements of a habitat metric.

S/N is defined as a comparison of the variance of a "habitat metric observed across a regional sampling of streams ("signal") with the variance resulting from [replicate] field measurements within the sampling season" ("noise"), and is computed as,

$$S/N=\text{Variance of a population} / \sum \text{Variance between replicated pairs.}$$

The larger the calculated S/N value the more precise the measurement. Generally, values <2.5 are considered imprecise, between 2.5 and 6 are moderately precise and >6 are precise (for discussion see Kaufman *et al.* 1999).

Repeatability, R, is a value between 0 and 1, and the closer R is to 1 the more precise the measurement. Repeatability is calculated as,

$$R = (\text{Variance among unit reaches})/(\text{Variance within replicates} + \text{Variance among unit reaches}).$$

RESULTS & DISCUSSION

Bankfull Width (BFW)

It is important to know whether bankfull width can be measured precisely and similarly among survey teams. An adjustment in BFW is one way streams respond to changes in sediment, hydrologic, or woody debris inputs over time. Channels, especially in alluvial settings, may widen or narrow over time, and the ability to detect real changes will be dependent, in part, on the variance of BFW measurement. For the purposes of this survey, other habitat measures of pools and large woody debris (LWD) are also dependent on the bankfull width measurement. Thus, it is important that BFW be precisely measured to minimize the probability that pool size or LWD classes are sampled differently among survey teams.

In 2001, BFW was sampled at the beginning, middle, and end of each unit reach. Previously in 2000, one BFW measurement was taken at the beginning of each reach surveyed and subsequent measurements were taken only if there was a significant observed change in BFW dimensions. In 2 of 7 QC reaches sampled in 2000, one survey team reported a change in BFW part way through the reach that changed the pool and wood size criteria. In two other reaches sampled in 2000, different BFW measurements required the selection of different pool size criteria for those QC reaches. Because of this inconsistency, these pools were not included in the analysis of pool measures. In 2001, no replicate reaches were placed in different pool classification categories.

Among paired observations (OR and QC replicate measures) of BFW in 23 reaches, there were no indications that BFW was different at α=0.05 (Wilcoxon matched pairs test, Table 1). The pooled replicate standard deviation, or σ_{rep}, S/N, and R for replicate surveys of BFW were also calculated. BFW measures for replicate surveys were found to be very precise (Table 2).

Riffle Wetted Width

Riffle wetted width (RWW) is influenced by stream discharge. It is important to sample at similar flow regimes since pool and channel dimensions, used to calculate pool and total channel area, may be affected. Wilcoxon matched pairs testing of mean RWW measures between teams was non-significant at α=0.05 (Table 1), suggesting samples were taken throughout the season at appropriate low flow conditions. The pooled replicate standard deviation, or σ_{rep}, S/N, and R for replicate surveys of RWW were also calculated. RWW measures for replicate surveys were found to be very precise (Table 2). However, replicate samples of mean RWW for two unit reaches (115.11 and 115.22) were significantly different ($t=2.4$, $p=0.04$, $n=6$ and $t=4.3$, $p<0.001$, $n=6$, respectively).

Table 1. Wilcoxon matched pairs test for differences between paired unit reaches (SWM staff vs. Contracted surveyors and OR vs. QC).

		Test/statistic	Significance	Test/statistic	Significance
	n	Wilcoxon matched pairs test/ z	p (a=0.05)	Sign test/ z	p (a=0.05)
SWM vs. Contractor					
Pool count	42	z=2.4	0.017	z=2.58	0.01
Pool frequency	42	z=2.5	0.01	z=2.75	0.006
Pool area, wetted %	42	z=0.7	0.5	z=1.2	0.24
Pool area, functional %	42	z=1.2	0.2	z=0.7	0.5
SWD frequency	32	z=2.9	0.003	z=2.8	0.006
LWD frequency	46	z=0.78	0.4	z=0.67	0.5
Fine sediments, %	34	z=2.4	0.016	z=2.9	0.003
OR vs. QC					
BFW	46	z=1.49	0.14	z=1.25	0.2
RWW	44	z=0.48	0.62	z=0.0	1
Pool count	42	z=1.0	0.3	z=0.25	0.8
Pool frequency	42	z=1.4	0.15	z=0.0	1
Pool area, wetted %	42	z=1.5	0.13	z=0.9	0.4
Pool area, functional %	42	z=1.7	0.09	z=0.9	0.4
SWD frequency	32	z=2.5	0.01	z=1.75	0.08
Bank instability, %	46	z=0.56	0.57	z=0.46	0.65

Shaded cells indicate a significant difference among matched pairs.

These two unit reaches were located on the same stream channel as five other replicate reaches, surveyed on the same day. Some factors that may be responsible for this difference include change in stream discharge that affected width only at certain locations, different measurement tools, and variance from established protocol.

