National Marine
Fisheries Service
NOAA

Fishery Bulletin
established in 1881

Spencer F. Baird
First U.S. Commissioner
of Fisheries and founder
of *Fishery Bulletin*

Abstract—In 2-stage fishery sampling, abundance is often estimated by using a primary sampling gear and total abundance is then partitioned into groups of interest by applying data on composition derived from a secondary sampling gear. However, the literature is sparse on statistical properties of estimates of run composition. We examined the accuracy and precision of estimators of composition of wild steelhead (*Oncorhynchus mykiss*) in the Snake River, in the Pacific Northwest. We simulated estimators, using pooled and time-stratified data. We compared confidence intervals (CIs) determined on the basis of asymptotical normality or a 2-stage bootstrap method. Stratified estimators were unbiased, except in a few cases. Joint CIs (all groups considered simultaneously) had coverages near nominal. Conversely, pooled estimators performed poorly; the proportion of biased estimates increased as the number of groups estimated increased. Using empirical data, we show that CIs met precision goals for most groups. Half-widths of CIs decreased and stabilized as the number sampled and group abundance increased. In complex scenarios, estimates of small groups will yield poor precision and some may be biased, but a stratified estimate with a conservative joint CI can be of practical use. The 2-step bootstrap approach is flexible and can incorporate other sources of variability or sampling constraints.

Manuscripts submitted 22 December 2015.
Manuscripts accepted 25 August 2016.
Fish. Bull. 115:1–12 (2017).
Online publication date: 18 October 2016.
doi: 10.7755/FB.115.1.1

Abundance estimates and confidence intervals for the run composition of returning salmonids

Kirk Steinhorst[1]
Timothy Copeland (contact author)[2]
Michael W. Ackerman[3]
William C. Schrader[2]
Eric C. Anderson[4]

Email address for contact author: tim.copeland@idfg.idaho.gov

[1] Department of Statistical Science
University of Idaho 415A Brink Hall
875 Perimeter Drive
Moscow, Idaho 83844-1104

[2] Idaho Department of Fish and Game
1414 East Locust Lane
Nampa, Idaho 83686

[3] Eagle Fish Genetics Laboratory
Idaho Department of Fish and Game
Pacific States Marine Fisheries Commission
1800 Trout Road
Eagle, Idaho 83616

[4] Fisheries Ecology Division
National Marine Fisheries Service
Southwest Fisheries Science Center, NOAA
110 Shaffer Road
Santa Cruz, California 95060

In 2-stage sampling for fisheries monitoring and research, abundance is estimated with a primary sampling gear and then partitioned into groups of management interest by applying compositional data (e.g., species, stock, sex, age, and size) derived from a secondary sampling gear. For example, biological samples obtained from gillnetting or electrofishing can be used to allocate abundance estimates from hydroacoustic counts to species (e.g., Tarbox and Thorne, 1996; Pritt et al., 2013; Rudstam et al., 2013; Hughes and Hightower, 2015). Alternatively, a more highly controlled sampling regime can be instituted by counting fish as they move past barriers (e.g., weirs or dams) and by collecting biological samples or data from some portion of the fish in order to partition counts (e.g., Wagner, 2007; Campbell et al., 2012). However, the complexities of fishery sampling programs and the relevant groups into which the fish are parsed present difficulties for estimating precision of the generated point estimates.

Steelhead (*Oncorhynchus mykiss*) are an important cultural, economic, and recreational resource in the Pacific Northwest of the United States. After the construction of hydroelectric dams on the Columbia and Snake rivers during the late 1960s and early 1970s, the abundance and survival of steelhead in the Snake River decreased (Raymond, 1988). In response, steelhead within the Snake River basin were listed as threatened under the Endangered Species Act in 1997. In recent years, abundances have increased slightly. However, the increase has been dominated by fish produced in hatcheries (intended to mitigate for reduced harvest opportunities and to supplement natural populations), while the returns of steelhead born in the natural environment remain critically low (Ford, 2011). Fishery biologists need to know how many wild versus hatchery-produced steelhead return in order to manage fisheries effectively, as well as to assess the conservation status of wild populations. Further, for wild fish, we need to know the numbers of fish returning by sex, age, and stock to inform viability analyses.

We collected data on the run composition of adult steelhead as they migrated past Lower Granite Dam (LGD) on the Snake River, 695 km from the ocean. Adults returning from the Pacific Ocean to spawn in tributaries of the Snake River must ascend fish ladders at 8 dams during their migration; Lower Granite Dam is the final dam they encounter before dispersing to spawn. An observation window on the LGD fish ladder (the primary sampling "gear") allows the enumeration of fish by species as they migrate upstream. A trapping facility (a diversion gate in the fish ladder with chutes leading to a holding tank) located above the observation window (the secondary sampling gear) allows the interception of fish and the collection of biological data (Harmon, 2003). Counting and sampling returning adult steelhead at the dam provide the data for calculating run composition (Schrader et al.[1]). Surprisingly, the primary literature is sparse on the statistical properties of estimates derived with current methods and applied to run composition.

In our study, we examined the properties of estimators of fish composition and confidence intervals (CIs) derived from weightings of counts of fish at the observation window, data on origin (wild versus hatchery) obtained from the samples taken at the trapping facility, and compositional data (sex, age, and genetically defined stock) collected from wild fish subsampled at the trapping facility. We considered counts at the observational window to provide a census of fish passing the dam. Initially, we assumed trapping rates (proportion of time the trap was open) could be precisely controlled to obtain a constant proportion of the fish throughout the run and, therefore, that data could be pooled across time to estimate abundance. However, logistical issues that affected trapping rates through time led us to investigate temporally stratified estimators. Individual CIs (for each group considered independently) and joint CIs (for all groups within a variable of interest considered simultaneously) were derived 1) by using closed-form asymptotically normal equations or 2) a 2-step bootstrap sampling method, by origin (hatchery or wild) of the fish, by using compositional data collected from wild fish. Using simulations, we compared the options for developing accurate estimates of abundance and CIs with good coverage; we then applied the preferred method from the simulations to empirical data to develop guidance for sampling and interpreting fisheries data on fish composition.

Materials and methods

We used data to describe the abundance and composition of wild adult steelhead migrating past LGD to spawn in the Snake River during spring 2011. Spawning year (SY) 2011 is defined as the year when adult steelhead migrate past LGD between 1 July 2010 and 30 June 2011. Although all steelhead in the Snake River basin spawn in the spring, the majority migrate past LGD during the previous fall, and a smaller portion migrates during the spring just before spawning. Later in this section, we describe the data set and estimation procedures and the simulations developed to test the bias of the estimators and the coverage of the associated CIs. A complete description of the collection methods and data used in this study is given by Schrader et al.[1]

Data collection

Primary sampling stage (counts from the observation window) Adult steelhead were counted as they passed a viewing window located in the LGD fish ladder, which they must ascend to migrate upriver. Counts of fish observed from the window were conducted during a majority of the year and occurred daily at 0400–2000 Pacific Time. Counts from videos were used in lieu of counts from the window in November, December, and March and occurred at 0600–1600. Most fish pass the window during the 10–16 h of daylight when counts are made. The ladder is drained and closed in January and February, and, as a result, adult steelhead cannot migrate upriver during those months. Count data were downloaded from the U.S. Army Corps of Engineers website (website). The steelhead count consists of all fish >30 cm in fork length identified as *O. mykiss*. Daily counts were aggregated on a weekly basis. We further combined weeks into longer time periods (up to 2 months) if few fish (<75 individuals) were passing the window during a week—a level observed typically early and late in the migration season.

Secondary sampling stage (trapping rates) Trapping rates were determined by a committee of co-managers balancing sampling requirements for multiple projects with fish handling concerns. The trap is operational 24 h/day and the trapping rate determines how long a trap gate remains open 4 times/h, such that a daily systematic sample (by time) is taken from the fish ascending the fish ladder. Thus, the trapping rate (proportion of an hour that the trap gate is open) approximates the desired proportion of the population to be sampled. Trapping rates are typically 10–20%; for the majority of the SY2011 run, it was set at 10%.

Trapped fish were anesthetized and examined to determine whether they were of hatchery or wild origin. In the Snake River basin, most hatchery-origin steelhead have a clipped adipose fin; however, some are released with an intact adipose fin to supplement natural populations. Therefore, unclipped steelhead were examined for the presence of dorsal or ventral fin erosion, which often occurs in hatchery-reared fish (Latremouille, 2003). Unclipped hatchery fish may also be identified by the presence of a coded wire tag, by

[1] Schrader, W. C., M. P. Corsi, P. Kennedy, M. W. Ackerman, M. R. Campbell, K. K. Wright, and T. Copeland. 2013. Wild adult steelhead and Chinook salmon abundance and composition at Lower Granite Dam, spawn year 2011. 2011 annual report. Idaho Dep. Fish Game, IDFG Rep. 13-15, 89 p. [Available at website.]

parentage determined genetically (Steele et al., 2013), or by a ventral-fin clip. Genotyping procedures for parentage were conducted after the trapping season and gives accuracy rates approaching 100% (Steele et al., 2013). Fish not determined to be of hatchery origin were treated as wild fish.

