NRC Research Press

ARTICLE

The spatial distribution of salmon and steelhead redds and optimal sampling design

Martin C. Liermann, Dan Rawding, George R. Pess, and Bryce Glaser

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Abstract: Redd surveys are used extensively to estimate spawner population size for Pacific salmon (*Onchorynchus* spp.). Because redds tend to be spatially aggregated, estimates of total redds based on subsamples of the potential spawning grounds can be uncertain unless the spatial structure is accounted for. Here we use known redd locations for three populations over several years to compare five different probability sampling designs through simulation. The coefficient of variation (CV) for estimates based on simple random sampling was high, with values well over 15% when sampling a third of the reaches. Moving to a spatially balanced sampling design (generalized random tessellation stratified; GRTS) produced improvements in two of the three watersheds (16%–22% reduction in CV). Estimates based on a stratified GRTS design and a GRTS design that included a census of all reaches close to the peak count had higher accuracy, with an approximate CV of one-half to one-third of GRTS alone. We show how these improvements are predicted by theory and under which conditions the different approaches are likely to perform well.

Résumé : Les dénombrements de nids de frai sont largement utilisés pour estimer la taille de populations de géniteurs de saumons du Pacifique (*Onchorynchus* spp.). Étant donné que les nids de frai sont souvent regroupés dans l'espace, les estimations de leur nombre total reposant sur des sous-échantillons des lieux de frai potentiels peuvent s'avérer incertaines si la structure spatiale n'est pas prise en considération. Nous utilisons des emplacements connus de nids de frai pour trois populations sur plusieurs années afin de comparer par simulation cinq plans d'échantillonnage aléatoire différents. Les coefficients de variation (CV) pour les estimations basées sur un échantillonnage aléatoire simple étaient élevés, les valeurs dépassant de beaucoup 15 % quand un tiers des tronçons était échantillonné. L'utilisation d'un plan d'échantillonnage spatialement équilibré (stratifié par tessellation aléatoire généralisée; GRTS) a produit des améliorations pour deux des trois bassins versants (réduction de 16 % à 22 % des CV). Les estimations basées sur un plan GRTS stratifié et un plan GRTS qui inclue le recensement de tous les tronçons situés à proximité du maximum de nids dénombrés présentaient une plus grande exactitude, les CV approximatifs allant de la moitié au tiers du CV pour le GRTS seul. Nous démontrons que ces améliorations sont prédites par la théorie et illustrons dans quelles conditions ces différentes approches sont susceptibles de donner de bons résultats. [Traduit par la Rédaction]

Introduction

Estimates of adult spawning escapement are essential for monitoring salmon population status and trends, developing biological reference points for fisheries, and evaluating the effectiveness of hatcheries and restoration actions (McElhany et al. 2000). Redd surveys (Gallagher et al. 2007) are one of the most frequently used methods to estimate escapement for salmon and steelhead populations in the Pacific Northwest (Washington Department of Fish and Wildlife 2002). In a typical redd survey, a portion of the spawning grounds are surveyed, and the resulting value is expanded to account for the unsurveyed spawning grounds and the number of adults per redd (Gallagher et al. 2007). Historically, the selection of survey reaches was often opportunistic, based on access and knowledge of spawning concentrations (e.g., Irvine et al. 1992; Hassemer 1993; Jacobs and Nickelson 1998). While this approach tended to produce indices that were effective at measuring temporal trends, the relationship with the total population size was uncertain, since the portion of the population represented by the surveyed sites was generally not known.

In the last few decades, sampling designs based on probability sampling have gained traction in fisheries surveys (e.g., Stevens et al. 2007), providing unbiased estimators of population size. While probability sampling can guarantee unbiased estimates, variance can be high owing to the patchy distributions common in salmonid spawning locations (e.g., Irvine et al. 1992; Isaak and Thurow 2006). Stratified designs that concentrate sampling effort in areas with higher redd density and spatially balanced designs that ensure an even spatial distribution of sample reaches have been shown to substantially reduce estimator variance (Courbois et al. 2008; Gallagher et al. 2010; Jacobs and Nickelson 1998; Jacobs et al. 2009). These results align with general sampling theory that predicts increased accuracy for designs that sample units with higher predicted variance more intensively (e.g., Cochran 1977) and constrain samples to be representative with respect to auxiliary variables that are related to the variable of interest (e.g., Tillé 2011).

Accuracy can also be improved through model-assisted estimators (Särndal et al. 1992), where the within-sample relationship between the variable of interest (e.g., redd density) and an auxiliary variable available for all reaches (e.g., reach gradient) is used to extrapolate to the larger population. For example, with a regression estimator (Cochran 1977), redd density in the unsampled reaches is estimated based on the estimated linear relationship

Received 10 April 2014. Accepted 20 October 2014.

Paper handled by Associate Editor Josef Michael Jech.

M.C. Liermann and G.R. Pess. Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 2725 Montlake Boulevard East, Seattle, WA 98112, USA.
D. Rawding and B. Glaser. Washington Department of Fish and Wildlife, 600 Capitol Way N., Olympia, WA 98501-1091, USA.
Corresponding author: Martin Liermann (e-mail: martin.liermann@noaa.gov).

Can. J. Fish. Aquat. Sci. **72**: 434–446 (2015) dx.doi.org/10.1139/cjfas-2014-0181

Published at www.nrcresearchpress.com/cjfas on 11 November 2014.

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

between redd density and reach gradient in the sampled reaches. While inference for design-based estimators is based solely on the design and sample, model-assisted estimators rely on assumptions (e.g., linearity for the regression estimator), which can result in bias.

There is considerable guidance available for those constructing redd sampling designs (e.g., Gallagher et al. 2007; Stevens et al. 2007). However, most studies examining redd sampling designs are limited to a single species or population, and focus on one or two approaches to design and estimation. Here we investigate the relative performance of five different approaches for estimating total redds. These include standard design-based estimators applied to simple random, spatially balanced, and stratified sampling as well as regression (model-assisted) estimators based on a census of the entire basin conducted close to the peak occurrence of redds. Comparison of the estimators is based on simulations using known redd locations for three populations representing two species, each with multiple years of data. We illustrate how the performance of particular approaches can be anticipated using sampling theory and knowledge of the spatial distribution of redds, and develop this into guidance for those implementing redd sampling designs in the many small Pacific Northwest basins.

Methods

Study site and species

The populations included two fall Chinook (*Oncorhynchus tshawytscha*) populations in the Coweeman River (CWM) and the East Fork Lewis River (EFL), and one winter steelhead (*Oncorhynchus mykiss*) population in Mill, Abernathy, and Germany (MAG) creeks (Fig. 1). The CWM is a tributary to the Cowlitz River and drains approximately 329 km^2. Sexually mature Chinook salmon enter the watershed from late August through late October, with spawning occurring shortly thereafter. Spawning has been observed from river kilometre (rkm) 9 up to the Washboard falls (rkm 50), a barrier to migratory salmon (Rawding et al. 2014). There is considerable annual variation in the upstream extent of the spawning distribution, which is likely influenced by discharge and spawner density (S. VanderPloeg, Washington Department of Fish and Wildlife, Olympia, Washington, personal communication, 2013). The EFL is located approximately 24 km north of Vancouver, Washington, and enters the Lewis River at rkm 5.6. It drains 543 km^2, and Chinook spawning is limited to the main stem from the start of suitable substrate at rkm 9.4 to Lucia Falls (rkm 34) (Steel et al. 2009). Chinook salmon entry is similar to that of the CWM but spawning timing is slightly later. Although there are no hatchery Chinook salmon releases in these two watersheds, approximately 25% of the spawners are of hatchery origin based on the absence of the adipose fin (Rawding et al. 2014).

The MAG creeks drain approximately 75, 75, and 58 km^2, respectively. They flow into the Columbia River west of Longview, Washington, at rkm 86.6, 87.3, and 90.5. Adult winter steelhead return to Lower Columbia River tributaries from November through May, with peak spawning occurring in April (Leider et al. 1984). The primary spawning locations include the main stems of the MAG creeks, with limited spawning occurring in their tributaries (Fig. 1). The Abernathy Fish Technology Center (AFTC), located at rkm 7, has operated a local winter steelhead broodstock program as part of a relative reproductive success study (USFWS 2006). An electric weir and a trap in the fish ladder, adjacent to AFTC, are used to collect broodstock. A mixture of hatchery and natural origin adults are released above AFTC based on the study design. The proportion of hatchery steelhead spawners below AFTC may exceed 50% in some years (Kinsel et al. 2009).

