
A Simple Stratified Design for Mark–Recapture Estimation of Salmon Smolt Abundance

Stan R. Carlson, Lewis G. Coggins Jr., and Charles O. Swanton

Reprinted from the
Alaska Fishery Research Bulletin
Vol. 5 No. 2, Winter 1998

The Alaska Fishery Research Bulletin can found on the World Wide Web
at URL: <http://www.state.ak.us/adfg/geninfo/pubs/afrb/afrbhome.htm> .

A Simple Stratified Design for Mark–Recapture Estimation of Salmon Smolt Abundance

Stan R. Carlson, Lewis G. Coggins Jr., and Charles O. Swanton

ABSTRACT: We describe a mark–recapture (M–R) technique in which a stratified design and sampling at 1 or 2 stream locations are used to estimate the abundance of a migrating salmon smolt population. The method consists of counting smolts captured at a designated downstream site and releasing marked smolts back into the population at an upstream site. Marked smolts subsequently recovered at the downstream site are counted to estimate capture probability (trap efficiency), which is used to estimate smolt abundance for a segment of the population. This procedure is temporally stratified such that each trap efficiency trial is discretely paired with one capture period; this can typically be accomplished by releasing marked smolts at relatively short intervals (a few days) with little chance of recaptured fish occurring in later strata. This approach accounts for potential temporal changes in capture probability under a fairly modest assumption of stratum consistency. The method simplifies the generalized 2-sample stratified design and provides some important advantages: (1) because marking occurs in discrete intervals, personnel costs are substantially reduced; (2) because each release of marked smolts corresponds to one capture period, only one type of mark is needed, which greatly simplifies marking procedures and recapture tallying; and (3) when only one capture site is used, material costs are reduced by about half. We present approximately unbiased abundance and variance estimators of the total smolt population and develop a method of estimating the number of smolts to be marked. A parametric bootstrap technique for quantifying precision is also developed. An example of the method is given using the 1997 sockeye salmon *Oncorhynchus nerka* smolts migrating from Akalura Lake, Kodiak Island, Alaska. The Akalura Lake study included a weir count of smolts, which we used to evaluate the accuracy of the M–R estimate.

INTRODUCTION

Effective management of salmon stocks often includes enumeration of smolts, which represent the freshwater production of a stock prior to entering the ocean environment. Smolt abundance, often with ancillary variables such as smolt size, can be used to forecast adult recruitment (Foerster 1954; Peterman 1981; Fried

and Yuen 1987; Macdonald and Smith 1987). Abundance, age composition, and size of smolts can be used to assess management and enhancement actions, freshwater rearing capacity, and potential problems in early life history development and survival (Hyatt and Stockner 1985; Koenings and Burkett 1987a; Kyle et al. 1988; Thedinga et al. 1994; Edmundson et al. 1997). Smolt numbers can also be used to assess the stock–

Authors: STAN R. CARLSON is a biometrician for the Alaska Department of Fish and Game, Division of Commercial Fisheries, 34828 Kalifornsky Beach Road, Soldotna, Alaska 99669. LEWIS G. COGGINS JR., former fishery biologist for the Alaska Department of Fish and Game, Commercial Fisheries Division, Kodiak, is a research fisheries biologist with the Alaska Biological Science Center, 1011 E. Tudor Road, Anchorage, Alaska 99503. CHARLES O. SWANTON, former area fishery biologist for the Alaska Department of Fish and Game, Division of Commercial Fisheries, Kodiak, is now a regional management biologist for the Alaska Department of Fish and Game, Division of Sport Fisheries, 300 College Road, Fairbanks, Alaska 99701.

Acknowledgments: Dana Schmidt, former principal limnologist for the Alaska Department of Fish and Game, encouraged the completion of this work. Gary Kyle, Gary Fandrei, Bruce King, Patrick Shields, and Gary Todd — helpful discussions.

Project Sponsorship: This research was partially funded from Exxon Valdez Oil Spill Restoration projects 258 and 259 with the balance funded by the Alaska Department of Fish and Game.

recruit relationship without the added variation of ocean mortality and problems with estimating adult recruitment (Koenings and Burkett 1987b; Hume et al. 1996).