Wood

In 2001, large woody debris (LWD) and stumps (wood classes 1, 2 and 3; Snohomish County 2001b) were enumerated and measured for each reach, as were quantities of small woody debris (SWD). In 2000, SWD was not enumerated and stumps were tallied, but not measured. Wood was grouped for the purposes of this analysis into the largest fraction, LWD, and smallest fraction, SWD, and results were summarized by reach frequency. Thus, only LWD classes 1 and 2 from 2000 and 2001 (the largest wood fraction of the survey and similarly surveyed between years) and SWD from 2001 are included.

σ_{rep}, S/N, and R of LWD and SWD were calculated and measures for replicate surveys were found to be marginally precise for both (Table 2). The results for the SWD fraction are complicated by the fact that SWM staff surveys enumerated more SWD than contracted surveyors in 14 of 16 QC reaches. A Wilcoxon matched pairs test on results, grouped by survey team, reveal highly significant differences at α=0.05. No such differences were observed by survey team for LWD (Table 1). The precision associated with measuring only the largest wood in a survey was approximately equal to SWD measurement precision, but given the apparent bias between survey teams, (i.e. SWM and contracted surveyors) measurement precision for LWD can be corrected more easily, and the data are more reliable in terms of quality.

Table 2. Estimates of Habitat Parameter Precision

QC vs. OR					
	n	Grand mean	RMSE	S/N	R
Habitat parameter					
Bankfull width, m	46	12.1	1.5	45	0.99
Riffle wetted width, m	44	6.5	1.1	27	0.96
Pool count	42	5.1	1	6.9	0.87
Pool frequency (per km)	42	15	4.4	8.2	0.89
Pool area, wetted %	42	14.2	4.2	7.2	0.88
Pool area, functional %	42	8.1	2	10.8	0.92
Small Woody Debris freq. (pieces/km)	32	190.5	62.5	3.7	0.8
Large Woody Debris freq. (pieces/km)	46	27.2	9.6	2.7	0.75
Fine sediment, %, all QC reaches	34	18.2	17.2	0.31	0.27
Fine sediment, %, select paired riffles	38	22.6	20	0.39	0.3
Bank instability, %	46	11	4.5	8.9	0.9

Potential factors that affect the precise counting of LWD and stumps include buried or obscured wood, wood that was not clearly in or out of the bankfull channel and differences in the location (along the length of a piece of wood) of diameter measurements. These differences may cause one team to include a piece, but the next team to exclude it. Potential factors that affect the precise counting of SWD, as reflected in the apparent bias between survey teams, include buried or obscured wood, wood that was not clearly in or out of the bankfull channel and a the proportion of SWD pieces in complex woody debris jams.

Pools

Precise pool measurements depend on two main factors. First, both OR and QC survey teams must be able to distinguish pool features from other stream features. Second, dimensions of an identified pool must be interpreted and measured in a similar fashion. The survey protocol limits surveyor interpretation of pool classification and instead applies rigid quantitative criteria for classification; that is, a pool must have a residual depth (measured to 0.05 m) and functional pool area (measured to 1.0 m) must meet the minimum criteria in the protocol to be recorded. The survey criteria used in 2000 (Johnston and Slaney 1996) limited the collection of data from small and shallow pools that would seem to be more difficult to distinguish from other stream features. The modified 2001 protocol used criteria from Pleus *et al*. (1999), which reduced qualifying pool sizes and residual depths by half. Although criteria differed between years, 2000 and 2001 pool survey results were pooled for these analyses because the protocol was consistent within year.

The σ_{rep}, S/N, and R were calculated for replicate surveys of pool conditions. Pool count and pool frequency (pools/km) estimates for replicate surveys were found to be quite precise and repeatable (Table 2). However, S/N and R are higher for pool frequency estimates, because frequency estimates are normalized to the length surveyed, which sometimes varied between replicate surveys. Pool count is not irrelevant, however, as additional analyses reported below will show. Wetted and functional pool percent area measurements were also precise (Table 2). The higher S/N and Repeatability values for functional pool area indicate that additional quantitative criteria for measuring functional pool size (i.e., minimum depth of 0.2 m on all sides of pools) positively influenced precision.

Differences in pool counts can be attributed primarily to three factors. First, the pool depth and area criteria established in the protocol make it possible for one survey team to count a pool and the other to pass it by based on slightly different pool dimension measurements. Second, because the pool size criteria are based on the bankfull width of the survey reach, differences in the bankfull width measurements between survey teams can affect the number of pools each team records. This occurred only once among the 23 replicate survey reaches, and this reach was excluded from analysis. The third potential reason for differences in pool counts is each team’s bias towards lumping and splitting two or more pools in sequence. While this will have little effect on pool area calculated for a given reach, it has the potential to affect pool frequency, especially if the bias is systematic.