Data

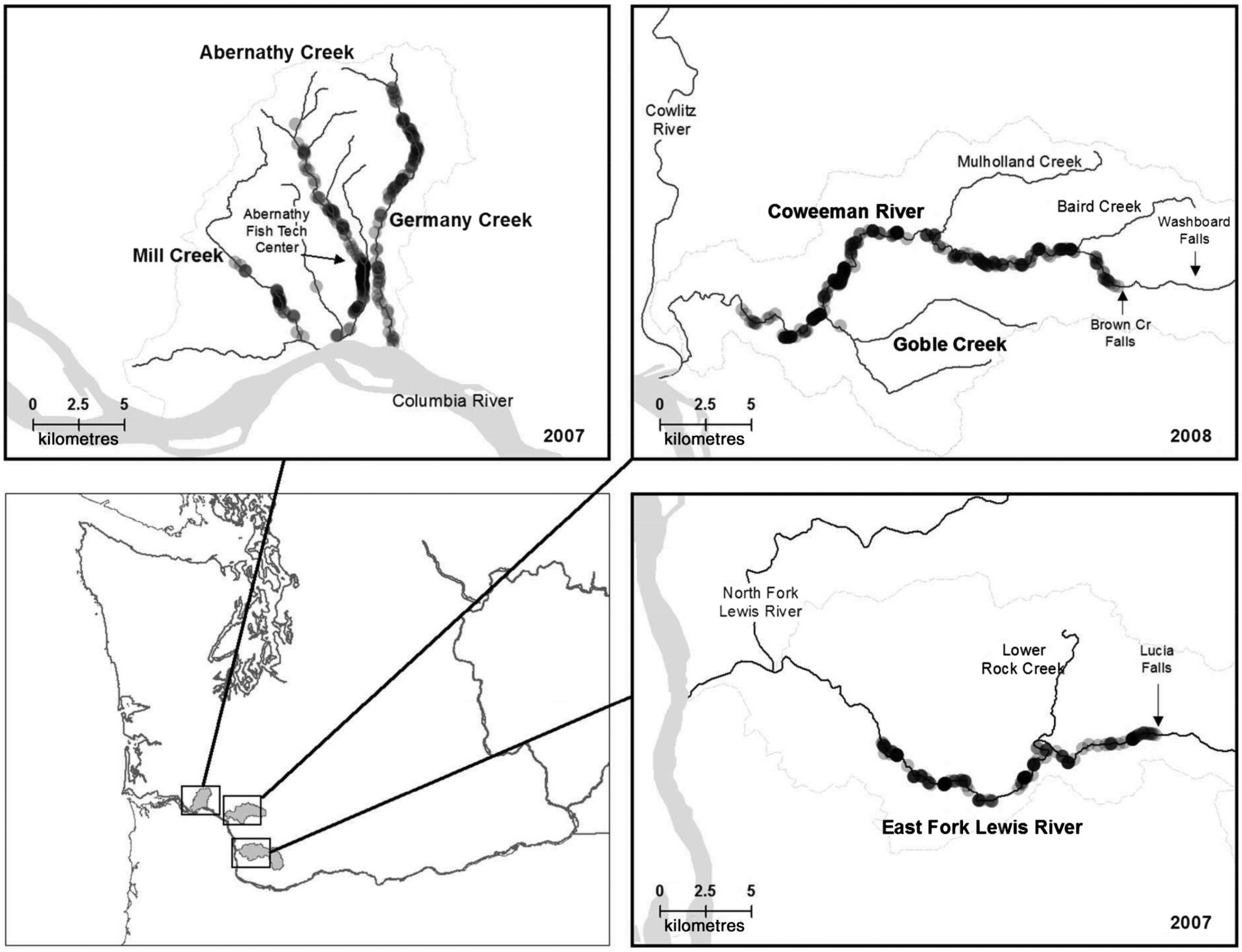
Redd location data for the three populations were collected from 2003 to 2011 (Fig. 1, Fig. 2). Redd surveys for both species followed the American Fisheries Society salmonid monitoring protocols (Gallagher et al. 2007), covered the entire spawning period and distribution, and were considered censuses. Redd surveys occurred weekly from mid-September through early November for fall Chinook salmon, and every 2 weeks from early March through late May for winter steelhead. During each foot survey, newly identified redds were flagged and georeferenced using recreational-grade global positioning system (GPS) units (Courbois et al. 2008; Jacobs et al. 2009). The reported accuracies from the GPS units generally ranged from 2 to 15 m. In subsequent surveys, previously flagged redds were inspected to determine whether they should be classified as “still visible” or “not visible”. A “still visible” redd would have been observed and identified without the flagging present. A “not visible” redd did not meet this criterion. This methodology allowed us to track the new redds and all visible redds in each survey. The individual redd coordinates were assembled into a geographic information system (GIS) database (Isaak and Thurow 2006; Jacobs et al. 2009). GPS points were moved to the closest point on the 1 : 24 000 stream layer, except in cases where the coordinate was clearly in error owing to poor satellite coverage. For these readings the points were placed at the midpoint along the stream layer between the previous and next redd. While observer error when enumerating redds can be a concern (Gallagher et al. 2007), it should be low for these populations since (*i*) observers were well trained and used standard protocols, (*ii*) superimposition is likely low because redd densities are low except downstream of the Abernathy weir, (*iii*) surveys on these small rivers and streams during the early fall (Chinook) and spring (steelhead) are conducted under low to moderate stream flow with good visibility, and (*iv*) redds were flagged and resighted across multiple visits. During 2010 and 2011, portions of Germany creek were not surveyed owing to problems with access (in 2010, rkm 6.43–9.41 was only surveyed on two of the six sampling events, while in 2011, rkm 8.0–9.41 was not surveyed at all). To correct for this we added redds (with random locations) to these segments to achieve relative densities comparable to those observed in other years when the segments were surveyed. The added redds constituted <10% of the population in both years.

The raw redd data was then collated into reach counts by breaking the geographic spawning range into evenly spaced 1 km reaches and tallying the number of redds in each reach. Reach boundaries were defined by starting at the lower extent of the spawning range within each stream and adding reaches going upstream until the upper extent of spawning was reached. The upper bound was adjusted up so that the total distance within each stream was rounded up to the nearest kilometre. These reaches in total then define the sampling frame (i.e., the population of reaches from which subsamples are selected). For the sampling event that coincided most closely with the peak of observable redds (for each population and year), the redd counts for all reaches in the sampling frame (a peak count census) was used as an auxiliary variable for the regression estimators and to construct the stratified designs. In practice, the date of peak count will have to be predicted based on data from previous years and the trajectory of counts from the current year.

Upon the completion of this process we have counts of total redds and redds at peak count for reaches constituting the total spawning areas for each population and year of data. We can therefore evaluate the different designs and estimators by comparing the true number of total observed redds to estimates based on subsamples of the reaches (as outlined below in “Assessing the performance of sampling designs”).

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Fig. 1. Redd locations for the Mill, Abernathy, and Germany complex (MAG) steelhead population in 2007 (upper left), the Coweeman (CWM) Chinook population in 2008 (upper right), and the East Fork Lewis (EFL) Chinook population in 2007 (lower right). A map locating the three populations in the Pacific Northwest (lower left). Redd locations are indicated with semitransparent gray dots, where dots falling on top of each other results in darker areas and indicates higher redd densities.



Characterizing the distribution of redd counts

We characterized the spatial distribution of redds using plots of redd location and reach counts, and two metrics based on the reach counts: (1) dispersion, which is the variance divided by the mean (Perry and Mead 1979), and (2) the lag-1 spatial autocorrelation (i.e., correlation of adjacent reach counts). We estimated spatial autocorrelation for the MAG population (which included several streams) by placing the streams end to end and ignoring their relative spatial positions in the network. While there are more rigorous methods for characterizing spatial structure on a network (e.g., Gresswell et al. 2006), the focus of this paper was not to characterize the spatial distribution, but to provide a rough illustration of its effects on different estimators. If each redd is placed independently and randomly on the stream network (i.e., all locations have equal probability), then redd counts will be Poisson distributed with expected dispersion of one and expected lag-1 spatial autocorrelation of zero. If instead redds are more likely to occur in certain locations and (or) the presence of a redd attracts others, then dispersion is likely to be larger. If the scale of the resulting spatial pattern of redds is large relative to the reach size, then spatial autocorrelation will likely also increase.

Sampling methods

We assessed five different sampling approaches (Table 1):

(*i*) Simple random sampling (SR).

(*ii*) Generalized random tessellation stratified (GRTS) sampling. GRTS shares the advantages of spatial systematic sampling (spatial balance) while allowing any two units to occur in the same sample (which makes the derivation of a design-based variance estimator possible) and providing a straightforward approach to adding or deleting units from an existing sample. This is accomplished by randomly placing a grid over the sample space, developing a list of randomized hierarchical addresses for each of the cells, and systematically sampling from these addresses.