A Peterson mark–recapture (M–R) estimate applied just one time to a smolt emigration may violate at least one important assumption: constant probability of capture. This could be due to several factors, including changes in stream flow, temporal variation in the age structure and size of smolts, and changes in sampling methods by the researcher (e.g., altering trap configurations). One approach that overcomes this problem is a generalized 2-sample stratified design for M–R estimation, which has been investigated by several authors. Darroch (1961) developed the maximum likelihood theory, and Seber (1982: 431–439) described the estimation methods. Dempson and Stansbury (1991) applied the technique to estimate the abundance of an Atlantic salmon *Salmo salar* smolt population. The 2-sample design involves capturing, marking, and releasing smolts upstream and subsequently capturing smolts downstream, where marked and unmarked fish are tallied. The technique requires a marking procedure that indicates the day of release and continuous recapturing. The method originally required pooling releases and recaptures into strata (Darroch 1961; Dempson and Stansbury 1991), but Schwarz and Dempson (1994) developed a model that overcomes this necessity by estimating daily capture probabilities.

Although the 2-sample design has clear advantages over other methods, in many situations it may be cost-prohibitive (particularly personnel time) or impractical. Rawson (1984) described a simpler technique that uses 1 sampling site and trap efficiency trials discretely paired with capture periods. In this paper we describe a modified version of this model, that uses 1 or 2 sampling locations, which we refer to as a *simple stratified M–R design*. The technique accounts for potential temporal changes in capture probability with a fairly modest assumption of stratum consistency. We discuss the estimators presented by Rawson (1984), which we believe were incorrect, and present approximately unbiased abundance and variance estimators. We also develop a bootstrap technique for quantifying estimation precision, provide a procedure for determining sample size, and present a method of estimating short-term mark survival. An application of the method is given using sockeye salmon *Oncorhynchus nerka* smolts migrating from Akalura Lake, Kodiak Island, Alaska, in 1997. The Akalura Lake study included a weir count of the smolts that we used to evaluate the accuracy of the M–R estimate.

SAMPLING TECHNIQUE

We describe 2 approaches to the sampling, which result in separate estimators: (1) similar to Rawson (1984), 1 capture site is used for all smolt sampling; and (2) 2 capture sites are used, one upstream for marking and releasing and one downstream for capturing marked and unmarked smolt. For the 1-site method marked smolts are released at a designated upstream site that corresponds to the mark–release site for the 2-site method. The downstream site in the 2-site method corresponds to the capture site in the 1-site method. The sampling method is stratified such that each trap efficiency trial corresponds to one downstream capture period or stratum; recaptures are assumed to occur in only one stratum. Trap efficiency trials are usually performed approximately midway through each period. The recovery rate of marked smolts in each stratum and the total number of (downstream) captures are used to estimate the size of the smolt population for that segment of the migration.

Field techniques for the 1-site method consist of continuously counting all smolts captured at a designated site. About midway through a stratum, a random sample is marked and released back into the population upstream. Smolt collected for marking are transported to the release site and typically held for at least 1 h to assess condition and remove dead individuals. The number of recaptures are recorded at the sampling site. For the 2-site approach, smolts are captured, marked, and released about midway through a stratum at the upstream site (the actual release site may be nearby) and captures of all marked and unmarked smolts are counted at the downstream site. Because most salmon smolt emigrations occur primarily at night (Hoar 1951; Hartman et al. 1967; Solomon 1981), counting and marking generally take place the following day. The site selected for releasing marked smolts must be an adequate distance upstream to ensure complete mixing of marked and unmarked smolts. The procedure is repeated at some minimally appropriate interval to provide one-to-one correspondence of the trap efficiency trial and capture period. Depending on the migration rate of the particular species or stock and the distance between sites, as few as 2–3 d can ensure that all marked smolts have migrated past the downstream site.