This potential bias was evaluated, as it was clear that surveys conducted by SWM staff generated higher pool frequency estimates in 18 of 20 reaches sampled (Figure 1). Whereas the higher pool frequencies does not affect precision, differences in the grand mean between survey teams was 3 pools/km. Wilcoxon matched pair testing on frequency estimates among replicate surveys was highly significant at α=0.05 (Table 1), indicating a bias in direction and difference. When only direction was tested (Sign test), differences were more extreme at α=0.05 (Table 1). This same bias was not reflected in testing of percent wetted and functional pool area, nor in mean functional pool area (Table 1), indicating the bias was likely due to SWM staff splitting smaller pools from larger pools or that SWM staff occasionally sampled smaller pools than did contracted surveyors. In the 2001 protocol, it was made explicit that splitting pools whenever possible was the preferred method. The fact that percent pool area measurements are more precise than pool frequency measurements suggests there is still some confusion about when to split pools. This bias (of lumping or splitting smaller pools) may have had indirect consequences on the stream riffles selected for sampling fine sediments.

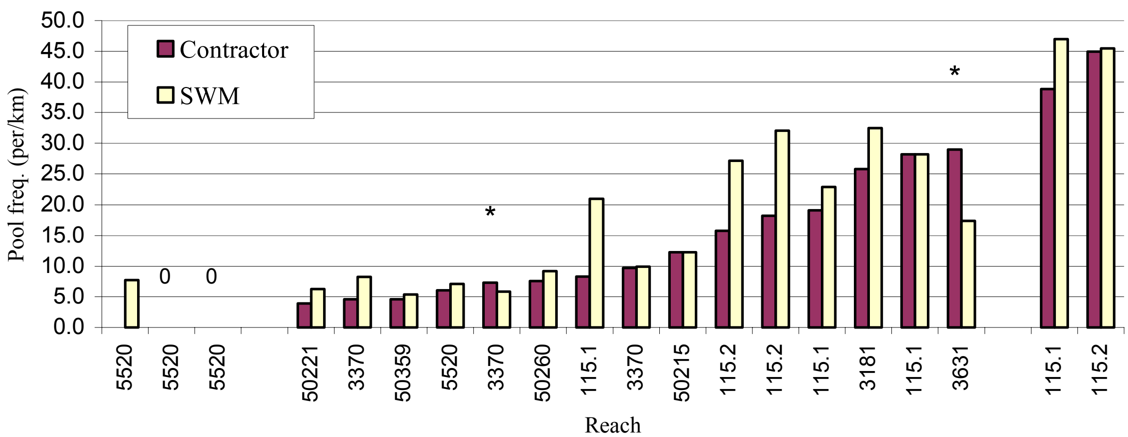

Figure 1. Estimated pool frequency by survey team. Asterisks denote those reaches where contractor estimates were greater than SWM. Missing reach numbers were removed from analysis.

Percent Surface Fine Sediment

Fine sediment samples were taken in stream riffles suitable for spawning (where gravel size, D_{50}, was visually estimated to be <128 mm) downstream from pools meeting protocol criteria for minimum functional pool area and residual pool depth. In 2000, all riffles downstream from qualifying pools were sampled (proving to be time intensive, but sample size did not exceed 11 riffles for any QC/OR reach). In 2001, survey teams sampled the first 4 riffles within a reach downstream from the first 4 qualifying pools, unless D_{50} >128 mm. To test the repeatability and effectiveness of the protocol for sampling and estimating fine sediments in wadable streams, unit reach mean fine sediment estimates from replicate samples were compared as were replicate samples taken within the same riffle (as determined by pairing station numbers and where upstream pool measurements corroborated sample location).

Because survey teams were likely to encounter the same variance in fine sediment composition within a unit reach (regardless of which four riffles were sampled), we assumed measurement of unit reach mean fine sediment was assumed to be repeatable between replicate surveys. The σ_{rep}, S/N, and R were calculated for replicate surveys. In particular, S/N was low, 0.3, as was repeatability, 0.3, due to the high variation between replicate surveys. The σ_{rep} was high relative to the grand mean for all replicate surveys (Table 2). Though the difference in unit reach mean fine sediment values between replicate OR/QC surveys was sometimes striking (Figure 2), detectable differences (α=0.05) were observed for only two of 17 reaches, due to high within-reach variance and low sample sizes in many cases (e.g., only one or two riffles sampled per replicate OR/QC reach, due to the lack of many qualifying pools). When plotting unit reach mean fine sediment values, it was apparent SWM survey teams usually measured higher fine sediment composition (in 15 of 17 cases and every reach sampled in 2001). Although within-reach comparisons of means were generally not significant (as described above; see Table 1), the pairwise ranked sign test of measurements indicated a significant difference between survey teams (Signed rank test; Z=2.4, p=0.016, Table 1).