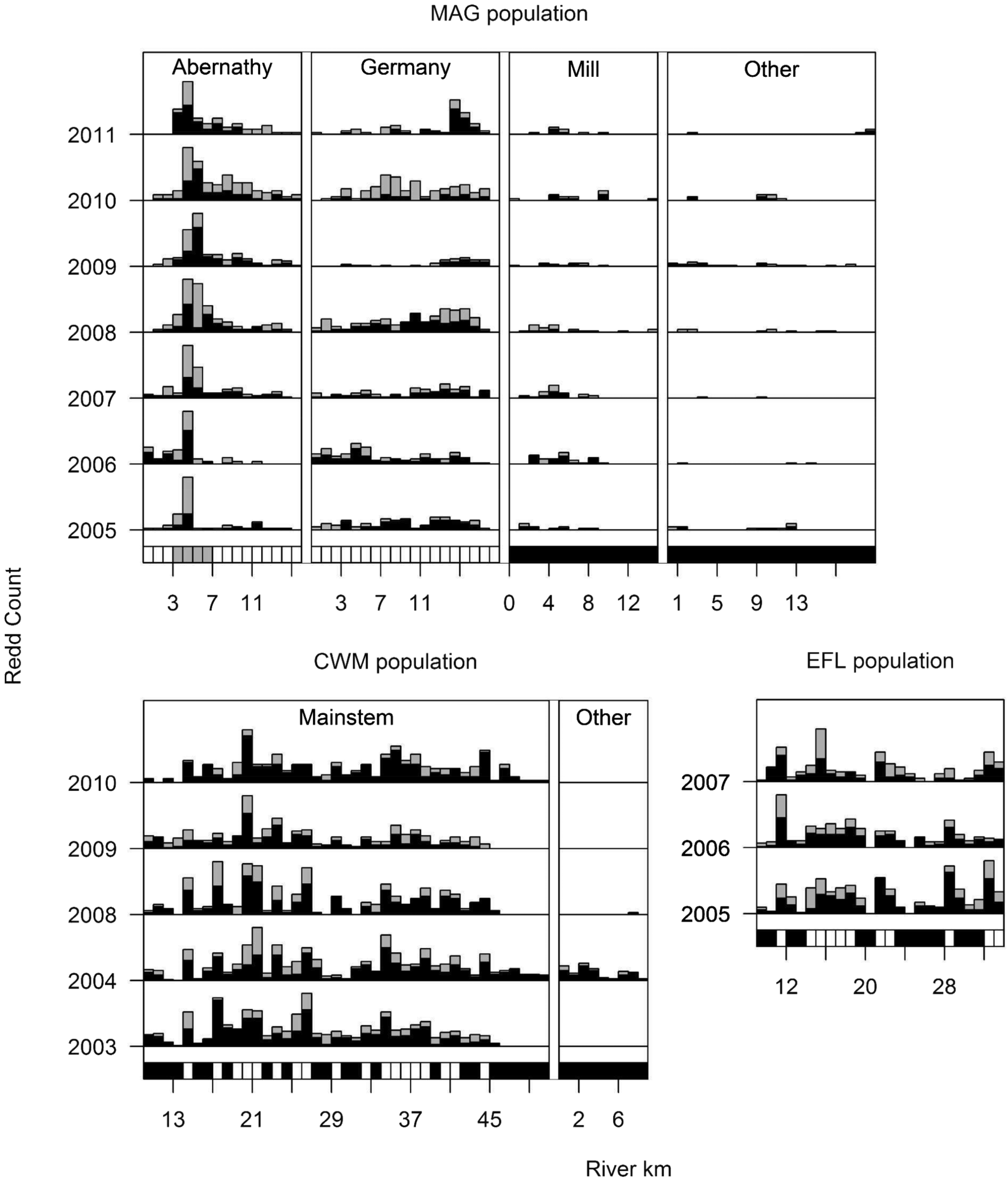
(*iii*) Stratified GRTS sampling (S-GRTS), where the sampling frame (population of reaches) is broken into groups (strata) based on geography or site characteristics and a separate GRTS sample is taken from each stratum.

(*iv*) A GRTS sample augmented with a census (all reaches sampled) during the estimated peak occurrence of redds (C-GRTS).

(*v*) A stratified GRTS sample (S-GRTS) augmented with a census (CS-GRTS).

For the MAG population we broke the sample frame into three geographic strata based on knowledge of the spawning distribution derived from two decades of surveys. These strata were reaches 4 through 7 in Abernathy Creek (rkm 3–7 associated with a concentration of redds below the hatchery weir), the rest of Abernathy Creek and Germany Creek combined, and everything else. Strata in the EFL and CWM populations were constructed

Fig. 2. Top panel: Steelhead redds per 1 km reach across years in the MAG population, where streams have been placed end to end. Bottom-left panel: Chinook redds per 1 km reach for the Coweeman River. Bottom-right panel: Chinook redds per 1 km reach across years for the EFL river. Bar heights are redds per reach as a proportion of the maximum for that year and population. The black portions of the bars represent the peak redd counts. Redds per reach were divided by the largest value for each year to emphasize the spatial distribution. The "Other" label at the top of the MAG and CWM figures represents a composite of minor streams. Reach membership in the different strata is indicated by the checkered strip at the bottom of each panel. For the CWM and EFL populations, white = high and black = low. For the MAG population, gray = weir, white = Germany and Abernathy, and black = other. Populations totals are 169, 245, 231, 325, 245, 248, and 169 for MAG; 601, 823, 289, 371, and 298 for CWM; and 366, 292, and 203 for EFL (all ordered by increasing year).



using the cumulative root frequency rule of Dalenius and Hodges (1959), where the stratifying variable was the normalized peak counts (all counts divided by the year total) averaged over years. The cumulative root frequency rule involves breaking the stratifying variable into many small bins and then finding contiguous groupings of these bins (into two strata in our case) such that that the sums of the square root of the bin frequencies are the same for each stratum. Technically, only peak count data from previous

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Table 1. A description of sampling designs, including the estimator for the population total, variance of the estimator, and references.

	Sampling method description						
Abbreviation	Simple random	GRTS	Stratified	Census	Estimator	Variance	Reference
SR	×						
GRTS		×			Standard	Neighborhood	Stevens and Olsen 2004
S-GRTS		×	×		Standard	Neighborhood	Stevens and Olsen 2004
C-GRTS		×		×	Regression	Standard	Cochran 1977
CS-GRTS		×	×	×	Regression	Standard	Cochran 1977

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

years would be available for constructing strata. However, this would produce different strata for the different years in this exercise, complicating interpretation and implementation. The effect of this simplification is likely to be small relative to the year-to-year and population-to-population variability. For all three populations the total number of samples was divided among the strata based on Neyman allocation (e.g., Cochran 1977), where stratum sample size is proportional to the stratum-specific normalized peak count standard error. To account for multiple years, we estimated the proportions allocated to the different strata for each year separately and then averaged across years and normalized. We assumed constant cost for all reaches, and GRTS sampling was used within all strata.

We applied three sampling fractions (1/5, 1/3, and 1/2) to all populations and approaches (Table 1) and four different reach lengths (0.25, 0.5, 1.0, and 2.0 km), with sampling fraction equal to one-third. The number of reaches sampled is therefore inversely proportional to the reach size. So, for example, there are 4 times as many 0.25 km reaches as 1.0 km reaches (since the sampling fraction is constant). Because the maximum reach size was 2 km, we expanded the population sample frames along each stream to be divisible by 2 km for the reach size analysis.

We use the terms balanced and spatially balanced throughout the manuscript to describe samples and sample designs. A spatially balanced sample (Stevens and Olsen 2004) is one that is evenly dispersed in space (i.e., along the stream network in this case). A sample is balanced relative to an auxiliary variable if for that variable the estimated mean based on the sample is equal to the population mean (Tillé 2011). A design is balanced or spatially balanced if all the samples it produces are balanced or spatially balanced. Both balance and spatial balance formalize the concept of a sample that is representative with respect to a variable. In practice, exact balance is rarely achieved.

Estimators

We use the standard estimators and their variances for SR, GRTS, and S-GRTS designs (Table 1). The C-GRTS and CS-GRTS designs include the standard sample of reaches which are visited multiple times during the spawning season (as is the case for all other sampling approaches), and a single visit to all reaches in the sampling frame (a census) close to the peak occurrence of redds. For these approaches we use the regression estimator:

$$\hat{T} = N[\bar{y} + \hat{b}(\bar{X} - \bar{x})]$$

where the estimate of the total population, $\hat{T}$, is the total number of reaches, N, times the mean number of redds per reach in the sample, $\bar{y}$, corrected by the difference between the mean of peak counts for the sample, $\bar{x}$, and all reaches, $\bar{X}$, multiplied by the estimated regression slope, $\hat{b}$ (Cochran 1977). Here, $\hat{b}$ is the standard least squares estimate of the slope based on the relationship between the peak count, x, and total redds, y, for the reaches in the sample. If the sample is representative of the larger population with respect to peak count (i.e., approximately balanced with respect to peak count), then $\bar{X} - \bar{x}$ will be close to zero and the estimate, $\hat{T}$, will be very close to the simple expansion estimator, $N\bar{y}$. However, for samples with peak counts that poorly represent the whole population (i.e., unbalanced with respect to peak count) the regression estimator will make a large correction to the standard expansion estimator (for example, when using SR sampling). Thus, for sample designs that produce approximate balance with respect to peak count (e.g., by stratifying on peak count) the regression estimator will tend to be less effective. Another way of thinking about the regression estimator is that the counts for reaches not in the sample are estimated using the estimated regression relationship between peak and total counts within the sample. Therefore, the precision of the regression estimator depends on the quality of this relationship (and not, for example, the degree of spatial autocorrelation). We did not calculate separate regression estimators for the different strata because sample sizes were small. The stratified regression estimator is therefore the same as above, with $\bar{x}$ and $\bar{y}$ replaced with their stratified estimators. Another common estimator, the ratio estimator, was not considered, since the relationship between x and y is not expected to go through zero (since $y_i \geq 0$ when $x_i = 0$). For the GRTS and S-GRTS designs we used the neighborhood variance estimator (Stevens and Olsen 2003), which uses the spatial proximity of points to produce variance estimates that are often smaller (and less biased) than the simple random sampling estimator. The neighborhood variance estimator is not currently available for regression estimators, so the variance estimates for the C-GRTS and CS-GRTS designs may be conservative. For all variance estimators we applied the finite population correction. Standard 95% confidence intervals based on the t distribution (e.g., Cochran 1977) were constructed for all estimates.

Assessing the performance of sampling designs

We assessed the relative performance of the different sampling approaches and estimators by sampling repeatedly from the known distribution of redds for each of the populations using the following steps:

(*i*) Select a sample from the population of reaches (sampling frame).

(*ii*) Estimate the total population size and standard error based on the sample.

(*iii*) Repeat steps 1 and 2 a total of 1000 times for each combination of population, year, and design.

(*iv*) Assess the performance of the sampling designs and estimators by comparing them to the true values (the known redd abundance for that year and population).

The estimators are compared with the known true values using the estimated mean squared error (denoted from here forward as MSE).

$$\mathrm{MSE}(\hat{T}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{T}_i - T)^2$$

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Table 2. Metrics describing the distribution of reach counts for the three populations.