For each stratum, the following 6 assumptions apply, which are modified slightly from the customary assumptions for the Peterson estimate (e.g., Seber 1982: 59). Assumption 1 — the population is closed; there-

fore, the abundance remains constant. This is probably met in most cases (any population changes are negligible); mortality observed during marking, capturing, and handling is censored or directly subtracted from the estimate. Assumption 2 — all smolts have the same probability of being marked or all smolts have the same probability of being examined for marks. This encompasses potential problems with random sampling, mixing of marked and unmarked smolts, and variable catchability (size selectivity, marking effects, etc.). Assumption 3 — the probability of capture is constant. This is more restrictive here than with the 2-sample method; i.e., the smaller the periods, the less likely this assumption will be violated. However, the estimator is unbiased if the estimated capture probability is unbiased for the average capture probability. Assumption 4 — marks are not lost between release and recovery. This concerns the survival of marked fish and mark retention; short-term mark survival can be estimated and incorporated into the population estimate. Assumption 5 — all marked smolts are reported on recapture. This should be met if the field crew is well trained and the marking method is adequate. Assumption 6 — all marked smolts released are either recovered or pass by the downstream capture site. This involves the problem of smolts occurring in subsequent strata; some knowledge of the migration habits of the particular species or stock may be necessary to meet this assumption.

ESTIMATION METHODS

Rawson's Analysis

Rawson (1984) described the 1-site method and based his estimators on a paper by Cochran (1978), who evaluated LaPlace's (1820) technique of estimating the population of France. Cochran defined the known total number of births in France as X , births in sampled communes as x , and the population size of the sampled communes as y . He then presented the ratio estimate of the population of France as $\hat{Y}_R = X(y/x)$, which is a Peterson estimate; the ratio x/y is the estimate of the birth rate, p . Characteristics of this estimator are well known. The analogy to M-R experiments is that X is the size of the marked sample, x is the number of recaptures, y is the total size of the second sample (Seber 1982: 104), and p is the proportion of marks in the population. For each stratum, Rawson initially defined the known number of marked releases as D , the number of recaptures as d , and the number of unmarked smolts in the total captures as n . The ratio d/D (recap-

ture rate) is the estimate of capture probability or trap efficiency. He then presented the ratio estimate of the number of smolts in the population as $\hat{N}_R = n(D/d)$, which is a Peterson estimate of the unmarked population (Seber 1982: 436). Rawson correctly assumed that marked smolts should not be included in the estimate because these fish were counted (included in n) prior to marking.

Rawson's errors were apparently initiated by a misinterpretation of Cochran's description of LaPlace's model. Rawson subsequently equated n with X and D with y , which is incorrect. Unfortunately, this mistake resulted in incorrect equations for Rawson's approximately unbiased estimators. In addition, it appears that Rawson misapplied Cochran's bias estimate. Assuming x is a binomial random variable with parameters $\{y, p\}$, Cochran showed that

$$E[\hat{Y}_R] \approx Y + X(1-p)/(yp^2).$$

The sample estimate of the bias, $X(y-x)/x^2$, should have been subtracted from the population ratio estimator, \hat{Y}_R . Rawson added this bias estimate to his ratio estimator, further compounding the notation error.

Notation

We define the following:

- h = stratum index (capture period and a corresponding trap efficiency trial),
- L = number of strata or periods ($h = 1, 2, \dots, L$),
- M_h = number of marked smolts released in stratum h (mortality censored),
- M = total number of marked smolts released ($= \sum M_h$),
- m_h = number of marked smolts recaptured in h ,
- u_h = number of unmarked smolts captured in h ,
- n_h = total number of smolts captured in h ($= m_h + u_h$),
- U_h = total smolt population size in h , excluding marked releases and minus observed mortality,
- N_h = total smolt population size in h , including marked releases ($= M_h + U_h$),
- U = total smolt population size, excluding marked releases ($= \sum U_h$),
- N = total smolt population size, including marked releases ($= \sum N_h$),
- p_h = population proportion of marked smolts in h ($= M_h/N_h$),
- e_h = smolt capture probability (trap efficiency) in h ($= n_h/N_h$), and

S_h = proportion of marked smolts in h that survive and retain their marks until passing by the downstream capture site.