Observed replicate survey differences in estimated unit reach mean fine sediment composition could be attributable to a number of sources, including survey teams sampling different riffles within the reach, naturally high variance in fine sediment composition among riffles low sample size and a difference between survey teams implementing the protocol and prescribed methodology for sampling fine sediment. By comparing station numbers within reaches and associated qualifying pool measurements, 19 individual riffles (among 55 riffles sampled in 17 replicate unit reaches) were selected where both survey teams sampled fine sediments.

Assuming survey teams were likely to encounter the same variance in fine sediment composition within a riffle, measurement of within-riffle mean fine sediment was assumed to be repeatable between replicate surveys. The σ_{rep}, S/N, and R was calculated for replicate riffles. Similar to unit reach measures of precision, S/N was low, 0.4, as was repeatability, 0.3, due to the high variation between replicate surveys. The σ_{rep} was high relative to the grand mean for all replicate surveys (Table 2). Low repeatability and S/N within riffles is partly explained by the differences measured between teams (sign test; Z=1.84, p=0.07, Table 1), which indicate SWM staff tended to measure a higher proportion of fines within the same riffle as compared to contracted surveyors. Within riffle comparisons of means between survey teams were usually not significant, due to high variation in transect samples. However, the differences in fine sediment measures at 4 of the 19 riffles were significant ($p < 0.05$), and of these, SWM staff sampled a higher proportion of fines in 3 of 4 riffles. These within-riffle mean differences are likely due to inconsistency between survey teams in sample transect placement along the length of the riffle. This inconsistency may be inconsequential, but might also be reflective of differences in perception of spawning gravel suitability between survey teams, even given the description provided in the protocol.

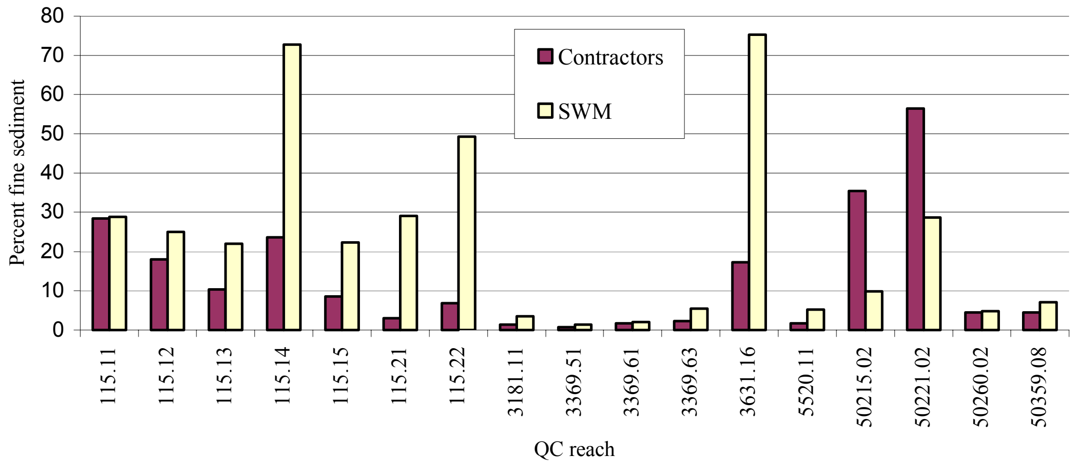

Figure 2. Mean replicate reach percent fine sediments sampled by survey team (Contractors and SWM).

SWM survey teams measured more pools per survey reach than contracted surveyors, resulting in a bias in pool frequency (per km) estimates per replicate reach (Wilcoxon matched pairs test; Z=2.53, p=0.01, Table 1). Pool frequency estimation was repeatable and S/N was higher, so this difference is related more to bias. The additional pools SWM staff measured were not likely to be large pools that contracted surveyors missed, but smaller pools that were marginal in meeting qualifying pool criteria. The difference in mean residual pool depths as measured by contracted surveyors and SWM (z=1.75, p=0.08, n=34) indicates that contracted surveyors tended to measure deeper pools. Thus, there was a greater probability that SWM staff sampled surface fine sediments at smaller pools. The limited dataset of 19 replicate riffles (with associated pools) indicates that both a high and low percentage of fine sediments were associated with smaller pools, but that only a low percentage of fine sediments was associated with larger pools. However, given that the same trend was evident in replicate riffles, additional direct field testing and comparison of survey team proficiency in fine sediment sampling and application of the protocol is necessary. Additional protocol controls on fine sediment sampling location (riffle selection, transect placement, and sample size) are probably warranted and a minimum riffle sample size (four riffles) must be obtained, even if unit reach length exceeds 30 channel widths.