Population statistic	Mill, Abernathy, and Germany creeks Mean	Min.	Max.	Coweeman River Mean	Min.	Max.	East Fork Lewis River Mean	Min.	Max.
Sum redds	233	169	325	476	289	823	287	203	366
Mean	3.28	2.38	4.58	9.53	5.78	16.46	11.48	8.12	14.64
SD	5.74	4.40	7.28	9.29	5.68	13.03	9.37	7.49	11.18
Variance/mean	10.23	7.00	15.34	9.30	5.42	12.49	7.70	6.91	8.53
Skewness	3.74	2.33	4.98	1.41	1.11	2.11	1.27	0.62	1.65
AutoCorr	0.55	0.31	0.78	0.32	0.16	0.45	0.03	–0.05	0.10
r	0.90	0.82	0.97	0.96	0.94	0.97	0.94	0.92	0.97
$\sqrt{1-r^2}$	0.40	0.25	0.57	0.29	0.25	0.35	0.35	0.24	0.40

Note: Mean, max. (maximum), and min. (minimum) values are for the multiple years. Reach size is 1 km. Sum redds is the sum of all redd counts, SD is the standard deviation, r is the correlation coefficient between the peak and total counts, AutoCorr is the lag-1 spatial autocorrelation, and $\sqrt{1-r^2}$ is the expected reduction in the coefficient of variation (CV) when applying the regression estimator.

Here n is the number of estimates (1000), $\hat{T}_i$ is the ith estimate, and T is the true value. To simplify between-population comparisons, we converted MSE to the coefficient of variation of root mean squared error (CVRMSE) by dividing the square root of MSE by the true value (i.e., total redds) to standardize.

$$\mathrm{CVRMSE}(\hat{T}) = \frac{\sqrt{\mathrm{MSE}(\hat{T})}}{T}$$

CVRMSE is the measure of accuracy used in this study. Bias was estimated as the mean difference between the estimator and the true value.

$$\mathrm{Bias}(\hat{T}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{T}_i - T)$$

For the design-based estimators applied to the SR, GRTS, and S-GRTS approaches (Table 1), the bias is zero, since these designs are probability sampling designs with known inclusion probabilities and the estimators do not rely on additional assumptions. Therefore, we estimated bias only for the model-based regression estimator used for the designs with a census (C-GRTS, CS-GRTS). Because MSE is the sum of the variance and squared bias of the estimator, when bias is zero MSE = variance and CVRMSE = CV. This allowed us to use CVRMSE to assess the performance of the estimators relative to the 15% CV standard proposed for salmon escapement estimates (Crawford and Rumsey 2011) when bias was small or zero. However, estimated spawning escapement includes additional sources of error, namely errors in identifying redds and uncertainty in estimating the number of adults per redd. Thus, a CV of 15% for the estimate of total redds is a necessary but not sufficient condition for achieving a CV of 15% for spawning escapement. The performance of the confidence interval around the estimate is assessed by calculating the true coverage of the 95% confidence interval. This is achieved by recording the percentage of times the confidence interval derived from the estimate includes the true value.

Sampling designs that include a peak count census (i.e., the C-GRTS and CS-GRTS) require additional sampling effort that should be accounted for when comparing the estimators. For a stream with n reaches, where a fraction of the reaches (f, the sampling fraction) is sampled k times, the total number of visits, or sampling effort, is $k \times f \times n$. For designs with a peak count census, the extra effort of the census can be accounted for by reducing the sampling fraction (f') so that the total number of visits is the same (i.e., $kfn = (k-1)f'n + n$, and therefore $f' = (kf-1)/(k-1)$). For example, if $k = 6$ and the original sampling fraction is 1/2 then the sampling fraction for the design with a peak count that results in the same number of reach visits would be $(6 \times 1/2 - 1)/(6 - 1) = 2/5$. Similarly a sampling fraction of 1/3 would be reduced to 1/5.

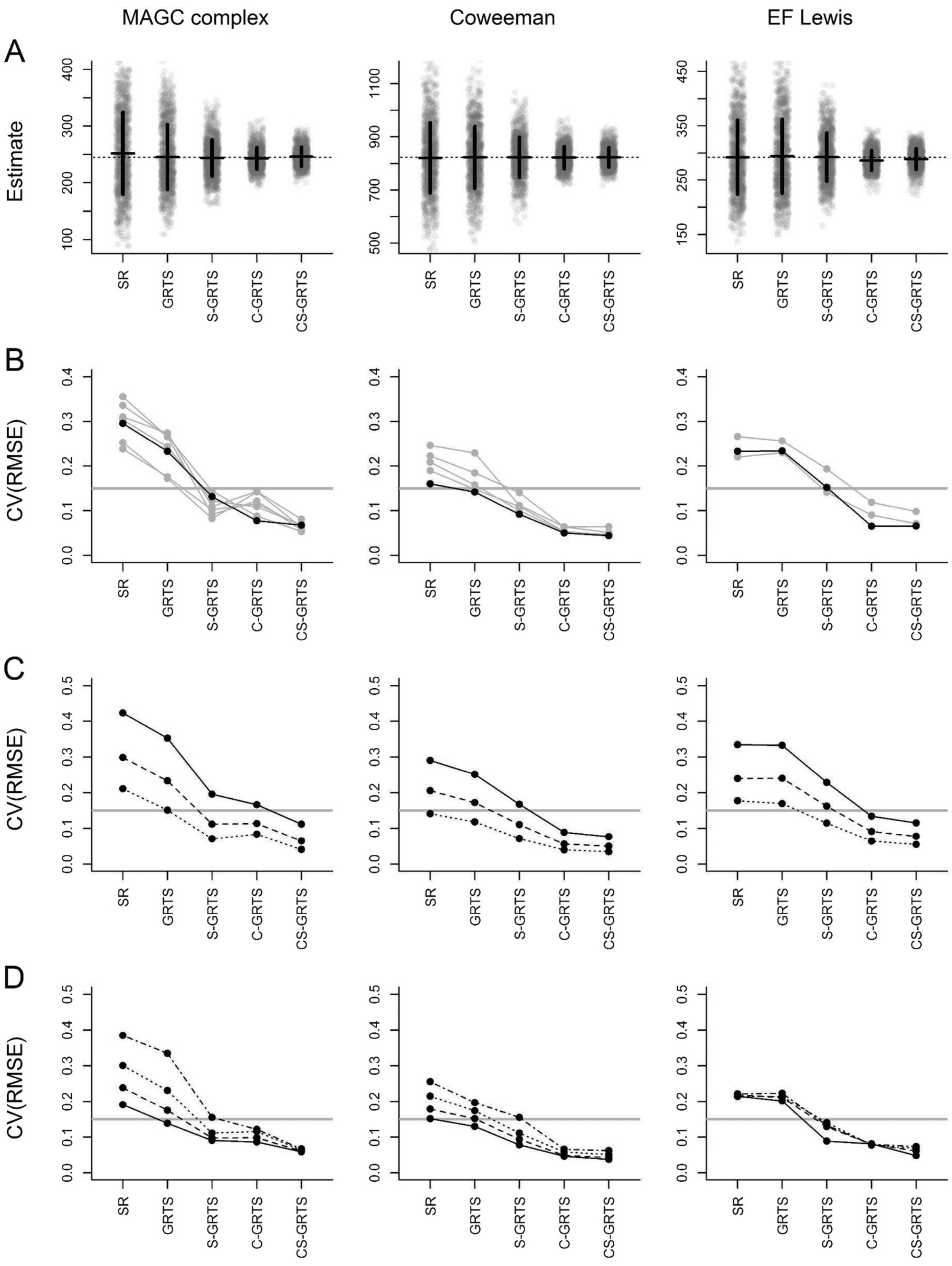
All analyses were carried out in R (R Development Core Team 2012). The R package spsurvey (Kincaid and Olsen 2012) was used to draw samples for all GRTS-based designs, and to provide the corresponding population estimates, standard errors, and confidence intervals for designs without a peak count census. The regression estimates and variances were calculated using the R package survey (Lumley 2012), and the strata boundaries for the CWM and EFL populations were defined using the R package stratification (Baillargeon and Rivest 2011).

Results

Geographic patterns were evident in the distribution of redds for the MAG, CWM, and EFL populations (Fig. 1). When the redds were aggregated into 1 km reaches the uneven distribution was still evident, with reaches closer together tending to have similar counts (Fig. 2). The unevenness occurred in all years, and the spatial pattern was somewhat consistent from year to year (Fig. 2). For example, a concentration of redds associated with a hatchery weir was visible for all years between rkm 2 and 7 on Abernathy Creek in the MAG population. These spatial patterns resulted in highly skewed redd counts, with mean variance to mean ratios of 10.2, 9.3, and 7.7 for the MAG, CWM, and EFL populations, respectively (Table 2). If redds were located at random (i.e., every location had equal probability) the expected variance to mean ratio would be 1. For simple random sampling these large variances translated into proportionally larger variances in the population estimators. The estimated spatial autocorrelation (averaged over years) for the MAG and CWM reaches was 0.55 and 0.32, respectively, while for the EFL population the spatial autocorrelation was 0.03 (Table 2).