The parameters L , M_h , and M are known, and the random variables are m_h , u_h , and n_h . For 1-site experiments all captures occur at a single location so that M and M_h refer to the upstream releases. We assume that observed smolt mortality is directly subtracted where appropriate. Specifically, the number of deaths counted during marking and releasing is censored from M_h , and observed mortality in the trap captures and during handling is subtracted from the estimate of U_h or N_h . We also develop estimators that take into account short-term mark mortality, in which case M_h and M are estimated.

Two-Site Experiments

The objective is to estimate N , the population of smolts including the marked releases. For each stratum, h , the fundamental estimator is therefore a Peterson estimate,

$$\tilde{N}_h = \frac{n_h M_h}{m_h},$$

which has well-known properties. The distribution of m_h , conditional on M_h and n_h , is a hypergeometric. Chapman (1951) has shown that \tilde{N}_h , although asymptotically the best estimate of N_h , is biased, and this bias can be large for small samples. His estimator,

$$\hat{N}_h = \frac{(n_h + 1)(M_h + 1)}{(m_h + 1)} - 1, \quad (1)$$

is exactly unbiased when $(M_h + n_h) \geq N_h$ and approximately unbiased when $(M_h + n_h) < N_h$ if $m_h \geq 7$ (Robson and Regier 1964; Seber 1982: 60). Seber (1970) derived an approximately unbiased estimate of the variance of N_h :

$$v(\hat{N}_h) = \frac{(M_h + 1)(n_h + 1)(M_h - m_h)(n_h - m_h)}{(m_h + 1)^2(m_h + 2)}. \quad (2)$$

Thus, for 2-site experiments the approximately unbiased stratum estimators are \hat{N}_h and $v(\hat{N}_h)$. The total smolt abundance estimate is

$$\hat{N} = \sum_{h=1}^L \hat{N}_h, \quad (3)$$

and the variance estimate is

$$v(\hat{N}) = \sum_{h=1}^L v(\hat{N}_h). \quad (4)$$

Because N is asymptotically normally distributed, an approximate 95% confidence interval is

$$\hat{N} \pm 1.96\sqrt{v(\hat{N})}. \quad (5)$$

One-Site Experiments

The objective is to estimate U because the total count of captures includes the sample of marked smolts. For each stratum, h , the fundamental estimator is therefore a Peterson estimate of the unmarked population,

$$\tilde{U}_h = \tilde{N}_h - M_h = \frac{u_h M_h}{m_h}.$$

This is the maximum likelihood estimate of U_h that treats the joint distribution of m_h (conditional on M_h) and u_h as the product of 2 binomial distributions (Darroch 1961; Seber 1982: 438). However, it follows from equation (1) that an approximately unbiased estimate of U_h is

$$\hat{U}_h = \hat{N}_h - M_h = \frac{u_h(M_h + 1)}{m_h + 1}. \quad (6)$$

Because $E\left[(\hat{N}_h - N_h)^2\right] = E\left[(\hat{U}_h - U_h)^2\right]$, an approximately unbiased variance estimate of \hat{U}_h is

$$v(\hat{U}_h) = \frac{(M_h + 1)(u_h + m_h + 1)(M_h - m_h)u_h}{(m_h + 1)^2(m_h + 2)}. \quad (7)$$

Thus, for 1-site experiments the approximately unbiased stratum estimators are \hat{U}_h and $v(\hat{U}_h)$. The total smolt abundance estimate is

$$\hat{U} = \sum_{h=1}^L \hat{U}_h, \quad (8)$$

and the variance estimate is

$$v(\hat{U}) = \sum_{h=1}^L v(\hat{U}_h). \quad (9)$$

An approximate 95% confidence interval for U is

$$\hat{U} \pm 1.96\sqrt{v(\hat{U})}. \quad (10)$$

Test for Consistency

The null hypothesis (H_0) of constant capture probability over all sampling periods can be tested using a chi-square test of homogeneity. If the test is nonsignificant, the strata can be pooled (equivalent to accepting H_0), and Chapman's (1951) estimator can be applied to estimate N or U (equation 1 or 6, respectively, without the stratum index). We recommend a fairly liberal significance level for accepting H_0 (e.g., $\alpha = 0.15$). The method consists of a test of homogeneity for the rows of an L -by-2 contingency table of M-R counts (Table 1). This test is similar to that given by Chapman and Junge (1956) and Darroch (1961), as described by Seber (1982: 438), for 2-sample stratified designs.