Subsampling of trapped wild steelhead Scale and tissue samples were then taken from a systematic subsample of trapped fish deemed wild. Percentages of the wild steelhead that were subsampled averaged around 50% during this study. Scale samples were used to determine age on the basis of visual examination of scale annuli. Age data collected at LGD were used to assign returning adults back to a brood year (BY, the year in which their parents spawned).

Tissue samples from the anal fin were used to determine sex and the stock of origin. Stock composition was determined by using individual assignment, a method of genetic stock identification (Pella and Milner, 1987; Shaklee et al., 1999) based on single nucleotide polymorphisms (SNPs). Adults were screened at 187 SNPs and with a sex-specific allelic discrimination assay (Campbell et al., 2012). Only individuals that were genotyped at >90% of SNPs were included. We used the maximum likelihood framework implemented in the program gsi_sim (Anderson et al., 2008; Anderson, 2010) to assign individuals to a stock. Each fish was assigned to the stock in which the probability of its genotype occurring was greatest by using the allocate-sum procedure (Wood et al., 1987). We did not attempt to identify out-of-basin strays. For this study, we assumed that the stock was determined without error (a future study will examine uncertainty in genetic assignments). In essence, we treated the genetic data in the same way as we did for the age data.

Ackerman et al.[2] defined 10 genetically determined stocks used for assignments at LGD. The locations of these stocks included 1) the upper Salmon River (UPS); 2) Middle Fork Salmon River (including Chamberlain and Bargamin creeks) (MFS); 3) South Fork Salmon River (SFS); 4) lower Salmon River (LOS); 5) upper Clearwater River (Lochsa and Selway rivers) (UPC) ; 6) South Fork Clearwater River (including Clear Creek) (SFC); 7) lower Clearwater River (LOC); 8) Imnaha River (IMN); 9) Grande Ronde River (GRR); and 10) Tucannon River, Asotin Creek, and other tributaries to the Snake River downstream of the Clearwater River confluence (LSN).

The sampling design produced 3 data sets: 1) a census of numbers of fish returning to and migrating past the dam (window counts), 2) a hatchery-versus-wild data set for all trapped fish (trap data), and 3) a data set containing sex, age, and stock for a subsample of wild fish that were trapped (for compositional data). These 3 data sets were used to produce estimates and CIs for the number of wild fish by sex or age or stock.

[2] Ackerman, M. W., N. V. Vu, J. McCane, C. A. Steele, M. R. Campbell, A. P. Matala, J. E. Hess, and S. R. Narum. 2014. Chinook and steelhead genotyping for genetic stock identification at Lower Granite Dam. Project progress report. 2013 annual report. Idaho Dep. Fish Game, IDFG Rep. 14-01, 60 p. [Available at website]

Estimator and confidence intervals

Abundance of wild steelhead The window-count data provided the abundance of adult steelhead migrating past LGD, but our focus was on wild fish; therefore, we first had to partition the overall abundance estimate into a wild-versus-hatchery abundance estimate. The proportions of wild and hatchery steelhead changed over the season, but within each weekly or monthly stratum, proportions were assumed to be relatively constant. Given the window counts by strata, $C_1, C_2, \ldots, C_S$, the number of wild steelhead $(\widehat{W})$ was estimated with the following equation:

$$\widehat{W}_{.} = \sum_{s=1}^{S} \hat{p}_s C_s, \tag{1}$$

where $\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_s$ = estimates of the proportion of wild steelhead by stratum from the trap data; and

$\widehat{p_s}$ = N_s/t_s (or denoted $\hat{p}_.$ for a pooled estimate).

These and all subsequent notations are defined in Table 1. Given the fixed numbers of adults counted at the dam for each stratum, we found a CI for the number of wild fish by using either an asymptotically normal interval or by a parametric bootstrap. The asymptotically normal interval is given as

$$\widehat{W}_{.} - Z_{\alpha/2} S_{\widehat{W}_{.}} \le W \le \widehat{W}_{.} + Z_{\alpha/2} S_{\widehat{W}_{.}}, \tag{2}$$

$$\text{where } S^2_{\widehat{W}_{.}} = \sum_{s=1}^{S} C_s^2 S^2_{\widehat{p_s}} = \sum_{s=1}^{S} C_s^2 \left(1 - \frac{t_s}{C_s}\right) \frac{\hat{p}_s(1-\hat{p}_s)}{t_s - 1}; \tag{3}$$

and $Z_{\alpha/2}$ = the (100α/2)th percentile of a standard normal distribution.

For the parametric bootstrap, we assumed the bootstrap number of wild fish has a binomial distribution, $N_s^* \sim \text{binomial}(t_s, \hat{p}_s | t_s)$ and $p_s^* = N_s^*/t_s$. We produced 500 sets (B) of $(p_1^*, p_2^*, \ldots, p_S^*)$, yielding $W_1^*, W_2^*, \ldots, W_B^*$. The 100α/2 and $(1 - \frac{\alpha}{2})$ percentiles of $W_1^*, W_2^*, \ldots, W_B^*$ gave us the 100(1 – α)% CI for the true number of wild steelhead that passed LGD for the year. In this study, we used 90% CIs as an acceptable tradeoff of the type-I error rate with the power to discern important differences.

Composition of wild steelhead When we had estimates of the number of wild fish migrating past LGD, we partitioned them into groups of interest for population assessments. There were 2 competing approaches for estimating numbers of wild fish by sex, age, or genetic origin: pooled or stratified. The compositional data set was about a tenth of the size of the hatchery-versus-wild data set (most fish were hatchery-origin; wild fish were subsampled at approximately 50%). There was ample data in the trap data set to estimate proportion of wild steelhead by stratum but not enough data for

Table 1

Definitions for notation used to describe parameters in estimating abundance and confidence intervals for the constituent groups that compose the run of steelhead (*Oncorhynchus mykiss*) in Snake River of the Pacific Northwest during spawning year 2011.

Parameter	Definition
A	Number of age groups
B	Number of bootstrap samples
C_s	Window count in stratum s; s=1,2,…S
F	Abundance of female steelhead
G	Number of genetic stocks
k	Number of categories in the compositional variable of interest
L	Lower bound of a confidence interval
M	Abundance of male steelhead
N_s	Number of wild steelhead trapped in stratum s; s=1,2…S
$\widehat{p_s}$	Estimated proportion of wild steelhead in stratum s; s=1,2…S
$\widehat{\pi_{is}}$	Estimated proportion of group i of the variable of interest (A, G, F / M) in stratum s; s=1,2…S
r_s	Number of wild steelhead subsampled in stratum s; s=1,2,…S
S	Number of time strata
t_s	Number of steelhead trapped in stratum s; s=1,2…S
U	Upper bound of a confidence interval
$\widehat{W}_s$	Estimated abundance of wild steelhead in period s; s=1,2…S

the wild compositional data for some strata. We could have pooled the compositional data over the season if we had assumed that we had sampled a fixed proportion of the wild fish for each stratum or if we had assumed that the composition proportions were constant. However, if neither condition was true, we had to focus on obtaining sufficient samples in each stratum to obtain stable estimates of composition proportions by stratum $\hat{\pi}_{is}$.

We defined the proportion of wild females in a stratum (π_{FS}) as P(*female* | *wild, stratum s*). Then we had (π_{F1}, π_{M1}), ..., (π_{FS}, π_{MS}) as the conditional probabilities for wild females and males for strata 1, ..., s. These proportions were estimated from data obtained from the subsample of the trapped fish; for example, for females $\hat{\pi}_{FS} = r_{FS} / (r_{FS} + r_{MS})$. Given estimates of these probabilities from wild fish examined in the trap, we used the following equation to estimate female abundance:

$$\hat{F} = \sum_{s=1}^{S} \hat{\pi}_{F_S} \hat{W}_S = \sum_{s=1}^{S} \hat{\pi}_{F_S} \hat{p}_s C_s, \tag{4}$$

and to estimate wild male abundance

$$\hat{M} = \sum_{s=1}^{S} \hat{\pi}_{M_S} \hat{W}_S = \sum_{s=1}^{S} \hat{\pi}_{M_S} \hat{p}_s C_s. \tag{5}$$

For the pooled estimators, we dropped the summation and subscript. Similar estimates were made for A (ages, BY2004–BY2008 in this study), G (genetically identified stocks), and A×G age groups for each stock. That is, the number of wild steelhead in any group is a weighted sum of the stratified window counts, in which the weights are estimates of the probabilities of being wild and being a member of a particular group, including any combinations of the compositional variables.