Bank Instability

Bank instability was measured along the right bank in OR unit reach surveys and both right and left banks during QC surveys during the 2000 field season (Snohomish County SWM 2000). In 2001, both banks were surveyed in OR and QC unit reaches (Snohomish County SWM 2001b). Since hydromodifications were not recorded separately from unstable banks during the 2000 field season they were included in the analysis. In order to test the null hypothesis that left and right bank instability are similar (high correlation coefficient value), only the 2001 data was used and hydromodifications were not included.

The pooled replicate standard deviation, or $\sigma_{rep,}$ S/N, and R were calculated for replicate surveys of bank instability. The results indicate replicate measurements are very precise, (Table 1). A Wilcoxon matched pair test of reach mean bank instability measured between survey teams was non-significant (Table 1).

In addition to testing repeatability of bank instability measures, an hypothesis was tested. The hypothesis was that, at a subbasin scale, the amount of bank instability on one bank is equal to the amount of instability on the other. Regression analysis was used to show correlation between right and left bank instability at three scales of measurement; unit reaches, geomorphic reaches and subbasins.

Right and left bank instability (including hydromodifications) based on QC reach data from 2000 and 2001 were correlated ($r=0.73$, $r^2=0.53$, $n=19$; Figure 3) but variance was high, the relationship was leveraged by one large value and results do not support the null hypothesis [slope of regression is not equivalent to 1 based on 95% confidence interval estimation].

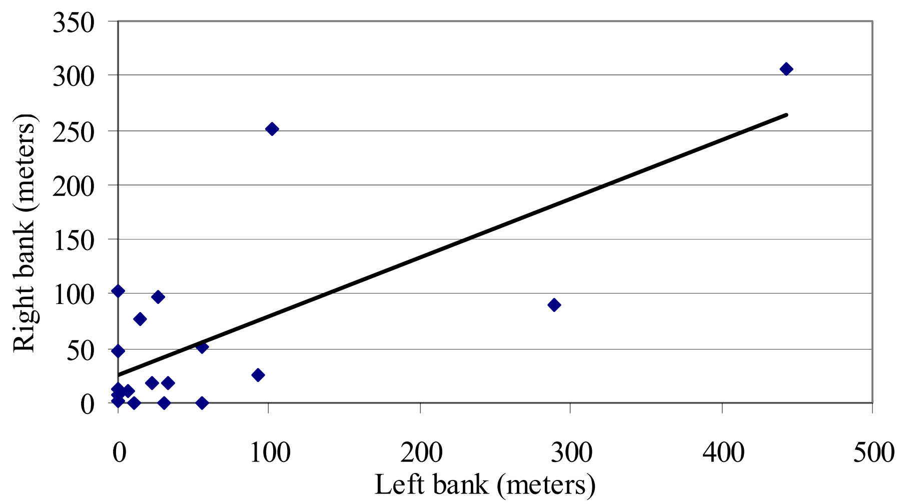

Figure 3. Relationship between QC unit reach right and left bank instability (including hydromodifications) from 2000 and 2001 (r=0.73, r2=0.53, y=0.54x +25.9, n=19).

All data collected in 2001 was used for regression analysis of right versus left bank instability at three scales: unit reach, geomorphic reach and subbasin. Hydromodifications were not included in these analyses. At the unit reach scale, right and left bank instability show no meaningful correlation ($r=0.05$, Figure 4) and the null hypothesis is rejected as the regression coefficient is not equal to one (95% confidence interval $-0.25<\text{slope}<0.39$). At the geomorphic reach scale, the correlation between right and left bank instability (excluding hydromodifications) improves ($r=0.7$) and the null hypothesis is not rejected.