Mark Survival

A simple experiment can be performed to estimate short-term mark survival of smolts for each stratum (S_h), which is the combined rate of mark retention and fish survival after marking. The technique was described by Thedinga et al. (1994) and consists of holding a sample of marked smolts (in a flow-through box or other suitable device) for at least 2 d and counting the remaining live fish with visible marks. The experiment is repeated for each stratum, and the estimated mark survival for stratum h is

$$\hat{S}_h = \frac{a_h}{A_h}, \quad (11)$$

where a_h is the remaining number of marked live fish, and A_h is the initial number of marked smolts held. The estimated number of marked smolts available for recapture in stratum h is therefore

$$\hat{M}_h = M_h \hat{S}_h, \quad (12)$$

Table 1. Contingency table of time strata by recapture fate of marked smolts used to test the null hypothesis of consistency in capture probability. The statistical analysis is a chi-square test of homogeneity for the rows of the table.

Recaptured (m_h)	Not Recaptured ($M_h - m_h$)	Total (M_h)
m_1	$M_1 - m_1$	M_1
m_2	$M_2 - m_2$	M_2
...
m_L	$M_L - m_L$	M_L
m	$M - m$	M

which replaces M_h in equation (1) (2-site method) or equation (6) (1-site method) to estimate the smolt population adjusted for mark survival. The bootstrap method, which is presented in the following section, can be used to estimate the variance and confidence intervals. Adjusting for mark survival will always lower the abundance estimate, a result of increasing the estimated trap efficiency. This method does not take into account other sources of mark mortality (e.g., mark-induced predation); thus, it should be viewed as a minimum adjustment.

Bootstrap Technique

In M-R experiments, nonparametric bootstrapping using sampling with replacement implies an underlying binomial or multinomial distribution (Buckland and Garthwaite 1991). Bootstrap methods that use sampling without replacement imply an underlying hypergeometric distribution (or multivariate extension). The latter approach follows Chapman's (1951) M-R model, where m_h is conditional on M_h and n_h and sampling is without replacement. Therefore, we develop a parametric bootstrap method for the simple stratified M-R design based on the hypergeometric distribution. We also present a bootstrap method for mark survival-adjusted estimates.

In stratified sampling, the strata are treated independently (Sitter 1992) so that bootstrap estimates for N_h or U_h can be obtained, as well as for N or U . The random variable of interest is the number of recaptures, m_h . The total number of captures, n_h , is treated as fixed, and the number of releases, M_h , (thus M) is treated as fixed, unless the abundance estimate is adjusted for mark survival. The method works for 1- and 2-site experiments because $\hat{N}_h = \hat{U}_h + M_h$. Let $j = 1$,

2, ..., B indicate the bootstrap sample; usually $B = 1,000$.
Let

$$\hat{e}_h = \frac{n_h}{\hat{N}_h} \quad (13)$$

and

$$\hat{p}_h = \frac{M_h}{\hat{N}_h}. \quad (14)$$

For mark survival-adjusted estimates, \hat{M}_h replaces M_h in equation (14). In general, the following steps are used to obtain stratified bootstrap estimates:

1. generate a bootstrap estimate of N_h or U_h for stratum 1;
2. repeat step 1 independently for each of the L strata;
3. compute the stratified bootstrap estimate of N or U ;
4. repeat steps 1–3 B times to get the bootstrap sample of \hat{N} or \hat{U} ;
5. compute precision estimates of the bootstrap sample (for N or U).

Stratum bootstrap samples may be obtained by skipping step 3 and performing steps 4–5 for N_h or U_h . The bootstrap variance is

$$v(\hat{\theta}) = \sum_{j=1}^B (\hat{\theta}_j - \bar{\theta})^2 / (B-1), \quad (15)$$

where θ is the parameter of interest (N , N_h , U , or U_h) and $\bar{\theta}$ is the sample mean of the bootstrap estimates $\hat{\theta}_j$. Robust confidence intervals are found using percentiles of the bootstrap distribution (Efron and Tibshirani 1993: 168–177). For example, the 2.5th and 97.5th percentiles would give a 95% confidence interval.