To find CIs for these estimates, we had to account for the variability of both the trap data and the compositional data. For the asymptotically normal interval, we used Goodman's (1960) estimated variance of a product:

$$s_{\hat{F}}^2 = \sum_{s=1}^{S} C_s^2 (\hat{p}_s^2 s_{\hat{\pi}_{F_s}}^2 + \hat{\pi}_{F_s}^2 s_{\widehat{p_s}}^2 - s_{\widehat{p_s}}^2 s_{\hat{\pi}_{F_s}}^2, \tag{6}$$

yielding

$$\hat{F} - Z_{\alpha/2} s_{\hat{F}} \le F \le \hat{F} + Z_{\alpha/2} s_{\hat{F}}, \tag{7}$$

where $s_{Fs}^2 = (1 - \frac{r_s}{\hat{W}_s}) \hat{\pi}_{Fs} (1 - \hat{\pi}_{Fs}) / (r_s - 1)$.

For the pooled case, we used the following equation:

$$s_{\hat{F}}^2 = \widehat{W}_{.}^2 s_{\hat{\pi}_F}^2 + \hat{\pi}_F^2 s_{\widehat{W}}^2 - s_{\widehat{W}}^2 s_{\hat{\pi}_F}^2, \tag{8}$$

where $s_{\hat{\pi}_F}^2 = (1 - \frac{r.}{\hat{W}.}) \hat{\pi}_F (1 - \hat{\pi}_F) / (r. - 1)$; and (9)

$\hat{\pi}_F$ = the proportion of wild females from the pooled sample.

Similar formulae for $\hat{M}$ and estimates of age, stock, and ages by stock follow in the next paragraph.

The bootstrap process described previously for obtaining the CI for the number of wild fish was extended by adding a conditional bootstrap loop based on the sex, age, and stock of wild fish in the trap. We defined pseudoreplicates parametrically, using

$$(F_s^*, M_s^*) \sim \text{binomial}(r_s, (\hat{\pi}_{Fs}, \hat{\pi}_{Ms}) \mid r_s, \text{wild}) \text{ and} \tag{10}$$

$$(\pi_{Fs}^* = \frac{F_s^*}{m_s}, \pi_{Ms}^* = \frac{M_s^*}{m_s}).$$

Bootstrap values for the total number of wild females, $F_1^*, \ldots, F_B^*$, were determined with the following equation:

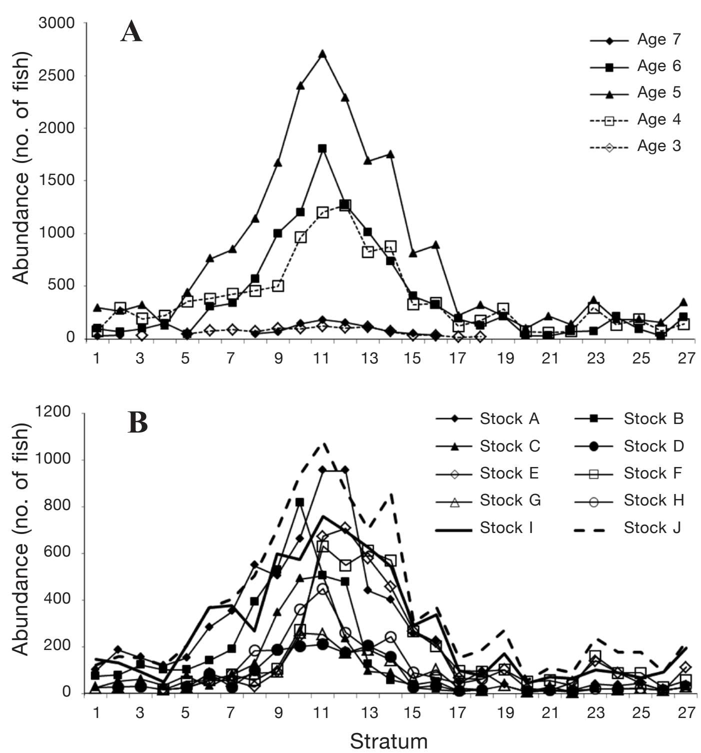

Figure 1

Simulated abundance of wild steelhead trout (*Oncorhynchus mykiss*) by weekly time strata for (**A**) 5 age groups and (**B**) 10 stocks, which are given generic designations, such as *Stock A*.

$$F^* = \sum_{s=1}^{S} \pi^*_{Fs} p^*_s C_s. \tag{11}$$

The percentile bootstrap CI for the true number of wild females, $[L_F, U_F]$, is determined by finding the $100\alpha/2$ and $100(1-\frac{\alpha}{2})$ percentiles. Similarly, we calculated a bootstrap CI for the number of wild males. Changing the binomial described previously to a multinomial, we generated B sets of $(\pi^*_{11},\ldots,\pi^*_{A1}),\ldots,(\pi^*_{1B},\ldots,\pi^*_{AB})$ to obtain bootstrap CIs for the true number of wild fish of ages 1, …, A. We followed the same procedure for stocks 1, …, G and $A \times G$ ages by stock (ages within stocks). To ensure accuracy across all groups being evaluated at a particular time, joint CIs for numbers of wild fish by sex or age or stock were calculated by using the methods of Mandel and Betensky (2008).

Simulations

Although our estimators and CIs are straightforward, we did not know their statistical properties. We designed a simulation of the sampling process to examine the properties of sex, age, and stock estimators and CIs, using the methods defined previously to analyze each simulated sample. We set the total passage of steelhead similar to the SY2011 observed count (200,000 fish). We set parameter values for all binomial and multinomial distributions similar to those of the stratified estimates obtained from the SY2011 data (Suppl. Tables 1 and 2). The percentage of wild steelhead ranged from 20% to 50%. The trapping rate varied from 3% to 14%, and the subsampling rate varied from 35% to 100%. Abundance and composition of the simulated population varied over 27 temporal strata loosely based on the character of the wild steelhead run in Snake River during SY2011 (Fig. 1). For simplicity, age and stock proportions in the population were generated by assuming age and stock are independent variables; therefore, we multiplied age and stock proportions to find age-by-stock proportions of the steelhead run.

We generated 500 samples from the population in the following manner. First, we simulated number of trapped fish (t_s) by generating binomial samples for each time stratum with the number of binomial trials equal to the number of fish returning during that stratum and with probability equal to the proportion of fish trapped within that stratum. Second, we simulated the number of trapped fish that were wild for each stratum by generating binomial samples with the number of trials equal to t_s and with probability equal to the true proportion of wild fish for that stratum. The remaining trapped fish were of hatchery origin. From these numbers, we generated a sample of trapped fish with 2 columns: time stratum and wild versus hatchery. The length of this data set was the sum of the numbers of wild fish trapped across the time strata ($\sum_{s=1}^{S} t_s$). Third, we calculated the number of wild fish whose sex, age, and stock had been determined in each stratum by multiplying the simulated number of wild fish trapped by the proportion subsampled. These numbers by stratum were the number of binomial or multinomial trials for sex or age or genetic stock (r_s). For example, we found the number of sampled fish that were wild females for a stratum by generating binomial trials of size r_s with probability equal to the true proportion of wild females during that stratum and with the remainder being males.

Knowing the random number of wild females and males trapped in each stratum, we put together a simulated subsample of fish by sex by forming a data set with two columns (for stratum and sex). The size of this sample was equal to the sum across the time strata of the numbers of handled fish, $\sum_{s=1}^{S} r_s$. Similar samples were simulated for age, stock, and stock by age (500 for each). The simulation generated 2 types of data from each sampling iteration: 1) a randomly generated trap sample with a random number of wild and hatchery fish, and 2) a randomly generated compositional sample with random numbers of females and males or numbers of fish of various ages or numbers of fish of various stocks.

Table 2

Number of simulations in which criteria for coverage levels for individual confidence intervals coverage were not met for combinations of estimator type (scenario). Simulations were conducted for 4 variables of interest with varying numbers of categories (k).