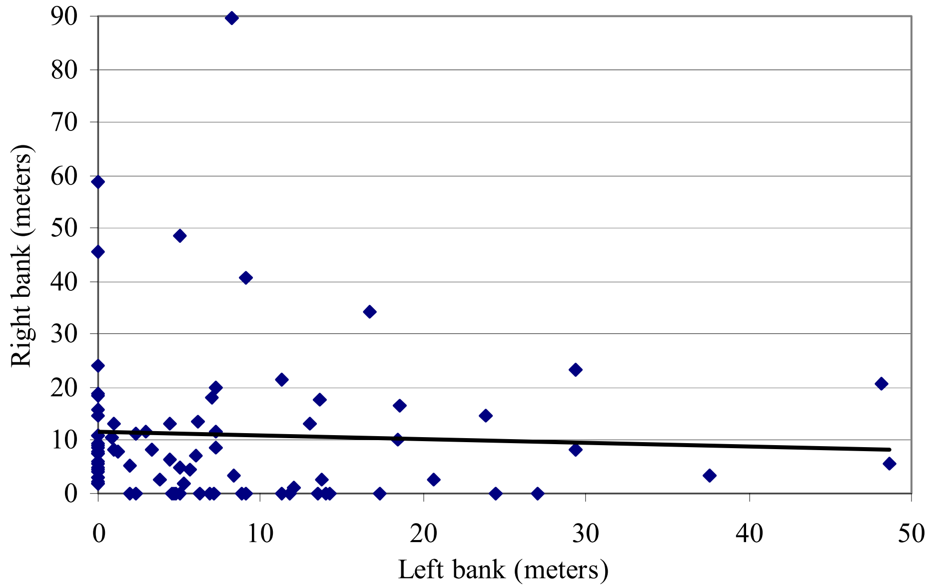

Figure 4. Relationship between 2001 unit reach right and left bank instability (r=0.05, r2=0.002, p=0.7; y=11.6-0.07, n=84).

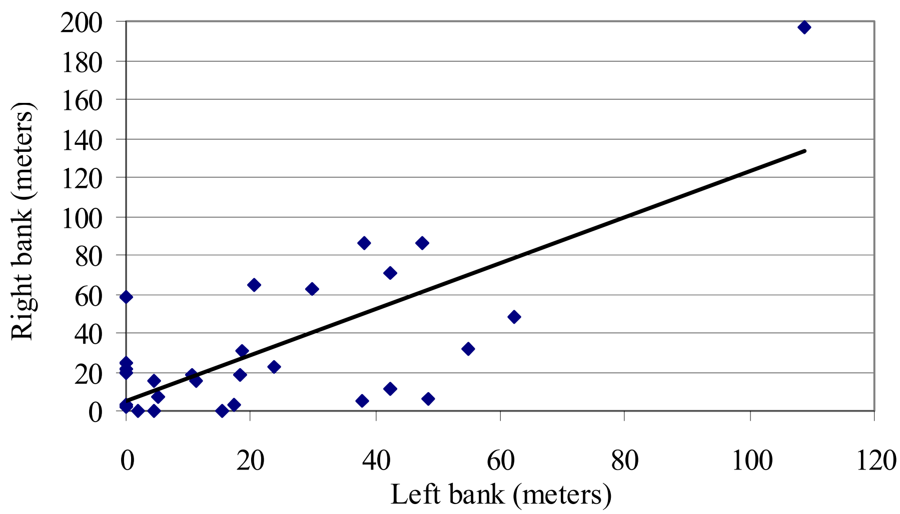

Figure 5. Relationship between geomorphic reach right and left bank instability (r=0.7, r2=0.52, p<0.01; y=1.2x+5.2, n=28).

At the subbasin scale, right and left bank instability (excluding hydromodifications) are highly correlated (r=0.97), and the null hypothesis is not rejected (Figure 6). At the subbasin scale, the extent (cumulative length or percentage with confidence interval) of bank instability should be predicted from sampling only one streambank.

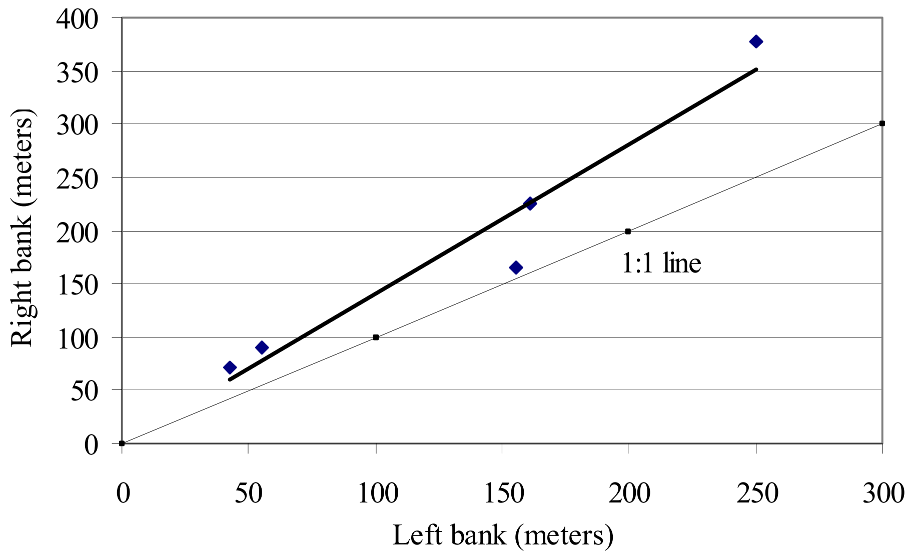

Figure 6. Relationship between subbasin right and left bank instability ($r=0.97$, $r2=0.93$, $p<0.001$; $y=1.4x+0.25$, $n=5$).