Parametric bootstrapping consists of generating B variates from a fitted probability distribution and estimating smolt abundance for each of the B iterations. This involves drawing m_{jh} from a probability distribution and calculating \hat{N}_{jh} or \hat{U}_{jh} using equation (1) or (6). This is repeated for each stratum to calculate \hat{N}_j or \hat{U}_j using equation (3) or (8). The whole process is repeated B times. Bootstrap variance and percentile confidence intervals are then used to quantify the precision of the abundance estimate.

The parametric method treats m_h as hypergeometrically distributed with parameters $\{N_h, M_h, n_h\}$; \hat{N}_h is substituted for N_h , and n_h is treated as fixed in the bootstrap sampling. Each iteration consists of drawing a random variate, m_{jh} , from the hypergeometric distribution and following the bootstrap sampling steps. This technique was mentioned in Buckland and Garthwaite (1991) for nonstratified sampling but apparently was not applied.

For mark survival-adjusted estimates, a_h is treated as binomially distributed with parameters $\{A_h, S_h\}$; \hat{S}_h is substituted for S_h and A_h is fixed. Each bootstrap iteration consists of first drawing a_{jh} from the binomial distribution. Equations (13) and (14) are used to calculate \hat{S}_{jh} and \hat{M}_{jh} . The random variate m_{jh} is then drawn from the hypergeometric distribution (as described above), replacing M_h with \hat{M}_h , and the remaining bootstrap sampling steps are followed.

Sample Size

An important objective in estimating smolt abundance is achieving specified levels of accuracy and precision. This involves sampling to obtain an estimate that is accurate within an absolute error or relative error, allowing for a small probability of exceeding the allowable error (Thompson 1992: 31). For example, the goal may be to estimate N or N_h within 10% (relative error) of the true value with 95% confidence ($\alpha = 0.05$). Determining a sample size to achieve a desired accuracy and precision involves some initial guesswork and assumptions. In particular, the technique we describe requires some knowledge about the smolt population, expected trap efficiency, and temporal characteristics of the run. The method also relies on the normal approximation.

For a precision level of $\alpha = 0.05$, the absolute error bound is defined as

$$d = 2\sqrt{v(\hat{\theta})}, \quad (16)$$

where θ is the parameter of interest. The relative error bound is defined as

$$r = \frac{d}{\theta} = \frac{2\sqrt{v(\hat{\theta})}}{\theta}. \quad (17)$$

Robson and Regier (1964) suggest using $\alpha = 0.05$ and $r = 25\%$ for management work and $r = 10\%$ for research purposes.

We first develop a method of estimating the number of smolts to mark in a single M–R experiment or stratum (M_h) for a selected relative error, r_h . The technique requires a guess of trap efficiency (e_h). In some instances the researcher may prefer to specify a relative error for each stratum, in which case this method is suitable for the simple stratified design. We also describe a method of estimating stratum sample sizes (M_h) to achieve a selected relative error, r_i , for estimating the total emigration (N or U). This technique requires a guess of the fraction of the total smolt emigration in each stratum. Define

$$e_h = \frac{n_h}{N_h} \text{ and} \quad (18)$$

$$E[m_h | M_h, n_h] = \frac{M_h n_h}{N_h} = M_h e_h. \quad (19)$$

Chapman (1951) has shown that, for the hypergeometric model,

$$V(\hat{N}_h) = V(\hat{U}_h) \approx N_h^2 \left[(M_h e_h)^{-1} + 2(M_h e_h)^{-2} + 6(M_h e_h)^{-3} \right]. \quad (20)$$

The relative error bound for estimating N_h is therefore

$$r_h = 2 \sqrt{N_h^2 \left[(M_h e_h)^{-1} + 2(M_h e_h)^{-2} + 6(M_h e_h)^{-3} \right]} / N_h. \quad (21)$$

Solutions for M_h are found by solving the equation

$$M_h = \frac{K}{e_h(100)}. \quad (22)$$

Table 2 provides values of K for relative errors ranging from 5 to 100% and $\alpha = 0.05$ and 0.10. The relationship between e_h and M_h for $\alpha = 0.05$ and relative errors of 5, 10, 25, and 50% are shown in Figure 1. Because the relationship is a negative power function, relatively small increases in trap efficiency will often effect a substantial reduction in the number of smolts required for marking. Assuming that $U_h/N_h \approx 1$, equation (22) also applies to estimating U_h (1-site experiments), although the sample sizes will be biased slightly low.