		Variable of interest (number of categories)			
Scenario	Criterion	Sex (k=2)	Brood year (k=5)	Stock (k=10)	Age×stock (k=50)
Pooled asymptotically normal					
	Coverage ≤0.85	0	1	5	18
	Coverage ≤0.80	0	0	3	6
Pooled parametric bootstrap					
	Coverage ≤0.85	0	0	5	13
	Coverage ≤0.80	0	0	3	2
Stratified asymptotically normal					
	Coverage ≤0.85	0	1	0	12
	Coverage ≤0.80	0	0	0	3
Stratified parametric bootstrap					
	Coverage ≤0.85	0	0	0	8
	Coverage ≤0.80	0	0	0	0

We obtained estimates of abundance and bootstrapped CIs for the groups of interest for every simulation iteration, using the window counts and the primary and secondary samples, as explained previously. After 500 simulation iterations, we had 500 estimates of the total number of wild steelhead, female and male wild, wild by age, and wild by stock. We also had 500 individual and joint CIs for each estimate. We saved the estimates and CIs for subsequent evaluation.

The evaluation of estimator performance was based on bias and CI coverage. We computed bias as the mean of the simulated estimates minus the true value. We computed the coverage of any individual CI by tallying the number of times the true population number fell inside the CI. For the pooled and stratified estimators, we tallied the number of cases for which the bias was ≥5%. Likewise, we tallied the number of cases for which the coverages for estimators were ≤0.85 (considered poor) and ≤0.80 (considered very poor).

Analysis of SY2011 data

We evaluated precision using the preferred estimator (determined from the simulations) that was applied to real data from SY2011; these data had more irregularities than the simulated data. For this application, we determined the age-by-stock proportions from the data, not as the product of age and stock proportions. We measured precision as the half-width of a (1-α)100% CI expressed as a percentage of the point estimate (P_{ind} for individual CIs or P_{joi} values for joint CIs). Researchers often set a stringent goal of a CI half-width ≤10% of the estimate. For management purposes, it is recommended that salmon stocks have unbiased abundance estimates with a coefficient of variation of 15% or less (Crawford and Rumsey[3]). For a 90% asymptotic CI (which indicates a critical value of the t distribution at 1.645), it follows that

$$\left|\widehat{W}-W\right| \leq 1.645se \leq 1.645(0.15)\widehat{W} \leq 0.25\widehat{W} \tag{12}$$

or

$$\left|\widehat{W}-W\right| / \widehat{W} \leq 0.25 \tag{13}$$

(i.e., half of the width of the CI interval should be ≤25% of the estimate). We compare the P_{ind} and P_{joi} values for all CIs with 0.10 and 0.25. To determine whether P_{ind} was related to the number of fish sampled or estimated size of the target group, we fitted power curves, using the results from all cases.

Results

Simulations

Performance of the pooled and stratified estimators was similar when the variable of interest had few categories, but the stratified estimators did better as complexity increased (Table 2). Detailed simulation results are provided in Suppl. Tables 3 and 4. All estimators produced acceptable accuracy and CI coverage when numbers of wild fish were estimated by sex. Similarly, all estimators provided acceptable accuracy

[3] Crawford, B. A., and S. M. Rumsey. 2011. Guidance for monitoring recovery of Pacific Northwest salmon and steelhead listed under the federal Endangered Species Act, 117 p. Northwest Region, National Marine Fisheries Service, NOAA. [Available at website.]

Table 3

Sample size (*n,* number of steelhead), abundance estimate, and individual and joint confidence interval half-width (%) for groups of wild steelhead trout (*Oncorhynchus mykiss*) that spawned in 2011. Values are given for groups defined by sex, brood year (BY), or stock (identified by the location where the stock spawns).

Group	*n*	Abundance	Individual	Joint
Total wild fish	4701	44,133	2.3	2.8
Females	1466	29,541	2.4	2.8
Males	732	14,592	2.8	3.2
BY2004	38	784	8.4	12.3
BY2005	520	11,239	3.6	4.8
BY2006	994	21,449	2.6	3.5
BY2007	473	10,103	3.7	5.0
BY2008	26	558	10.1	13.8
Grande Ronde	472	9442	7.1	11.9
Imnaha	168	3318	11.9	21.0
Lower Clearwater	173	3421	12.4	20.3
Lower Salmon	98	1941	16.3	26.5
Lower Snake	219	4374	10.3	17.5
Middle Fork Salmon	214	4312	10.8	15.9
South Fork Clearwater	233	4228	10.4	15.8
South Fork Salmon	135	2512	13.8	20.6
Upper Clearwater	215	3885	11.4	18.2
Upper Salmon	340	6699	8.1	12.8

when numbers by age were estimated, but the asymptotically normal CIs had poor coverage in one case for each estimator type. Average CI coverage among stocks was similar between the pooled estimators: 87.7% for the pooled asymptotically normal estimator and 88.2% for the pooled bootstrap estimator. Average CI coverage was slightly higher for the stratified estimators: 88.1% for the stratified asymptotically normal estimator and 89.0% for the stratified bootstrap estimator. The pooled estimators had unacceptable bias and very poor CI coverage for 3 of the 10 stocks, whereas the stratified estimators had acceptable accuracy for all stocks. Average CI coverage among stocks was similar for the pooled estimators: 81.5% for the pooled asymptotically normal estimator and 82.2% for the pooled bootstrap estimator. In contrast, average CI coverage was higher for the stratified estimators, although it was similar between them: 88.4% for the stratified asymptotically normal estimator and 89.0% for the stratified bootstrap estimator.

Problems with pooled estimators became even more prevalent when we addressed age by stock; however, the performance of the stratified estimators also began to suffer as the number of groups to be estimated increased to 50 (Table 2). The pooled estimators had unacceptable levels of bias in 21 cases, whereas the stratified estimators had unacceptable bias in 3 cases. Poor performance was most common in groups composed of steelhead from the least abundant BYs. Instances of poor CI coverage were usually, but not always, associated with unacceptably high bias. Overall, stratified estimators performed better than pooled estimators. Further, the bootstrap CIs had better coverage than the asymptotically normal CIs; in 3 instances, asymptotically normal CIs had very poor coverage (<80%), but there were no such instances for the bootstrap CIs. For this reason, we applied the stratified bootstrap estimator to the SY2011 data to develop guidelines for sampling and interpretation of such data.

Application of the stratified bootstrap estimator to data from SY2011

During SY2011, 208,296 steelhead were counted at LGD. Of these fish, 44,133 steelhead were estimated to be wild (21.2%, Table 3). The 90% CI was 43,152–45,140 wild steelhead. There were approximately twice as many females as males. Sex ratio varies annually, but the ratios seen in 2011 were typical. The middle age groups had more returning fish than the youngest and oldest age groups. Stocks were not evenly represented (e.g., the GRR stock had almost 4 times the number as the LOS stock) (Table 4). There were 46 stock-by-age groups in SY2011; estimated abundance ranged from 4912 individuals in the GRR stock in BY2006 to 21 individuals in the LSN stock in BY2004 (Table 4). The composition of a real steelhead run was not as balanced as that in the simplification used to generate the simulated data in this study, and this uneven distribution was most apparent in the age-by-stock groups.

Effects of using individual versus joint CIs depended on the complexity (i.e., number of groups) in the vari-

Table 4

Sample size (*n,* number of steelhead), abundance estimate, and individual and joint confidence interval half-width (%) for age groups (brood years) within stocks of wild steelhead trout (*Oncorhynchus mykiss*) that spawned in 2011. Groups are identified by brood year (BY) and the location where the stock spawns.