The potential relationship between hydromodifications and bank instability was not determined in this report. However, when hydromodification data are included with bank instability (for 2001 data only), the relationship between right and left bank instability is weaker at both the geomorphic reach ($r=0.4$, $y=0.76x+30$, $n=30$) and subbasin scales ($r=0.58$, $y=0.67x+200$, $n=5$). Conversely, when hydromodifications are considered alone, there is greater correlation between right and left bank modification (not shown) than right and left bank instability at the scales considered; unit reach ($r=0.78$, $n=28$), geomorphic reach ($r=0.73$, $n=20$), and subbasin ($r=0.99$, $n=4$), provided survey data from two reaches are reserved from the analysis. In these two cases, hydromodifications dominated one streambank of the unit reach to protect a road grade in upper North Fork Skykomish subbasin. The unit reach and geomorphic reach did not include areas not dominated by these hydromodifications, so at these scales it could not be determined whether hydromodifications may have caused bank instability or have been countered by other hydromodifications up- or down-stream on the opposite bank. Based on these results, total streambank hydromodification may be better predicted from only one streambank than bank instability at the unit or geomorphic reach, especially if the survey area is not dominated by one condition.

DISCUSSION

There are no major gaps in the protocol that prevent addressing two critical management questions as they apply to wadable streams. The protocol allows for the summary of habitat conditions and their comparison against those habitat conditions and criteria being applied regionally. Furthermore, data collected for many other stream habitat parameters or indicators have not as yet been evaluated. These data, such as LWD size, type, decay condition, and association with woody debris jams will be analyzed for their contribution to the relative abundance and quality of wadable stream habitats nested within the context of overall subbasin conditions (e.g., land cover types and watershed groups). Based on this forthcoming work, the wadable stream habitat protocol may be revised if data are not useful or not used. This analysis and reporting is intended to be the subject of future wadable stream habitat reports.

Bankfull Width- Apparent imprecision in year 2000 replicate survey BFW measures is primarily the result of an inconsistent method of recording BFWs and of reliance on one BFW measure per survey. The protocol and field sheets for bankfull width measurements were modified slightly in 2001 to eliminate confusion about when and where to measure bankfull width within a unit reach. With these changes bankfull width measurements in 2001 were found to be precise and repeatable between replicate surveys.

Riffle Wetted Width- Two out of 23 replicate surveys of unit reaches were found to have significantly different mean RWW. This may have occurred because the reach lengths surveyed were non-overlapping. In general, measurements of RWW were highly repeatable.

Wood- Calculated S/N and repeatability (R) values for LWD and SWD are marginally precise. Measures of LWD were more precise than measures of SWD because LWD was measured and frequency was lower within a reach. However, the precision of LWD relative to SWD measures also reflects a bias between survey teams when enumerating SWD.

Pools- While individual pool dimensions were not analyzed, the aggregate parameters, pool frequency (pools/km) and percent pool area, had relatively high S/N and repeatability values. At this scale of analysis (unit reach), pool measurements were found to be repeatable among replicate surveys. There was a directional bias toward SWM staff enumerating more qualifying pools, but this bias was not evident in estimates of pool area.

Surface Fine Sediment- Percent surface fine sediment measurements have low precision and repeatability compared to the other parameters. In reaches with high variability of fines, there was a corresponding high variation between survey teams. Streams with either very low or very high levels of fines seem to have higher repeatability than streams with moderate levels of fines. Not only is the high variability of surface fines within unit reaches and individual transects a barrier to precision in fine sediment sampling, there are factors influencing when and where surveyors sample fines in a reach. These include pool criteria, pool tailout substrate size criterion, transect location within the spawnable area of a pool tailout and transect width due to the exclusion of slack water and dry bars within the pool tailout.

Instability- Bank instability measurements have high precision and repeatability suggesting that survey methods are appropriate for collecting instability measurements. Right and left bank instability was strongly correlated at the subbasin scale; less so at the survey unit reach or geomorphic reach, supporting the null hypothesis presented. Interestingly, right and left bank hydromodification were well correlated at all scales where a single hydromodification was not dominant within the survey unit reach.

RECOMMENDATIONS

Two parameters showed high variability and low precision in repeatability, wood and surface fine sediments. It was determined that some of the variability could be accounted for by additional training throughout the sampling period. Further discussion of training and recommended changes to the methods is described for each parameter below.

BFW – Continue with 2001 protocol.

RWW – Continue with 2001 protocol.