Table 2. Method for estimating the sample size of fish to mark in mark–recapture experiments based on the equation $M = K[e(100)]^{-1}$, where M is the number of fish to mark and e is the trap efficiency. Solutions for K are given for 2 probability levels (α) of exceeding the desired relative error.

Relative Error (%)	Numerator (K)	
	$\alpha = 0.05$	$\alpha = 0.10$
5	160,250	108,914
6	111,385	75,757
7	81,896	55,734
8	62,744	42,722
9	49,604	33,790
10	40,201	27,396
11	33,274	22,692
12	27,998	19,106
13	23,886	16,310
14	20,620	14,087
15	17,982	12,291
16	15,835	10,834
17	14,052	9,623
18	12,556	8,605
19	11,287	7,742
20	10,203	7,003
21	9,277	6,374
22	8,472	5,828
23	7,769	5,349
24	7,150	4,928
25	6,603	4,556
30	4,648	3,227
35	3,470	2,426
40	2,705	1,906
45	2,181	1,550
50	1,807	1,295
60	1,318	961.7
70	1,023	759.1
80	830.0	625.9
90	696.7	532.8
100	600.0	464.7

The relationship between r_h and the total relative error is

$$r_t = \sqrt{\sum_{h=1}^L r_h^2 (\phi_h^2)}, \quad (23)$$

where ϕ_h is the fraction of the smolt run passing the sampling site in stratum h . If r_h is constant among strata, that is, $r_1 = r_2 = \dots = r_L$, then