Group	*n*	Abundance	Individual	Joint
Grande Ronde BY2004	6	116	19.0	34.1
Grande Ronde BY2005	67	1501	8.6	16.2
Grande Ronde BY2006	222	4912	7.4	12.8
Grande Ronde BY2007	133	2875	7.7	14.2
Grande Ronde BY2008	2	38	31.9	54.5
Imnaha BY2004	2	43	35.1	63.3
Imnaha BY2005	22	537	17.2	32.0
Imnaha BY2006	82	1781	12.9	23.7
Imnaha BY2007	43	957	13.8	26.0
Lower Clearwater BY2004	1	32	46.8	84.2
Lower Clearwater BY2005	24	522	14.7	25.7
Lower Clearwater BY2006	78	1693	12.7	23.2
Lower Clearwater BY2007	49	1097	13.5	24.2
Lower Clearwater BY2008	4	78	30.2	55.3
Lower Salmon BY2004	1	23	57.5	106.1
Lower Salmon BY2005	18	409	19.5	35.4
Lower Salmon BY2006	40	909	17.9	30.6
Lower Salmon BY2007	26	557	17.9	31.1
Lower Salmon BY2008	2	44	48.3	87.6
Lower Snake BY2004	1	21	50.8	91.9
Lower Snake BY2005	13	324	16.2	29.5
Lower Snake BY2006	99	2268	11.4	20.2
Lower Snake BY2007	72	1614	11.4	29.4
Lower Snake BY2008	6	147	23.1	40.5
Middle Fork Salmon BY2004	10	222	16.2	`31.4
Middle Fork Salmon BY2005	99	2245	11.7	21.7
Middle Fork Salmon BY2006	72	1543	11.8	23.2
Middle Fork Salmon BY2007	6	303	22.1	40.4
South Fork Clearwater BY2004	3	59	28.2	50.3
South Fork Clearwater BY2005	44	906	12.1	22.5
South Fork Clearwater BY2006	144	3010	10.5	19.6
South Fork Clearwater BY2007	12	254	22.6	40.9
South Fork Salmon BY2004	6	132	28	51.8
South Fork Salmon BY2005	75	1612	14.3	25.2
South Fork Salmon BY2006	31	695	18.3	33.5
South Fork Salmon BY2007	3	73	34.2	62.2
Upper Clearwater BY2004	5	99	25.1	47.8
Upper Clearwater BY2005	120	2425	11.0	20.1
Upper Clearwater BY2006	63	1283	12.2	22.7
Upper Clearwater BY2007	3	54	29.8	54.0
Upper Clearwater BY2008	1	24	53.0	97.6
Upper Salmon BY2004	2	39	29.7	54.2
Upper Salmon BY2005	35	758	10.6	21.0
Upper Salmon BY2006	156	3356	8.8	16.3
Upper Salmon BY2007	107	2319	9.1	17.9
Upper Salmon BY2008	10	227	15.4	28.1

able of interest (Table 3). Widths of the joint CIs for females and males were not markedly wider than those of the individual CIs. For age, the widths of the joint CIs were 1.2–3.4 times the widths of the individual CIs. For stock, the joint CIs were 1.2–2.5 times wider. For some age-by-stock groups, the joint CIs were considerably wider than those of the individual CIs (Table 4).

Attainment of precision goals depended on the group and whether individual or joint CIs were considered (Tables 3 and 4). All sex and age groups met the 10%

research goal for precision if P_{ind} was used, except for the BY2008 group (P_{ind}=10.1%). With the use of P_{joi}, the 10% goal was met for all sex and age groups, except the BY2004 and BY2008 groups. Stock groups did not meet the 10% goal, except the GRR and UPS stocks, if P_{ind} was used. All sex, age, and stock groups met the 25% management goal for individual and joint CIs, except the LOS stock if P_{joi} was used. Of the stock by age groups, 11% met the 10% precision goal if P_{ind} was used, but none met the goal if P_{joi} was used. There was wider disparity in attainment of the 25% goal; 70% of the age-by-stock groups met the goal if P_{ind} was used, but only 35% of the age-by-stock groups met the goal if P_{joi} was used.

In general, half-widths of the CIs declined and stabilized as the number of fish sampled and estimated abundance of the groups involved increased (Fig. 2). Precision scaled approximately with the cube root of sample size, indicating that reducing the CI width by half would require approximately 8 times as many samples. Values of P_{ind} within a percentage point of the 10% precision criterion were obtained when there were 26–233 fish from a given category in the subsample (mean=140) and when there were 558–4374 steelhead in the category (mean=2794). Values of P_{ind} closest to the 25% criterion were obtained when there were 5 and 6 fish in a subsample and when there were 99–147 steelhead in a category. The power functions parameterized from the group estimates yielded values of 96 samples and 1982 steelhead at the 10% criterion and 7 samples and 147 steelhead at the 25% precision criterion.

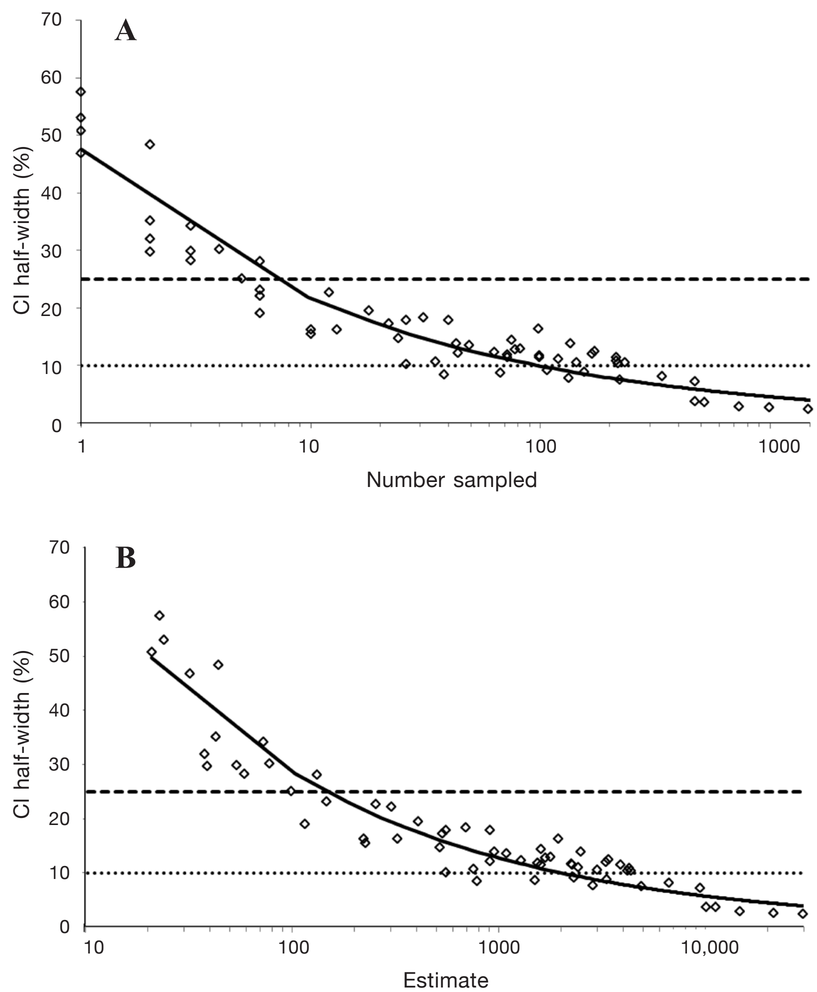

Figure 2

Relationship of confidence interval (CI) half-width for groups of wild steelhead (*Oncorhynchus mykiss*) spawning in 2011 in Snake River, Pacific Northwest, to (**A**) number of fish sampled (CI half-width=47.48[number sampled]$^{-0.341}$, coefficient of multiple determination [R^2]=0.872), and (**B**) abundance estimate (CI half-width=144.74[abundance]$^{-0.352}$, R^2=0.870). Precision criteria levels of 10% (dotted line) and 25% (dashed line) CI half-widths are shown for reference.

Discussion

The stratified estimators had biases <5%, except for a few cases in the most complex analysis (age by stock). Individual CIs for most constituent groups had coverages very near the nominal 90%. For conservation assessments, the greatest need is data on abundance of each stock. Coverage of CIs was good even for the smallest stock. Age structure is important for computation of productivity for each stock (i.e., for summarizing the adult progeny from a brood year returning across years and for computing progeny per parent). Accuracy of productivity estimates typically are largely controlled by the most abundant age groups, which had unbiased estimates in our study. When the stratified estimators were used, only a few of the smallest stock-by-age groups had bias ≥5.0% or CI coverage <85%.

Conversely, pooled estimators performed poorly except for the simplest variables of interest: sex and age. As variable complexity increased from age (5 cases) to stock (10 cases) to age-by-stock (50 cases), the proportion of estimates biased >5% increased from 0% to 30% to nearly half, respectively. Initially, we expected pooled estimators to be acceptable because of the highly controlled sampling regime at LGD, which was very consistent for SY2011. The realized sampling rate in the simulation (the product of trap rate and subsample rate) averaged 4.7% (standard deviation 1.3%). However, stock and age composition changed through the run, as did the number of steelhead crossing the dam. As the variable of interest became more complex, accuracy and precision of pooled estimators decreased.

Steinhorst, et al. (2010) estimated the run composition of fall-run Chinook salmon (*O. tshawytscha*) at LGD on the basis of counts from observation windows from 18 August through 15 December (their method 1). They used a stochastic model based on a fast Fourier transform to model the distribution of daily window counts, which were summed to obtain total abundance. Steinhorst, et al. (2010) used 2 bootstrap steps—a nonparametric bootstrap associated with the Fourier model and a parametric bootstrap applied to an estimate of composition pooled over the season. They did not report composition by stratum because composition was calculated with a complex accounting algorithm that could not be applied to individual strata. In essence, they assumed that either the proportions of their sex-by-age-by-origin groups were fairly uniform over the season or that a constant proportion of the run was sampled for each stratum. However, if the groups of interest returned at different times, a pooled estimate of composition applied to total escapement would not be accurate, especially over longer temporal spans (e.g., the steelhead run). In our study, the simulation results from the pooled estimators indicated precisely that outcome.