LWD – Revise protocol to provide clear description of channel dimensions and the channel location from which LWD is sampled. Field training of survey teams should emphasize wood measurement location and precision. Precision may be improved by using a logger's diameter tape for diameter measurements whenever possible. Similar analysis (S/N, R, σ_{rep}, and z) should be conducted to estimate the repeatability of stump counts (LWD class 3) and assess the more qualitative measurements of wood decay class, wood type, rootwad presence and logjam association (Snohomish County, 2001b).

SWD – Revise protocol to provide clear description of channel dimensions and the channel location from which SWD is sampled. Field training should emphasize visual SWD classification.

Pools – Continue with 2001 protocol. Field training should emphasize a consistent approach to identify and precisely measure pools, particularly those smaller pools that may barely meet pool criteria.

Fines – Clarify protocol regarding location of surface fines transects. Modify the protocol to ensure a minimum number of riffles are sampled per unit reach. Increase the number of riffles sampled for unit reaches with high surface fines variability.

Instability - Continue with 2001 protocol by measuring both right and left banks. Increase lengths of QC reaches within subbasins to continue to evaluate right bank and left bank instability at a subbasin scale. In the future, it may be most appropriate to predict the extent of hydromodifications as a reflection of human response to bank instability rather than predict the extent of bank instability based on survey data from only one streambank.

GLOSSARY

Geomorphic Reach: Stream reaches in selected sub-basins were delineated according to Level I stream classification methods (Rosgen 1996 type A, B, C, D, DA, E, F, or G) within wadable, fish-bearing waters. Within each geomorphic reach, potentially numerous unit reaches could be surveyed.

Large woody debris (LWD): A piece of wood must be of a length greater than or equal to 7.6 m (length class 1) and a diameter of at least 30 cm (USFS 1999).

Original survey (OR): Data collected during the original stream survey.

Quality control survey (QC): Data collected during the replicate stream survey (quality control).

Repeatability (R): A quotient having a value between 0 to 1 showing the similarity between repeated measurements on the same parameter. R represents the proportion of the variation in the data that could occur among different sampling groups where;

$$R = \text{Variance among groups}/(\text{Variance within groups} + \text{Variance among groups})$$

Rootwad: The root ball portion of woody debris with a mean diameter greater than or equal to one meter.

Signal-to-Noise (S/N): A comparison of the variance of a habitat metric observed across a regional sampling of streams (the “signal”) with the variance resulting from replicate field measurements within the sampling season (the “noise”) where;

$$S/N = \text{Variance of a population} / \sum \text{Variance between replicated pairs}$$

Small woody debris (SWD): All wood that is not classified in the LWD or SSR class, but is at least 10 cm diameter for at least 2 meters length (Pleus et al. 1999).

Unit reach: A survey reach length of 30 bankfull channel widths.

LITERATURE CITED

Johnston, N.T. and P.A. Slaney. 1996. Fish habitat assessment procedures. British Columbia Ministry of Environment, Lands and Parks and British Columbia Ministry of Forests, Watershed Restoration Program, Tech. Circ. No. 8.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger, and D.V. Peck. 1999. Quantifying physical habitat in wadeable streams. EPA/620/R-99/003. U.S. Environmental Protection Agency, Washington, D.C. July 1999.

Krebs, C.J. 1989. Ecological methodology. Harper & Row, Publishers, New York. 654 p. Second edition.

Pleus, A. E., D. Schuett-Hames, and L. Bullchild. 1999. TFW Monitoring Program method manual for the habitat unit survey. Prepared for the Washington State Dept. of Natural Resources under the Timber, Fish, and Wildlife Agreement. TFW-AM9-99-003. DNR #105.

Rosgen, D. L. 1996. Applied River Morphology. Wildland Hydrology Books, Pagosa Springs, CO.

Snohomish County. 2000. Physical Habitat survey and monitoring protocol for wadable streams. Version 5.0. Snohomish County, Public Works Department, Surface Water Management Division. Everett, Washington.

Snohomish County. 2001a. Wadable Stream Habitat survey preliminary results for the year 2000. Snohomish County, Public Works Department, Surface Water Management Division. Everett, Washington. Also available at: http:www.co.Snohomish.wa.us/publicwk/swm/reports/index.htm

Snohomish County. 2001b. Physical Habitat Survey and Monitoring Protocol for wadable streams. Version 6.0. Snohomish County, Public Works Department, Surface Water Management Division. Everett, Washington. Also available at: http:www.co.Snohomish.wa.us/publicwk/swm/reports/index.htm

Statsoft Incorporated. 1999. STATISTICA [data analysis software system], version 5.5.

United States Forest Service. 1999. *Stream inventory handbook: Level I & II*. USDA, Region 6, Version 9.9.

Zar, J.H. 1984. Biostatistical Analysis. Prentice-Hall, Englewood Cliffs, N.J. Second edition.