Applying the different sampling approaches and estimators to the MAG, CWM, and EFL populations produced more than threefold differences in CVRMSE (Fig. 3B; Table 3). CVRMSE for SR sampling was high (Fig. 3B; Table 3) in spite of the high sampling fractions. As expected (given the population-specific spatial autocorrelation values), moving from SR sampling to the spatially balanced GRTS method reduced CVRMSE for the MAG and CWM populations (78% and 84% of SR, respectively) while having little impact on the EFL population (100% of SR) (Table 3, sample fraction = 1/3). Moving from GRTS to stratified GRTS sampling (S-GRTS) resulted in declines in CVRMSE of a third for CWM and EFL populations (65% and 67% of GRTS for CWM and EFL, respectively), while CVRMSE more than halved (49% of GRTS) for MAG. For the MAG population the regression estimator applied to GRTS samples with a census (C-GRTS) performed comparably to the stratified estimator (51% of GRTS). However, for the CWM and EFL populations the regression estimator CVRMSE was almost half

Fig. 3. (A) Estimators for the years indicated by the bold line in panel B based on 1000 simulations jittered slightly on the *x* axis to better show the distribution. The vertical bars are the standard deviations, and the short solid horizontal bars are the means. The dotted horizontal bar is the true population size. (B) The root mean squared error divided by the population total, CV(RMSE), for the five different approaches using a sampling fraction of 1/3 with a reach length of 1 km. Each line represents a different year. The bold line represents 2006, 2004, and 2006 for the MAG, CWM, and EFL populations, respectively. (C) CV(RMSE) for different sample fractions, averaged across years. Solid line = 1/5; dashed line = 1/3; and dotted line = 1/2. (D) CV(RMSE) for four different reach sizes, averaged across years with sample fraction of 1/3. Solid line = 0.25 km; dashed line = 0.5 km; dotted line = 1 km; dash-dotted line = 2 km. SR = simple random sampling; GRTS = generalized random tessellation stratified sampling; S-GRTS = stratified GRTS; C-GRTS = regression estimator with GRTS; CS-GRTS = stratified GRTS with regression estimator.



Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Table 3. Root mean squared error (%) and bias (%) for the different estimators for sampling 1/5, 1/3, and 1/2 of the total reaches.

	Mill, Abernathy, and Germany creeks			Coweeman River			East Fork Lewis River		
Sampling approach	1/5	1/3	1/2	1/5	1/3	1/2	1/5	1/3	1/2
(*a*) Root mean squared error (%) for sampling 1/5, 1/3, and 1/2 of the total reaches.									
GRTS									
Mean	83	78	71	87	84	84	100	100	95
Min.	74	68	55	80	76	78	98	96	94
Max.	93	88	85	94	93	94	101	104	96
S-GRTS									
Mean	56	49	48	67	65	61	68	67	68
Min.	42	34	34	53	49	46	61	62	66
Max.	64	58	60	77	76	73	72	75	69
C-GRTS									
Mean	50	51	59	36	33	34	40	38	38
Min.	29	33	33	31	28	31	28	28	28
Max.	97	83	97	40	36	36	48	46	45
CS-GRTS									
Mean	33	29	28	31	30	29	35	32	33
Min.	22	20	19	23	22	24	27	28	32
Max.	44	37	42	35	35	37	41	38	33
(*b*) Bias (%) for sampling 1/5, 1/3, and 1/2 of the total reaches.									
C-GRTS									
Mean	2.7	1.7	1	0.8	0.3	0.2	2	1.3	0.7
Min.	0.1	0.5	0	0.2	0.1	0.1	0.2	0.3	0.4
Max.	6	3.3	2.2	1.5	0.6	0.4	3.2	1.9	1.1
CS-GRTS									
Mean	0.7	0.3	0.1	0.4	0.2	0.1	1.2	0.4	0.2
Min.	0.3	0.1	0	0.1	0	0.1	0.4	0	0
Max.	1.1	0.7	0.4	0.7	0.5	0.2	1.9	1	0.4

Note: Root mean squared error (%) for the generalized random tessellation stratified (GRTS) results is reported as a percentage of root mean squared error (RMSE) for simple random sampling and as a percentage of RMSE for GRTS for the three other designs. The absolute value of bias is reported as a percentage of the mean for the two regression estimators. The other estimators are guaranteed to be unbiased (design unbiased).

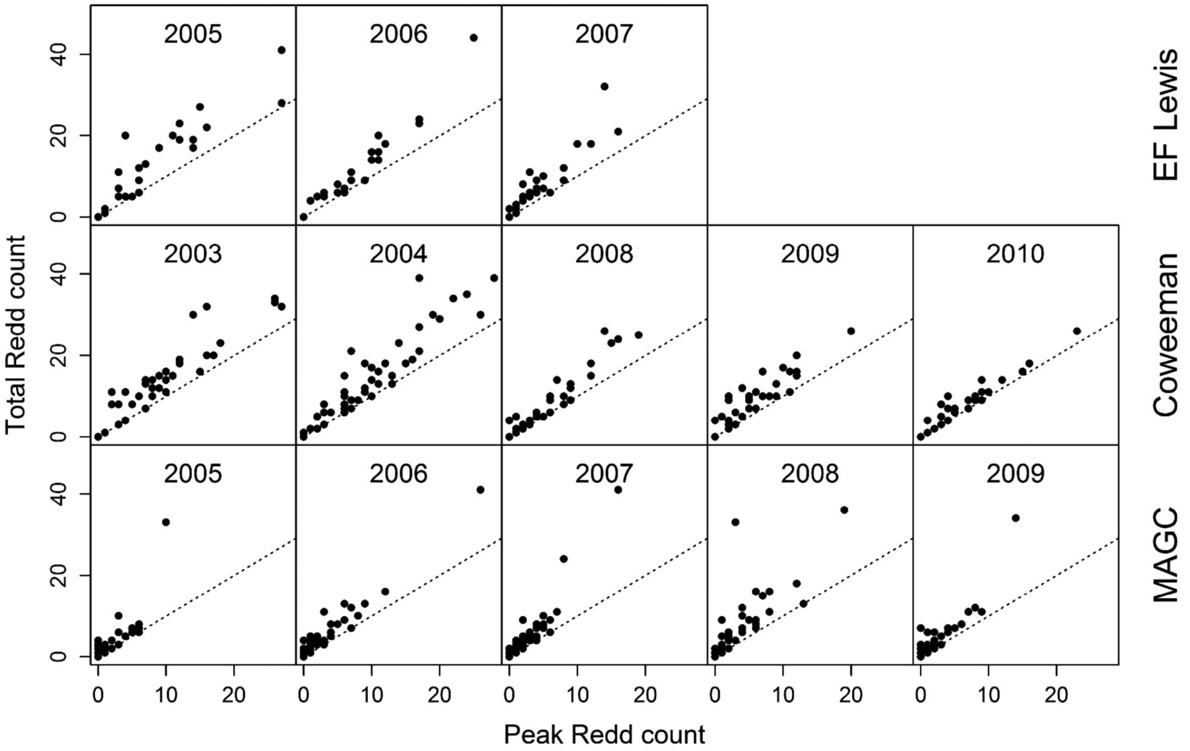
that of the stratified estimator (33% versus 65% and 38% versus 67% of GRTS, respectively). The performance of the regression estimator can be explained by the high correlation and approximate linearity of the relationships between peak census counts and total counts for the different populations and years (Fig. 4; Table 2). An illustration of how CVRMSE of the regression estimator is affected by the quality of the relationship between peak and total counts was evident in the MAG population for 2006, where the exceptionally good relationship between peak and total counts (Fig. 4) resulted in the year with lowest CVRMSE for the regression estimator (Fig. 3B, bold line). Finally, stratifying improved the regression estimator performance for the MAG population (29% versus 51% of GRTS), but had little effect for the CWM and EFL populations (30% versus 33% and 32% versus 38% of GRTS, respectively; Table 3, sample fraction =1/3; Fig. 3B). The estimators based on the SR, GRTS, and stratified GRTS designs are all guaranteed to be unbiased. The two regression estimators, however, can be biased. For this study we saw no evidence of substantial bias in the regression estimators (Fig. 3A; Table 3), with even the worst scenarios resulting in bias of 6% of population total.

The regression estimator requires additional sampling effort to collect the peak count census data. As described in the methods, a sample fraction of 1/2 and 1/3 without the peak count census would be equal to sample fractions of 2/5 and 1/5, with the peak count census assuming that there are six surveys across the season. Comparing CVRMSE of the regression estimators with sampling fractions of 1/3 and 1/5 to the CVRMSE for the other estimators with sampling fractions of 1/2 and 1/3 places the estimators on approximately equal footing in terms of reach visits. The regression estimator still performs better than all other estimators, with the exception of the GRTS stratified estimator for the MAG population (Fig. 3C). Sampling during the census requires less travel to and from reaches, therefore this additional cost is overstated by simply counting total reaches surveyed.

As expected, sampling a larger proportion of the population reduced CVRMSE (Fig. 3C). However, the relative performance of the different approaches was comparable across sample fractions. The analysis was repeated for four reaches sizes (0.25, 0.5, 1, and 2 km) at a sampling fraction of 1/3. For the MAG and CWM populations, smaller reach sizes performed better for SR and GRTS (Fig. 3 D), with the largest reach size resulting in twice the CVRMSE of the smallest for the MAG population. This difference was less pronounced for the stratified and regression estimators, and was negligible for all approaches in the EFL population where spatial autocorrelation was low.

The true 95% confidence interval coverage was almost always 95%, although generally greater than 80% and often closer to 90% (Fig. 5). Coverage varied by estimator, sample fraction, and reach size, but these patterns were not consistent across populations. Coverage was lowest for the nonstratified designs applied to the MAG population. This is due to the large positive skew (Table 2) caused by the high concentration of redds in the small area immediately surrounding the weir on Abernathy Creek. Samples that underrepresented the reaches from this area tended to produce low estimates for both the total and variance, resulting in confidence intervals that fell below the true value (Fig. 5A). Much better coverage was achieved with stratified sampling (Fig. 5), since every sample included all four weir reaches (for sample fractions of 1/2 and 1/3; Table 4).

Fig. 4. Redds per reach observed at the sampling event with highest mean number of visible redds (i.e., peak count) versus the total redds observed over all sampling events across the spawning season. Only the first 5 of the 7 MAG years were plotted to conserve space. The horizontal and vertical axes were constrained to emphasize the majority of the data. As a consequence some points are outside of the plotting region and were not displayed. The 1:1 line is included to assist in interpretation.