Precision may be computed for each group of interest, one at a time (i.e., P_{ind}), or more conservatively across all groups within a variable of interest (P_{joi}), minimizing study-wide error. However, the conservative approach resulted in wider CIs; for example, joint CIs were 14–17% wider than the individual CIs for sex in the SY2011 run. For stocks, P_{joi} values were about 47–76% wider than P_{ind} CIs. Given the number of stocks, we were not paying a large penalty for computing joint CIs. For age-by-stock groups, the joint CIs were on average 85% wider than the individual CIs (range: 71–158%). This difference likely was due to the uneven distribution of numbers by age and the large number of age-by-stock groups. Because we were trying to achieve joint coverage across so many groups simultaneously, a much greater expansion of the CIs was necessary. The results of our study show the cost to statistical power caused by the inclusion of many groups in an analysis. Investigators must consider whether the more conservative approach affects the usefulness of the resulting estimates and which groups are truly of management interest. For the latter consideration, investigators may combine some groups or decide that loss of precision is acceptable for their application. In our case, we combined strata to achieve greater sample sizes and used total age rather than the combinations of years spent in freshwater and years spent in saltwater that salmon biologists often use (Quinn, 2005).

Precision is related to the amount of information available, and the quality of this information declines as group size becomes smaller or as the realized sample rate is reduced. The problem in our case was that the steelhead run in Snake River is protracted over time, compounded by the complexity of the life history and stock structure of steelhead. Therefore, multinomial proportions must be estimated for many groups over many time strata unless the groups of interest can be simplified. Even so, we generally met the research goal of half the 90% CI width within 10% of the estimate for sex and age groups present in SY2011. For stock groups, we met the management goal of half the 90% CI width within 25% of the estimate for sex and age groups but our 10% precision goal was not attained, except with the 2 largest stocks when the less stringent P_{ind} measure was used.

To develop guidance for interpretation of the estimates we obtained from the data, we relied on P_{ind} because it was not affected by the number of other groups in the analysis. Precision of the estimates for individual groups declined rapidly when group abundance was <2500 individuals or when <100 individuals from that group were collected in the subsample. However, if there were few fish in a group, analysts and managers probably would be content with a more lenient precision criterion. For example, if our estimate was 50 and the CI was 20–80, the percent half-width would be 60% of the estimate but the fact that the true number is between 20 and 80 should be sufficiently precise for management purposes, especially if the numbers of fish in other groups are decidedly larger. With P_{ind} as a measure of precision, the 10% research precision goal could be reached if group abundance were to exceed 2000 individuals or if >100 samples from that group were collected. The 25% management precision goal was much more attainable and was achieved at group abundances >150 individuals and when very few samples were collected (<10). These values can be used as thresholds for the lenient precision criterion in our application.

Our results have implications for monitoring fish populations. If the interest is on the largest groups in a mixed population, most sampling programs will yield sufficient results. However, weak stocks are frequently a problem for conservation and fisheries management, and precision of abundance estimates of smaller groups becomes important. Obviously, there is a tradeoff between sample size and number of subdivisions that can be maintained. Previous work by Gerritsen and McGrath (2007) has supported this notion, but their criteria for success focused on overall (average) precision. Thompson (1987) found that a sample size of 510 fish should suffice under a worst case scenario for α = 0.05 (equal proportions among groups, but number of groups does not matter) as long as desired precision is expressed in absolute terms. If desired precision is expressed in relative terms (as was done in our study), no sample size will be sufficient if group size approaches zero. However, our results provide useful guidance for determining the groups that will have reliable estimates in complex scenarios. Even at lower fish abundances, we can define a lenient precision criterion with practical value.

The estimation approach described in this article is very flexible and can be customized for many scenarios. In this study, we assumed that window counts were a census, and stock and age were determined without error. In most applications, abundance is determined from a sample rather than a census. Further, there is uncertainty in the determination of stock and age for each fish. Additional uncertainty (e.g., genetic variability or noncensus estimates of total abundance) can be incorporated into our framework by adding additional bootstrap steps to reflect the source of the additional variance (e.g., method 2 in Steinhorst et al., 2010). Another practical issue is that samples for genetic analysis can now be processed en masse, but ages must be read from scale samples individually; hence, more fish can be identified to stock than can be aged. The bootstrap routine can be altered such that all genetic information is used to estimate stock abundance, and age composition is applied within each stock estimate (i.e., age composition is conditional on stock).

The stratified estimator in our study produced unbiased estimates, and the parametric bootstrap CIs had good coverage and acceptable precision. In complex scenarios, estimates of abundance of small groups will have poor precision and some may be biased, but a stratified estimate with a conservative joint CI can be of practical use if the numbers of fish in other groups are much larger. The 2-step bootstrap approach is flexible and can be adapted to incorporate other sources of variability or sampling constraints.

Acknowledgments

Funding for this project was provided by Bonneville Power Administration under project numbers 1990-055-00, 1991-073-00, and 2010-026-00. M. Campbell reviewed earlier drafts of this article. We gratefully acknowledge the thorough critiques given by 3 anonymous reviewers.

Literature cited

Anderson, E. C.
2010. Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. Mol. Ecol. Resour. 10:701–710. Article

Anderson, E. C., R. S. Waples, and S. T. Kalinowski.
2008. An improved method for predicting the accuracy of genetic stock identification. Can. J. Fish. Aquat. Sci. 65:1475–1486. Article

Campbell, M. R., C. C. Kozfkay, T. Copeland, W. C. Schrader, M. W. Ackerman, and S. R. Narum.
2012. Estimating abundance and life history characteristics of threatened wild Snake River steelhead stocks by using genetic stock identification. Trans. Am. Fish. Soc. 141:1310–1327. Article

Ford, M. J. (ed.).
2011. Status review update for Pacific salmon and steelhead listed under the Endangered Species Act: Pacific Northwest. NOAA Tech. Memo. NMFS-NWFSC-113, 281 p.

Gerritsen, H. D., and D. McGrath.
2007. Precision estimates and suggested sample sizes for length-frequency data. Fish. Bull. 105:116–120.

Goodman, L. A.
1960. On the exact variance of products. J. Am. Stat. Assoc. 55:708–713. Article

Harmon, J. R.
2003. A trap for handling adult anadromous salmonids at Lower Granite Dam on the Snake River, Washington. North Am. J. Fish. Manage. 23:989–992. Article

Hughes, J. B., and J. E. Hightower.
2015. Combining split-beam and dual-frequency identification sonars to estimate abundance of anadromous fishes in the Roanoke River, North Carolina. North Am. J. Fish. Manage. 35:229–240. Article

Latremouille, D. N.
2003. Fin erosion in aquaculture and natural environments. Rev. Fish. Sci. 11:315–335. Article

Mandel, M., and R. A. Betensky.
2008. Simultaneous confidence intervals based on the percentile bootstrap approach. Comput. Stat. Data Analys. 52:2158–2165. Article

Pella, J. J., and G. B. Milner.
1987. Use of genetic marks in stock composition analysis. *In* Population genetics and fisheries management (N. Ryman and F. Utter, eds.), p. 274–276. Univ. Wash. Press, Seattle, WA.

Pritt, J. J., M. R. DuFour, C. M. Mayer, P. M. Kocovsky, J. T. Tyson, E. J. Weimer, and C. S. Vandergroot.
2013. Including independent estimates and uncertainty to quantify total abundance of fish migrating in a large river system: walleye (*Sander vitreus*) in the Maumee River, Ohio. Can. J. Fish. Aquat. Sci. 70:803–814. Article

Quinn, T. P.
2005. The behavior and ecology of Pacific salmon and trout, 388 p. Univ. Wash. Press. Seattle, WA.

Raymond, H. L.
1988. Effects of hydroelectric development and fisheries enhancement on spring and summer Chinook salmon and steelhead in the Columba River basin. North Am. J. Fish. Manage. 8:1–24. Article

Rudstam, L. G., J. M. Jech, S. L. Parker-Stetter, J. K. Horne, P. J. Sullivan, and D. M. Mason.
2013. Fisheries acoustics. *In* Fisheries techniques, 3rd ed. (A. V. Zale, D. L. Parrish, and T. M. Sutton, eds.), p. 597–636. Am. Fish. Soc., Bethesda, MD.