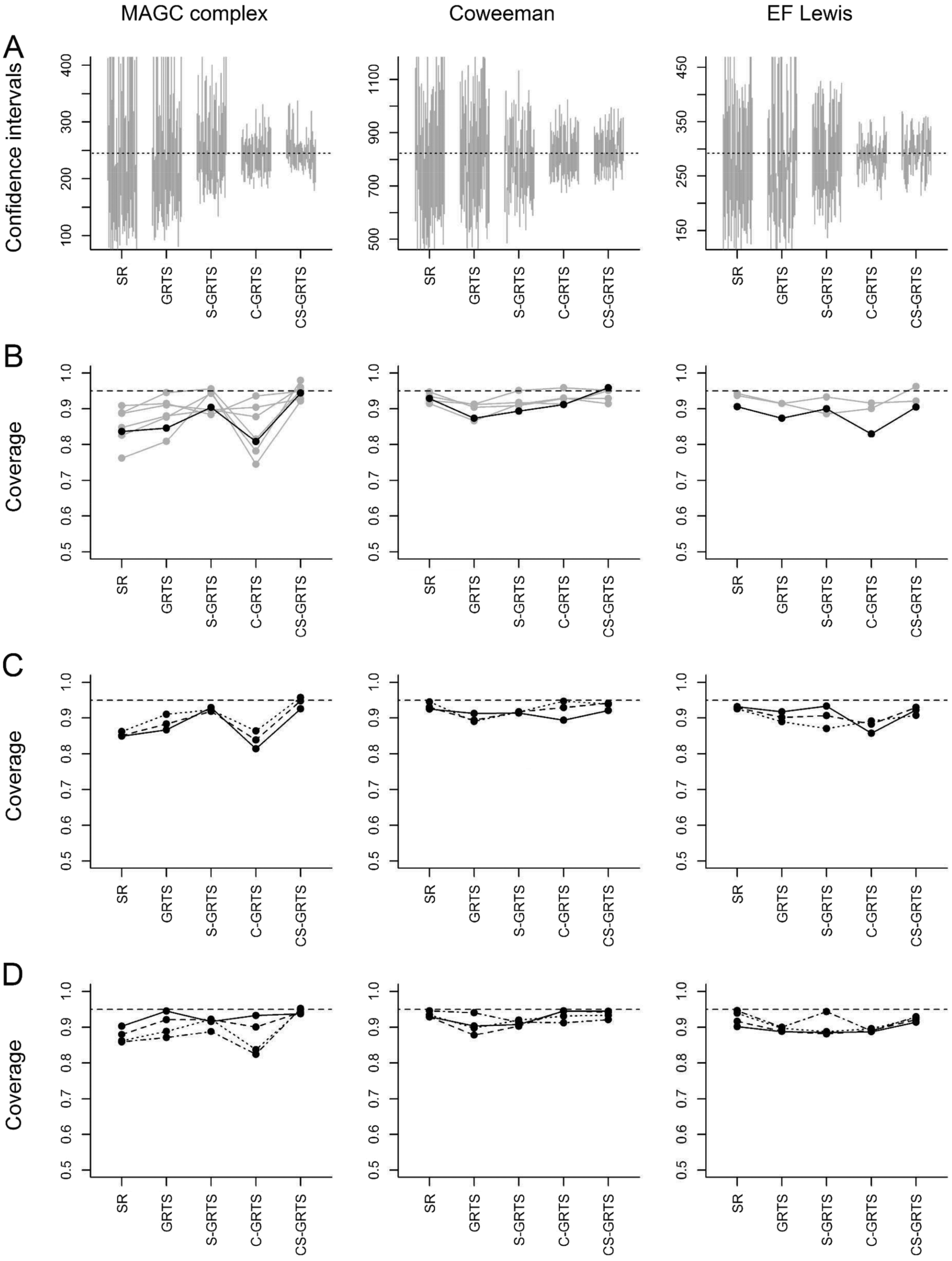
Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Discussion

The spatial distribution of salmon redds tends to be uneven owing to complex interactions between physical habitat, physiological requirements at different life stages, natural variability in fish behavior, and anthropogenic influences such as hatcheries (e.g., Irvine et al. 1992; Isaak and Thurow 2006). Consequently, redd surveys tend to produce reach counts that exhibit spatial autocorrelation and higher variability and right skew than would be expected if redds were distributed at random. All three populations in this study (MAG, CWM, and EFL) had uneven redd distributions. This led to reach counts with high variance for all populations (variance/mean > 5) and moderate to high spatial autocorrelation for two of the populations (averaging 0.55, 0.32, and 0.03 for MAG, CWM, and EFL, respectively). While this resulted in high CVRMSE for SR (simple random) sampling-based estimators, estimates based on stratification and the peak count census produced over threefold decreases in CVRMSE to well below 15%. For the MAG and CWM populations, where spatial autocorrelation was evident, both spatial balance (the GRTS design) and smaller reach sizes produced lower CVRMSE. Higher sampling fractions produced the expected decrease in CVRMSE but had little effect on the relative performance of the different approaches. The presence of right skew in a population (as seen in the redd counts) results in lower than their defined coverage for standard confidence intervals (e.g., Boos and Hughes-Oliver 2000). This was true for the populations in this study where the true coverage of the 95% confidence interval was almost always below 95%. While there are a number of approaches to improving coverage for skewed distributions (e.g., Zhou and Dinh 2005; Shilane et al. 2010), without additional information, constructing confidence intervals with good coverage can be very difficult for small samples from a skewed population (e.g., Shilane et al. 2010).

When there is auxiliary information related to the redd distribution beyond that provided by the sample, this information can be used to improve accuracy through integration in the sampling design or estimator. This knowledge can take the form of verified visual observations of where redd concentrations occur (as in the reaches in Abernathy creek associated with the weir) or auxiliary variables available for all reaches that are thought to be related to redd density (such as stream gradient or the peak count census in this study). Information available prior to sampling can be incorporated in the sampling design by (*i*) focusing sampling on reaches with higher predicted uncertainty and (*ii*) constraining samples to be approximately balanced (i.e., representative) with respect to an auxiliary variable (e.g., peak counts). Stratified designs are commonly used to achieve both these goals (e.g., Jacobs and Nickelson 1998; Courbois et al. 2008). By breaking the population into strata, the variability within each stratum can be estimated (based on the auxiliary variable), and this in turn can be used to sample more intensively in the strata with higher predicted variability. Also, by basing the strata boundaries on the auxiliary variable, balance with respect to the auxiliary variable is improved over simple random sampling, since each sample of reaches will be broken up into the different strata in the same way. For example, with a sampling fraction of 1/3 any sample of reaches from the stratified CWM design will include exactly 10 reaches from the high stratum and 7 reaches from the low stratum (Table 3), constraining the sample-based estimate of mean peak count to be closer on average to the true mean than under simple random sampling. In this study we found that adding stratification to the simple GRTS design reduced CVRMSE by at least a third for all populations. Stratification was particularly effective for the MAG population (S-GRTS CVRMSE < half of GRTS CVRMSE), where the most extreme values associated with the weir could be

Fig. 5. (A) 95% confidence intervals for the years indicated by the bold line in panel B based on 25 simulations arranged sequentially along the *x* axis for each estimator. The dashed line is the true population size. The confidence intervals were truncated above and below. (B) The coverage for the five different approaches using a sampling fraction of 1/3. Each line represents a different year. The bold line represents 2006, 2004, and 2006 for the MAG, CWM, and EFL populations, respectively. (C) The coverage for different sample fractions, averaged across years. Solid line = 1/5; dashed line = 1/3; dotted line = 1/2. (D) The coverage for four different reach sizes, averaged across years with sample fraction of 1/3. Solid line = 0.25 km; dashed line = 0.5 km; dotted line = 1 km; dash-dotted line = 2 km. SR = simple random sampling; GRTS = generalized random tessellation stratified sampling; S-GRTS = stratified GRTS; C-GRTS = regression estimator with GRTS; CS-GRTS = stratified GRTS with regression estimator.



Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

Table 4. Sample sizes for the different sample fractions, populations, and strata.

	Mill, Abernathy, and Germany creeks				Coweeman River			East Fork Lewis River		
Sample fraction	*N*	AG	Other	Weir	*N*	High	Low	*N*	High	Low
1/5	14	8	3	3	10	6	4	5	3	2
1/3	24	14	6	4	17	10	7	8	5	3
1/2	36	22	10	4	25	14	11	12	8	4
Mean		4.3	0.8	17.7		18.5	4.5		18.9	5.6
SD		3.7	1.5	15.1		9.5	4.4		8.7	4.4

Note: The strata are defined as follows: Weir, the reaches associated with the weir in Abernathy Creek; AG, Abernathy and Germany creeks (except the weir reaches); Other, all other reaches; High and Low, the strata with high and low predicted densities, respectively. *N* is the total number of samples. The mean and standard deviation (SD) of reach redd counts was calculated for each year, population, and stratum and then averaged across year.

isolated in a stratum in which all reaches were included in every sample (Fig. 2). For the EFL and CWM populations, a mean of the normalized peak count censuses across years was used as a stratifying variable. This could produce strata that change from year to year as additional peak count censuses are collected. Maintaining strata over years will simplify implementation and logistics of sampling, possibly at the cost of ignoring temporal shifts in spatial distributions of redds. In practice managers may include a number of additional sources of information to design strata such as cost and expert knowledge. The variance of the estimator based on a stratified design can be predicted if estimates of within-stratum variance are available (Cochran 1977). While other approaches are available for both producing samples that are approximately balanced with respect to an auxiliary variable and focusing sampling on reaches with high predicted variance (e.g., Deville and Tillé 2004), stratification is easy to implement and explain, and is well established in the fisheries literature. It is therefore likely to remain a popular method for using auxiliary data to improve sampling designs.

In addition to increasing accuracy through changes to the sample design, auxiliary information (such as the peak count census) can produce accuracy gains through modifications to the estimator (after sampling). In particular, regression estimators can improve upon the accuracy of standard estimators when the auxiliary variable is correlated with the variable of interest (Cochran 1977; Särndal et al. 1992). Because these estimators relying on an assumed model between the variable of interest and the auxiliary variable (e.g., a linear relationship for the regression estimator), these estimators are referred to as model based or model assisted as opposed to design based. For the populations in this study, the correlation, r, between peak and total redds was high, almost always over 0.9 (Fig. 4, Table 2). This was reflected in a decreases in CVRMSE of as much as two-thirds when applying the regression estimator (Table 3, Fig. 3). Under ideal conditions (large sample, linearity) the CVRMSE of the regression estimator is expected to be approximately $\sqrt{1-r^2}$ that of the standard estimator (Table 2) (Cochran 1977). In spite of the small sample sizes in our study, improvements predicted by this formula were comparable to those found through simulation (Tables 2 and 3), suggesting that this formula may be useful for assessing the utility of a regression estimator under similar conditions. While the spatial structure responsible for spatial autocorrelation may also increase the correlation between total and peak redd counts, improvements in precision derived from the regression estimator depend only on the correlation between peak and total counts (as is evident from the formula above). This is illustrated by the EFL population where the estimated spatial autocorrelation is essentially zero but the regression estimators are particularly effective (Fig. 3B; Table 3). Regression estimation can be especially effective for small sample sizes (e.g., Brewer 1999), where the chance inclusion of a couple extreme values can lead to samples that poorly represent the population. However, while design-based estimators (used for SR, GRTS, and S-GRTS in this study) are guaranteed to be unbiased, regression estimators that rely on model assumptions (e.g., linearity) can be substantially biased (Cochran 1977). Bias is proportional to the degree of nonlinearity and is exacerbated by small sample sizes (Cochran 1977). It is therefore important to check the relationship between the auxiliary variable and the variable of interest (Fig. 4). Fortunately, as illustrated by our study, bias can be low even when sample sizes are small (Table 3). Regression estimators can be generalized to include multiple auxiliary variables, nonlinearity, and different variance structures falling under generalized regression or calibration estimators (e.g., Särndal et al. 1992; Deville and Särndal 1992).

In this study we used a peak count census as an auxiliary variable. The high correlation between the peak and total counts for the reaches (Fig. 4) resulted in improvements in accuracy even when accounting for the cost of the peak count census. In addition to providing an auxiliary variable for the regression estimator, the peak count census was useful in defining strata and allocating the samples to the strata. Moreover, the peak count census provides an extensive survey that can be used to characterize the spatial distribution of redds, and the extent of spawning distribution for that year (Peacock and Holt 2012). However, as the sample frame (e.g., river basin size) increases in size, the sampling fraction necessary to achieve a fixed sample size (and therefore CV) will decrease, increasing the relative cost of adding a peak count census and perhaps limiting the use of peak count censuses to smaller populations. As basin size increases, differences in spawn timing across the different sub-basins may also be substantial, requiring separate regression relationships for each stratum. Here we were able to use the data to choose the timing of our peak count census. In practice the surveyors will have to use the timing of peak redd occurrence from previous years and the current year's trajectory to decide when to conduct the census. This is not likely to have a large effect, since the dates of peak count tended to be within 1–2 weeks of each other and the redds generally remain visible for well over 2 weeks (Kinsel et al. 2009). Also, the dates we used likely did not coincide exactly with the peaks, since we were constrained to counts that occurred once every week or 2 weeks. Correlation between peak and total counts from different years was substantially lower than those from the same year, averaging 62% versus 90%, 53% versus 94%, and 67% versus 96% for the MAG, EFL, and CWM populations, respectively. There are many other potential auxiliary variables that can be derived from GIS coverages (e.g., Busch et al. 2011) that may be useful in improving the accuracy of estimates and involve minimal additional effort. Also, if a variable derived from an extensive survey does not change substantially from year to year (e.g., substrate composition), the cost of the survey can be spread across several years by not surveying every year.

Even without specific knowledge of where redd density is high or low, if the spatial distribution is known to be clumpy, then spatial autocorrelation is likely, and theory predicts that moving to spatially balanced designs like GRTS that produce evenly distributed samples will tend to result in more accurate estimates

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

(e.g., Cochran 1977). This is because reaches closer together (common in SR sampling for example) have similar values, and therefore including both in the sample introduces redundancy. The magnitude of this effect depends on both the type and magnitude of the spatial autocorrelation (Cochran 1946), which is difficult to accurately characterize without a large data set. However, simplifying assumptions can be used to get a rough idea of how much can be gained from a spatially balanced design. For example, assuming an exponential correlation structure, CVRMSE for a systematic sample of every third reach will on average be 62% of the CVRMSE for the SR sample when spatial autocorrelation is 0.54 (the mean value for the MAG population; Table 2, Cochran 1946). This is larger than the observed reduction to 78% using the spatially balanced GRTS design (Table 3), likely due to improper characterization of the spatial structure and difficulty in achieving good spatial balance with a discrete sampling frame and high sample fraction. While in our study, spatial balance (GRTS) was less effective in improving accuracy than strata and the peak count census, implementing a spatially balanced design (GRTS) requires little additional effort with few down sides and potentially large benefits. Systematic sampling can produce better spatial balance and smaller CVRMSE than GRTS (e.g., Courbois et al. 2008; Jacobs et al. 2009), but this difference is generally small; GRTS has the added advantage of having an estimable variance and straightforward method of adding sample points when necessary (e.g., Stevens and Olsen 2004).

Another approach to more evenly distributing the sampling effort in space is to divide the total length sampled into smaller reaches (for example, one hundred 100 m reaches versus ten 1 km reaches). As expected, the reduction in CVRMSE from moving to smaller reaches was related to the degree of spatial autocorrelation in the population. For the MAG population with the highest mean spatial autocorrelation, moving to smaller reaches dramatically reduced CVRMSE (Fig. 3 D), while no effect was evident for the EFL population where spatial autocorrelation was negligible. This is consistent with the work of Courbois et al. (2008), who also found that smaller reach sizes improved accuracy. The benefits of dividing sampling effort into smaller reaches will depend on the degree to which the added travel cost is offset by increased accuracy. In the EFL population, for example, the added travel cost would have to be very small, since decreasing reach size has little effect on accuracy.

Patchy spatial distributions, which are common to salmon and steelhead redds, can lead to high uncertainty in estimates of total redds. For the populations in this study, sampling a third of the reaches produced CVs higher than 15% in all cases for SR sampling and in almost all cases for GRTS sampling. However, leveraging knowledge of this spatial structure using designs that were stratified, spatially balanced, and included the peak count census reduced CVs to less than 8%. These improvements are predicted by theory (Cochran 1977) and previous empirical studies (Jacobs and Nickelson 1998; Courbois et al. 2008; Jacobs et al. 2009; Gallagher et al. 2010), and are the result of improved balance (more representative samples), concentration of effort where there is more uncertainty, and the inclusion of additional information and assumptions (regression estimators). Given the degree of spatial pattern in this and other redd data, spatially balanced designs are likely to lead to improvements in accuracy and involve little additional cost. The usefulness of stratifying and regression estimators depends critically on the auxiliary information used. We found that a peak count census proved to be a very effective auxiliary variable even when the additional cost of collecting this information was accounted for. However, the magnitude of these improvements was a product of the high correlation between peak and total redd counts. Correlation between redd counts and habitat-based variables will likely be more modest (e.g., Gallagher and Gard 1999). While we used sample fraction as a measure of sampling intensity, sample size and the size of the population will be the primary determinants of the CV for the population estimate. Therefore, emphasis should be placed on the relative performance of the estimators when generalizing. However, the three populations in this study include substantially different spatial distributions (Table 2) and are similar in size to many populations in the Pacific Northwest for which redd surveys are used to estimate escapement. Combined with the approximate sampling theory predictions, these empirical results provide a good indication of the improvements possible with the described methods.

There are many opportunities for future research in redd-based escapement estimation. These include more robust methods for expanding redd counts to total escapement, consideration of temporal panel designs and associated estimators, examination of the tradeoffs between the number of visits over the sampling season and the number of reaches sampled, other auxiliary variables, more robust methods of constructing confidence intervals (such as the jackknife and model-based estimators with distributional assumptions), and models that use information from several populations to better describe and explain between population differences and leverage this information to provide more accurate estimates for individual populations.

Acknowledgements

We thank the many technicians and biologists that collected the redd data and Steve VanderPloeg for providing substantial GIS assistance. Constructive reviews of the manuscript by John McMillan, Eric Buhl, and three anonymous reviewers, as well as helpful comments from Alistair Coulthard, led to substantial improvements in the manuscript. Funding support for this project was provided by Pacific Coast Salmon Recovery Fund, Pacific Salmon Commission's Letter of Agreement (Chinook Technical Committee) and Southern Boundary Fund, NOAA-Fisheries Mitchell Act, the Washington State General Fund, and the Bonneville Power Administration through the Pacific Northwest Aquatic Monitoring Partnership.

References

Baillargeon, S., and Rivest, L.P. 2011. The construction of stratified designs in R with the package *stratification*. Surv. Methodol. **37**(1): 53–65. Available from www.statcan.gc.ca/pub/12-001-x/2011001/article/11447-eng.pdf [accessed January 2014].

Boos, D.D., and Hughes-Oliver, J.M. 2000. How large does *n* have to be for *Z* and *t* intervals? Am. Stat. **54**: 121–128. doi:10.2307/2686030.