Shaklee, J. B., T. D. Beacham, L. Seeb, and B. A. White.
1999. Managing fisheries using genetic data: case studies from four species of Pacific salmon. Fish. Res. 43:45–78. Article

Steele, C. A., E. C. Anderson, M. W. Ackerman, M. A. Hess, N. R. Campbell, S. R. Narum, and M. R. Campbell.
2013. A validation of parentage-based tagging using hatchery steelhead in the Snake River basin. Can. J. Fish. Aquat. Sci. 70:1046–1054. Article

Steinhorst, K., D. Milks, G. P. Naughton, M. Schuck, and B. Arnsberg.
2010. Use of statistical bootstrapping to calculate confidence intervals for the fall Chinook salmon run reconstruction to Lower Granite Dam. Trans. Am. Fish. Soc. 139:1792–1801. Article

Tarbox, K. E., and R. E. Thorne.
1996. Assessment of adult salmon in near-surface waters of Cook Inlet, Alaska. ICES J. Mar. Sci. 53:397–401. Article

Thompson, S. K.
1987. Sample size for estimating multinomial proportions. Am. Stat. 41:42–46. Article

Wagner, P. G.
2007. Fish counting at large hydroelectric projects. *In* Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations (D. H. Johnson, B. M. Shrier, J. S. O'Neal, J. A. Knutzen, X. Augerot, T. A. O'Neil, T. N. Pearsons, eds.), p. 173–195. Am. Fish. Soc., Bethesda, MD.

Wood, C. C., S. McKinnell, T. J. Mulligan, and D. A. Fournier.
1987. Stock identification with the maximum-likelihood mixture model: sensitivity analysis and application to complex problems. Can. J. Fish. Aquat. Sci. 44:866–881. Article

Supplementary Table 1

Values used to simulate the sampling process at Lower Granite Dam, in Washington, in order to examine the properties of estimators of sex and age composition. Values are given by weekly time stratum for total number of steelhead trout (*Oncorhynchus mykiss*) passing the dam, the trap rate (proportion of fish sampled per stratum), the proportion of wild fish in the sample, the proportion of wild fish subsampled per stratum, and the simulated proportions of wild fish by sex and age.

Time stratum	Total number	Trap rate	Proportion of wild fish	Sub-sample rate	Proportion of females	Proportion by age				
						Age 7	Age 6	Age 5	Age 4	Age 3
1	1957	0.04	0.25	0.96	0.82	0.05	0.20	0.60	0.15	0.00
2	2631	0.04	0.25	0.96	0.82	0.05	0.10	0.40	0.45	0.00
3	2559	0.04	0.25	0.90	0.82	0.00	0.15	0.50	0.30	0.05
4	1627	0.04	0.30	0.96	0.82	0.00	0.30	0.25	0.45	0.00
5	2953	0.03	0.30	1.00	0.58	0.00	0.05	0.50	0.40	0.05
6	5074	0.03	0.30	0.96	0.58	0.00	0.20	0.50	0.25	0.05
7	6784	0.07	0.25	0.35	0.58	0.00	0.20	0.50	0.25	0.05
8	9138	0.12	0.25	0.38	0.71	0.02	0.25	0.50	0.20	0.03
9	16,686	0.12	0.20	0.37	0.63	0.02	0.30	0.50	0.15	0.03
10	24,056	0.12	0.20	0.40	0.75	0.03	0.25	0.50	0.20	0.02
11	30,062	0.12	0.20	0.45	0.69	0.03	0.30	0.45	0.20	0.02
12	25,408	0.12	0.20	0.45	0.65	0.03	0.25	0.45	0.25	0.02
13	18,781	0.12	0.20	0.50	0.59	0.03	0.27	0.45	0.22	0.03
14	17,532	0.11	0.20	0.50	0.67	0.02	0.21	0.50	0.25	0.02
15	8122	0.11	0.20	0.50	0.59	0.03	0.25	0.50	0.20	0.02
16	6462	0.11	0.25	0.50	0.59	0.02	0.20	0.55	0.21	0.02
17	2193	0.13	0.25	0.50	0.59	0.00	0.35	0.40	0.22	0.03
18	2549	0.10	0.25	0.50	0.59	0.00	0.20	0.50	0.27	0.03
19	3556	0.03	0.20	0.50	0.664	0.00	0.30	0.30	0.40	0.00
20	986	0.09	0.20	0.50	0.664	0.00	0.15	0.50	0.35	0.00
21	1516	0.14	0.30	0.45	0.664	0.00	0.10	0.70	0.20	0.00
22	1379	0.13	0.30	0.50	0.664	0.00	0.25	0.50	0.25	0.00
23	2447	0.10	0.30	0.50	0.664	0.00	0.10	0.50	0.40	0.00
24	1754	0.13	0.30	0.45	0.664	0.00	0.40	0.35	0.25	0.00
25	1537	0.12	0.30	0.50	0.664	0.00	0.20	0.40	0.40	0.00
26	860	0.11	0.30	0.50	0.664	0.00	0.10	0.60	0.30	0.00
27	1391	0.10	0.50	0.50	0.664	0.00	0.30	0.50	0.20	0.00

Supplementary Table 2

Proportions of 10 simulated stocks by time stratum used to simulate numbers by stock of steelhead (*Oncorhynchus mykiss*) trapped at Lower Granite Dam in Washington.

Week	Stock A	Stock B	Stock C	Stock D	Stock E	Stock F	Stock G	Stock H	Stock I	Stock J
1	0.20	0.15	0.05	0.00	0.00	0.00	0.05	0.00	0.30	0.25
2	0.28	0.12	0.08	0.04	0.00	0.00	0.04	0.00	0.20	0.24
3	0.2381	0.1905	0.0952	0.0476	0.00	0.00	0.00	0.0476	0.1429	0.2381
4	0.2414	0.2069	0.069	0.00	0.0345	0.0345	0.0345	0.0345	0.1034	0.2414
5	0.1714	0.1143	0.0857	0.0286	0.0286	0.0571	0.0286	0.0286	0.2286	0.2286
6	0.1852	0.0926	0.0185	0.0556	0.037	0.037	0.0556	0.037	0.2407	0.2407
7	0.2063	0.1111	0.0476	0.0159	0.0317	0.0476	0.0317	0.0476	0.2222	0.2381
8	0.2393	0.1718	0.0552	0.0429	0.0123	0.0245	0.0368	0.0798	0.1166	0.2209
9	0.1508	0.1587	0.1032	0.0556	0.0278	0.0317	0.0278	0.0556	0.1786	0.2103
10	0.1373	0.1701	0.1015	0.0418	0.0507	0.0567	0.0537	0.0746	0.1194	0.194
11	0.1585	0.0839	0.0839	0.035	0.1119	0.1049	0.042	0.0746	0.1259	0.1795
12	0.1875	0.0938	0.0455	0.0341	0.1392	0.108	0.0341	0.0511	0.1364	0.1705
13	0.1167	0.0333	0.025	0.0542	0.1542	0.1625	0.05	0.05	0.1667	0.1875
14	0.1138	0.0163	0.0244	0.0447	0.1301	0.1626	0.0407	0.0691	0.1545	0.2439
15	0.1563	0.0234	0.0156	0.0156	0.1641	0.1641	0.0391	0.0547	0.1797	0.1875
16	0.1354	0.0313	0.0104	0.0208	0.1354	0.125	0.0625	0.0417	0.2083	0.2292
17	0.0784	0.0196	0.0196	0.0196	0.1176	0.1765	0.0392	0.0784	0.1765	0.2745
18	0.1463	0.0244	0.0244	0.0244	0.0976	0.1463	0.0244	0.0976	0.122	0.2927
19	0.0476	0.00	0.00	0.00	0.1429	0.1429	0.0476	0.00	0.2381	0.381
20	0.0357	0.00	0.00	0.0357	0.1071	0.25	0.0357	0.0357	0.2143	0.2857
21	0.0455	0.00	0.00	0.0455	0.0455	0.1818	0.0909	0.00	0.2273	0.3636
22	0.0426	0.00	0.0213	0.0213	0.1702	0.1277	0.0213	0.0426	0.234	0.3191
23	0.0541	0.027	0.027	0.00	0.1892	0.2162	0.00	0.027	0.1351	0.3243
24	0.0667	0.00	0.00	0.0333	0.1667	0.1667	0.0333	0.0333	0.1667	0.3333
25	0.0952	0.00	0.00	0.00	0.0952	0.1905	0.0476	0.0476	0.1429	0.381
26	0.05	0.00	0.00	0.00	0.05	0.10	0.05	0.05	0.35	0.35
27	0.04	0.04	0.00	0.00	0.16	0.08	0.04	0.04	0.28	0.32

Supplementary Table 3

Percent bias by estimator type (pooled or stratified) of all sex, age, stock, and stock-by-age groups determined from simulations of the sampling process at Lower Granite Dam in Washington. Bias was computed as the mean of the estimates from 500 simulations minus the true value. Absolute percent bias values >5% are presented in bold type.