Brewer, K.R.W. 1999. Design-based or prediction-based inference? Stratified random vs stratified balanced sampling. Int. Stat. Rev. **67**: 35–47. doi:10.2307/1403564.

Busch, S.H., Sheer, M., Burnett, K., McElhany, P., and Cooney, T. 2011. Landscape-level model to predict spawning habitat for lower Columbia River fall Chinook Salmon (*Oncorhynchus tshawytshca*). River Res. Appl. **29**: 297–312. doi:10.1002/rra.1597.

Cochran, W.G. 1946. Relative accuracy of systematic and stratified random samples for a certain class of populations. Ann. Mat. Stat. **17**: 164–177. doi:10.1214/aoms/1177730978.

Cochran, W.G. 1977. Sampling techniques. Wiley and Sons, New York.

Courbois, J.-Y., Katz, S.L., Isaak, D.J., Steel, E.A., Thurow, R.F., Wargo Rub, A.M., Olsen, T., and Jordan, C.E. 2008. Evaluating probability sampling strategies for estimating redd counts: an example with Chinook salmon (*Oncorhynchus tshawytscha*). Can. J. Fish. Aquat. Sci. **65**(9): 1814–1830. doi:10.1139/F08-092.

Crawford, B.A. and Rumsey, S. 2011. Guidance for monitoring recovery of Pacific Northwest salmon and steelhead listed under the Federal Endangered Species Act [online]. Internal report of the National Marine Fisheries Service Northwest Region. Available from http://www.pnamp.org/document/3338 [accesses January 2014].

Dalenius, T. and Hodges, J.L., Jr. 1959. Minimum variance stratification. J. Am. Stat. Assoc. **54**: 88–101. doi:10.1080/01621459.1959.10501501.

Deville, J.-C., and Särndal, C.E. 1992. Calibration estimators in survey sampling. J. Am. Stat. Assoc. **87**: 376–382. doi:10.1080/01621459.1992.10475217.

Deville, J.-C., and Tillé, Y. 2004. Efficient balanced sampling: the cube method. Biometrika, **91**: 893–912. doi:10.1093/biomet/91.4.893.

Gallagher, S.P., and Gard, M.F. 1999. Relationship between chinook salmon (*Oncorhynchus tshawytscha*) redd densities and PHABSIM-predicted habitat in the Merced and Lower American rivers, California. Can. J. Fish. Aquat. Sci. **56**(4): 570–577. doi:10.1139/f98-198.

Gallagher, S.P., Hahn, P.K., and Johnson, D.H. 2007. Redd counts. *In* Salmonid field protocols handbook: techniques for assessing status and trends in

Can. J. Fish. Aquat. Sci. Downloaded from www.nrcresearchpress.com by AMI HOLLINGSWORTH on 02/23/15
For personal use only.

salmon and trout populations. *Edited by* D.H. Johnson, B.M. Shrier, J.S. O'Neil, J.A. Kntuzen, X. Aguerot, T.A. O'Neil, and T.N. Pearsons. American Fisheries Society, Bethesda, Md. pp. 197–234.

Gallagher, S.P., Adams, P.B., Wright, D.W., and Collins, B.W. 2010. Performance of spawner survey techniques at low abundance levels. N. Am. J. Fish. Manage. **30**(5): 1086–1097. doi:10.1577/M09-204.1.

Gresswell, R.E., Torgersen, C.E., Bateman, D.S., Guy, T.J., Hendricks, S.R., and Wofford, J.E.B. 2006. A spatially explicit approach for evaluating relationships among coastal cutthroat trout, habitat, and disturbance in small Oregon streams. *In* Landscape influences on stream habitats and biological assemblages. *Edited by* R.M. Hughes, L. Wang, and P.W. Seelbach. American Fisheries Society, Bethesda, Md. pp. 457–471.

Hassemer, P. 1993. Salmon spawning ground surveys, 1989-1992. Idaho Department of Fish and Game Project F-73-R-15. Pacific Salmon Treaty Program Award NA17FP0168-02.

Irvine, J.R., Bocking, R.C., English, K.K., and Labelle, M. 1992. Estimating coho salmon (*Oncorhynchus kisutch*) spawning escapements by conducting visual surveys in areas selected using stratified random and stratified index sampling designs. Can. J. Fish. Aquat. Sci. **49**(10): 1972–1981. doi:10.1139/f92-219.

Isaak, D.J., and Thurow, R.F. 2006. Network-scale spatial and temporal variation in Chinook salmon (*Oncorhynchus tshawytscha*) redd distributions: patterns inferred from spatially continuous replicate surveys. Can. J. Fish. Aquat. Sci. **63**(2): 285–296. doi:10.1139/f05-214.

Jacobs, S.E., and Nickelson, T.E. 1998. Use of stratified random sampling to estimate the abundance of Oregon coastal coho salmon [online]. Report of the Oregon Department of Fish and Wildlife, Project F-145-R-09. Available from http://www.oregonstate.edu/dept/ODFW/spawn/reports.htm [accessed March 2010].

Jacobs, S.E., Gaeuman, W., Weeber, M.A., Gunckel, S.L., and Starcevich, S.J. 2009. Utility of a probabilistic sampling design to determine bull trout population status using redd counts in basins of the Columbia River plateau. N. Am. J. Fish. Manage. **29**: 1590–1604. doi:10.1577/M08-210.1.

Kincaid, T.M., and Olsen, A.R. 2012. Spsurvey: Spatial Survey Design and Analysis. R package version 2.3 [online]. Available from http://www.epa.gov/nheerl/arm/.

Kinsel, C., Hanratty, P., Zimmerman, M., Glaser, B., Gray, S., Hillson, T., Rawding, D., and VanderPloeg, S. 2009. Intensively monitored watersheds: 2008 fish population studies in the Hood Canal and Lower Columbia stream complexes. FPA 09-12, Washington Department of Fish and Wildlife, Olympia, Wash.

Leider, S.A., Chilcote, M.W., and Loch, J.J. 1984. Spawning characteristics of sympatric populations of steelhead trout (*Salmo gairdneri*): evidence for partial reproductive isolation. Can. J. Fish. Aquat. Sci. **41**(10): 1454–1462. doi:10.1139/f84-179.

Lumley, T. 2012. *Survey*: analysis of population survey samples. R package version 3.28-2.

McElhany, P., Ruckelshaus, M.H., Ford, M.J., Wainright, T.C., and Bjorkstedt, E.P. 2000. Viable salmonid populations and the recovery of evolutionarily significant units. US Dep. Commer. NOAA Tech. Memo. NMFS-NWFSC-42.

Peacock, S.J., and Holt, C.A. 2012. Metrics and sampling designs for detecting trends in the distribution of spawning Pacific salmon (*Oncorhynchus* spp.). Can. J. Fish. Aquat. Sci. **69**(4): 681–694. doi:10.1139/f2012-004.

Perry, J.N., and Mead, R. 1979. On the power of the index of dispersion test to detect spatial pattern. Biometrics, **35**: 613–622. doi:10.2307/2530252.

R Development Core Team. 2012. R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing, Vienna. Available from www.R-project.org/ [accessed January 2014].

Rawding, D., Wilson, J., Glaser, B., VanderPloeg, S., Holowatz, J., Buehrens, T., Gray, S., and Gleizes, C. 2014. Fall Chinook salmon escapement estimates and coded-wire-tag recoveries in Washington's Lower Columbia River tributaries in 2010. Washington Department of Fish and Wildlife, Olympia, Wash.

Särndal, C.E., Swensson, B. and Wretman, J. 1992. Model assisted survey sampling. Springer-Verlag, New York.

Shilane, D., Evans, S.N., and Hubbard, A. 2010. Confidence intervals for negative binomial random variables of high dispersion. Int. J. Biostat. **6**(1). Article 10. doi:10.2202/1557-4679.1164.

Steel, E.A., Fullerton, A.H., Caras, Y., Sheer, M.B., Olson, P., Jensen, D., Burke, J., Maher, M.T., and McElhany, P. 2009. A spatially explicit decision support system for managing wide-ranging species. Ecol. Soc. **13**(2): 50.

Stevens, D.L., Jr., and Olsen, A.R. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics, **14**: 593–610. doi:10.1002/env.606.

Stevens, D.L., Jr., and Olsen, A.R. 2004. Spatially balanced sampling of natural resources. J. Am. Stat. Assoc. **99**(465): 262–278. doi:10.1198/016214504000000250.

Stevens, D.L., Jr., Larsen, D.P., and Olsen, A.R. 2007. The role of sample surveys: why should practitioners consider using a statistical sampling design? *In* Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. *Edited by* D.H. Johnson, B.M. Shrier, J.S. O'Neil, J.A. Kntuzen, X. Aguerot, T.A. O'Neil, and T.N. Pearsons. American Fisheries Society, Bethesda, Md. pp. 11–24.

Tillé, Y. 2011. Ten years of balanced sampling with the cube method: an appraisal. Surv. Methodol. **37**: 215–226. Statistics Canada, Catalogue No. 12-001-X.

USFWS. 2006. Natural reproductive success and demographic effects of hatchery origin steelhead in Abernathy Creek, Washington. United States Fish and Wildlife Service 2006 Annual Report, Project No. 200306300, 65 electronic pages, (BPA Report DOE/BP-00016522).

Washington Department of Fish and Wildlife. 2002. Fisheries Management Salmonid Stock Inventory (SaSI). Salmon and steelhead stock reports [online]. Available from wdfw.wa.gov/conservation/fisheries/sasi/ [accessed January 2014].

Zhou, X.H., and Dinh, P. 2005. Nonparametric confidence intervals for the one- and two-sample problems. Biostatistics, **6**(2): 187–200. doi:10.1093/biostatistics/kxi002.