	Estimator type			Estimator type			Estimator type	
Group	Pooled	Stratified	Group	Pooled	Stratified	Group	Pooled	Stratified
Female	−0.322	−0.058	Age 7, Stock G	**10.17**	4.11	Age 5, Stock J	−0.74	−0.33
Male	0.034	−0.841	Age 7, Stock H	2.46	−1.86	Age 4, Stock A	**−5.04**	0.24
Age 7	3.044	2.411	Age 7, Stock I	**5.50**	3.72	Age 4, Stock B	**−7.07**	0.65
Age 6	0.064	0.354	Age 7, Stock J	4.82	2.56	Age 4, Stock C	−3.39	0.52
Age 5	0.113	0.028	Age 6, Stock A	−0.76	0.96	Age 4, Stock D	−0.48	1.34
Age 4	−1.57	−1.398	Age 6, Stock B	−2.30	2.04	Age 4, Stock E	**5.09**	−2.21
Age 3	−1.35	−1.255	Age 6, Stock C	0.27	0.48	Age 4, Stock F	**5.53**	−1.96
Stock A	−2.767	0.97	Age 6, Stock D	1.62	0.51	Age 4, Stock G	−2.25	−1.08
Stock B	**−6.047**	1.115	Age 6, Stock E	**6.70**	−0.38	Age 4, Stock H	3.40	1.37
Stock C	−2.691	0.809	Age 6, Stock F	**7.01**	0.35	Age 4, Stock I	**−5.09**	−3.90
Stock D	1.224	0.672	Age 6, Stock G	1.79	1.22	Age 4, Stock J	−4.20	−2.67
Stock E	**8.009**	−0.047	Age 6, Stock H	4.46	2.16	Age 3, Stock A	**−8.54**	−0.46
Stock F	**7.087**	−0.251	Age 6, Stock I	0.13	−2.03	Age 3, Stock B	**−10.48**	3.43
Stock G	0.511	0.421	Age 6, Stock J	−1.19	−1.29	Age 3, Stock C	**−7.44**	−0.21
Stock H	2.963	1.985	Age 5, Stock A	−2.79	0.69	Age 3, Stock D	−4.65	0.66
Stock I	−2.876	−2.299	Age 5, Stock B	**−6.40**	1.21	Age 3, Stock E	**7.50**	1.09
Stock J	−1.633	−1.153	Age 5, Stock C	−3.90	−0.04	Age 3, Stock F	2.46	0.64
Age 7, Stock A	1.29	4.65	Age 5, Stock D	1.32	1.24	Age 3, Stock G	**−5.42**	2.29
Age 7, Stock B	**5.95**	−0.18	Age 5, Stock E	**7.85**	−0.42	Age 3, Stock H	−1.05	−0.02
Age 7, Stock C	**7.62**	−2.79	Age 5, Stock F	**7.57**	−0.31	Age 3, Stock I	**−9.62**	−4.11
Age 7, Stock D	**9.30**	4.82	Age 5, Stock G	0.86	−0.19	Age 3, Stock J	**−8.45**	−1.91
Age 7, Stock E	4.56	**13.57**	Age 5, Stock H	2.35	2.13			
Age 7, Stock F	1.16	**7.32**	Age 5, Stock I	−2.24	−1.22			

Supplementary Table 4

Confidence interval coverage by combinations of estimator (pooled or stratified) and confidence interval type (asymptotically normal or parametric bootstrap) for of all sex, age, stock, and stock-by-age groups based on simulations of the sampling process at Lower Granite Dam in Washington. Confidence interval coverage is the proportion of 500 simulations in which the true value was within the confidence interval. Bold type indicates poor coverage (<0.85). **Bold** italic type indicates very poor coverage (<0.80).

	Estimator and confidence interval type					Estimator and confidence interval type			
Group	Pooled asymp.	Pooled bootstrap	Stratified asymp.	Stratified bootstrap	Group	Pooled asymp.	Pooled bootstrap	Stratified asymp.	Stratified bootstrap
Female	0.896	0.906	0.886	0.894	Age 6, Stock H	0.898	0.898	0.886	0.89
Male	0.896	0.916	0.864	0.872	Age 6, Stock I	0.884	0.892	0.864	0.876
Age 7	0.884	0.882	0.888	0.896	Age 6, Stock J	0.91	0.91	0.904	0.906
Age 6	0.91	0.91	0.902	0.908	Age 5, Stock A	0.856	0.856	0.89	0.894
Age 5	0.88	0.89	0.894	0.90	Age 5, Stock B	**0.82**	**0.83**	0.90	0.898
Age 4	**0.836**	0.852	**0.85**	0.864	Age 5, Stock C	0.86	0.874	0.898	0.896
Age 3	0.876	0.874	0.87	0.882	Age 5, Stock D	0.876	0.892	0.878	0.878
Stock A	**0.846**	**0.846**	0.888	0.89	Age 5, Stock E	***0.788***	***0.798***	0.868	0.878
Stock B	***0.736***	***0.758***	0.888	0.898	Age 5, Stock F	**0.842**	**0.836**	0.898	0.91
Stock C	0.858	0.864	0.874	0.876	Age 5, Stock G	0.902	0.904	0.906	0.906
Stock D	0.854	0.854	0.884	0.89	Age 5, Stock H	0.908	0.914	0.882	0.89
Stock E	***0.694***	***0.706***	0.896	0.90	Age 5, Stock I	0.864	0.884	0.872	0.886
Stock F	***0.722***	***0.72***	0.894	0.91	Age 5, Stock J	0.904	0.914	0.896	0.898
Stock G	0.90	0.902	0.878	0.88	Age 4, Stock A	**0.83**	**0.848**	0.88	0.892
Stock H	0.87	0.874	0.90	0.896	Age 4, Stock B	**0.808**	**0.822**	0.89	0.892
Stock I	**0.814**	**0.832**	0.854	0.862	Age 4, Stock C	0.852	0.866	0.87	0.88
Stock J	0.858	0.866	0.888	0.894	Age 4, Stock D	0.89	0.894	0.886	0.908
Age 7, Stock A	0.894	0.884	0.862	0.88	Age 4, Stock E	0.892	0.896	0.898	0.898
Age 7, Stock B	**0.846**	0.874	**0.844**	0.872	Age 4, Stock F	0.878	0.878	0.884	0.902
Age 7, Stock C	***0.742***	0.908	**0.82**	0.878	Age 4, Stock G	0.874	0.904	0.906	0.916
Age 7, Stock D	**0.832**	**0.804**	**0.85**	**0.84**	Age 4, Stock H	0.89	0.888	0.876	0.886
Age 7, Stock E	**0.858**	0.89	0.876	0.886	Age 4, Stock I	**0.848**	0.852	0.852	0.866
Age 7, Stock F	**0.81**	**0.844**	**0.822**	**0.844**	Age 4, Stock J	0.866	0.876	0.866	0.874
Age 7, Stock G	0.872	**0.836**	**0.844**	**0.842**	Age 3, Stock A	0.86	0.86	0.88	0.89
Age 7, Stock H	***0.732***	0.896	***0.77***	**0.848**	Age 3, Stock B	**0.81**	0.882	0.86	0.874
Age 7, Stock I	0.896	0.882	0.86	0.876	Age 3, Stock C	***0.722***	0.902	***0.79***	**0.826**
Age 7, Stock J	0.864	0.888	0.89	0.884	Age 3, Stock D	**0.806**	***0.776***	**0.844**	**0.838**
Age 6, Stock A	0.86	0.866	0.884	0.89	Age 3, Stock E	0.90	0.896	**0.84**	0.864
Age 6, Stock B	0.856	0.866	0.892	0.906	Age 3, Stock F	***0.796***	**0.838**	**0.84**	0.866
Age 6, Stock C	0.89	0.902	0.88	0.89	Age 3, Stock G	**0.83**	**0.83**	0.88	0.872
Age 6, Stock D	0.892	0.918	0.9	0.908	Age 3, Stock H	***0.774***	0.92	***0.796***	**0.83**
Age 6, Stock E	0.872	0.862	0.868	0.882	Age 3, Stock I	0.876	0.876	**0.836**	**0.848**
Age 6, Stock F	0.852	0.846	0.904	0.902	Age 3, Stock J	**0.81**	**0.826**	0.882	0.88
Age 6, Stock G	0.892	0.896	0.89	0.894					