$$r_t = r_h \sqrt{\sum_{h=1}^L \phi_h^2}, \quad (24)$$

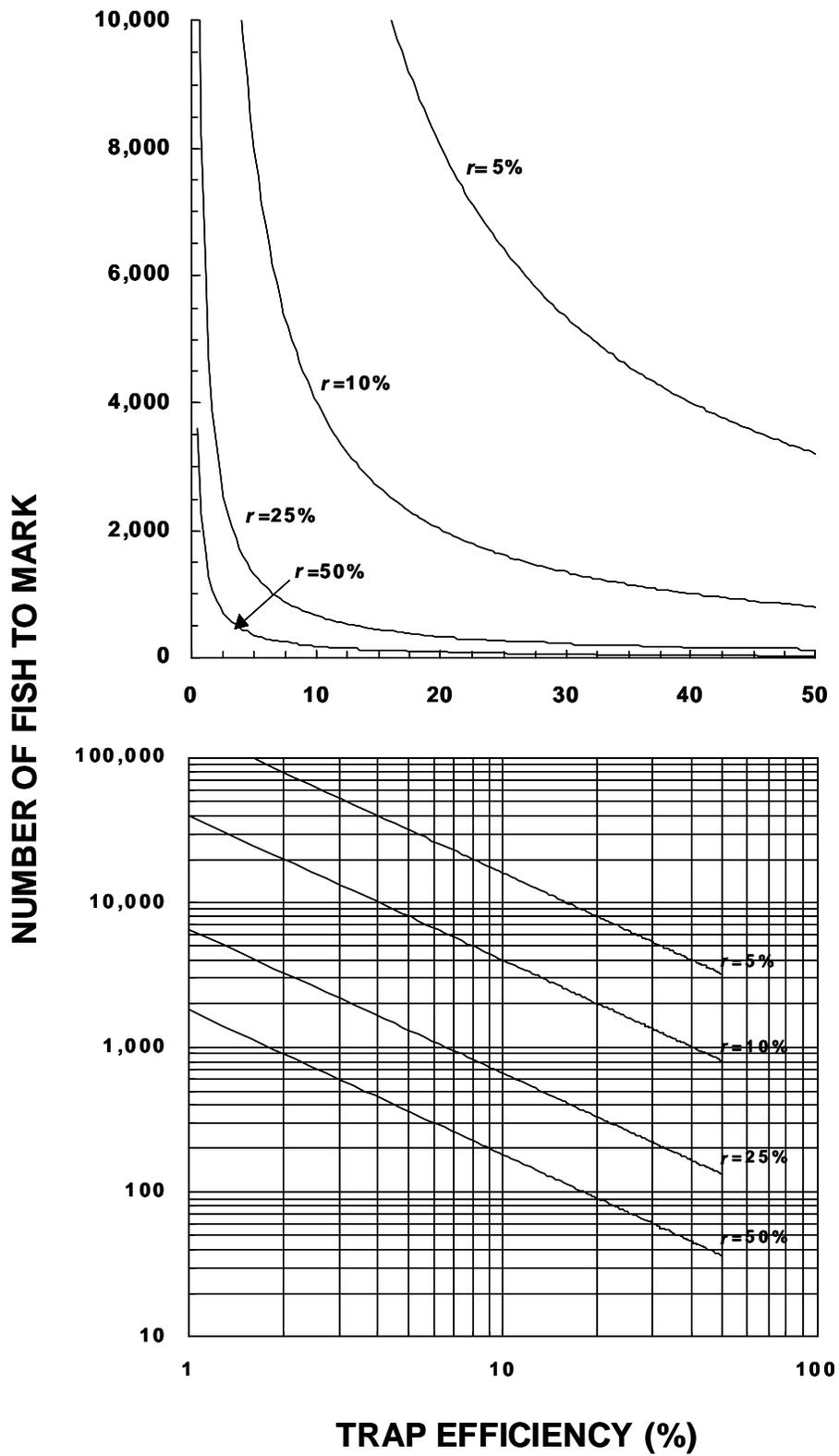


Figure 1. Relationship between trap efficiency (capture probability) and the number of salmon smolts to mark for relative errors (r) of 5, 10, 25, and 50%. The probability of exceeding r (α) is 5%. The lower plot shows the relationships on logarithmic scaled axes.

and thus,

$$r_h = \frac{r_t}{\sqrt{\sum_{h=1}^L \phi_h^2}} \quad (25)$$

After the ϕ_h are determined, r_h can be solved for the desired r_t , and Table 2 and equation (22) used to find the stratum sample sizes, M_h . Note that M_h will increase as a larger fraction of the run occurs in fewer strata. Chapman (1951) recommends designing M–R experiments so that at least 10 marked smolt are recovered (m_h) to ensure negligible bias in \hat{N}_h ; this minimum should always be adopted.

EXAMPLE

Methods

We conducted a study in 1997 to estimate the number of sockeye salmon smolts migrating from Akalura Lake, Kodiak Island, Alaska. One capture site was used in the experiment; thus, the appropriate estimators are for parameters U_h and U . Details of the methods are given in Coggins (1998). The release site was approximately 1 km upstream from the trap site and located in relatively low-velocity flow ($<0.5 \text{ m}\cdot\text{s}^{-1}$). A single Canadian fan trap (Ginetz 1977) covering about 30% of the stream width was located in Akalura Creek approximately 5.6 km below the lake outlet. A perforated-plate lead was attached from the trap to the nearest stream bank to prevent smolt passage on the shore side of the trap (Figure 2). The smolts were marked with Bismark Brown Y dye and held for 1 h in a perforated live box. Before release, dead smolts and those behaving abnormally were removed and censored from the data. Five trap efficiency trials were conducted approximately weekly and sampling took place from April 30 to June 12. Trap efficiency trials were paired with capture periods 1–2 d before the release and 1–3 d after the last observed recapture of marked smolts. An experiment to estimate short-term mark survival was conducted for each period and included in the analysis. The mark-survival test consisted of holding 100 marked fish for 5 d in a perforated live box and enumerating mark mortalities.

We estimated stratum sample sizes for marking to achieve a total relative error, r_t , of 15% ($\alpha = 0.05$). Based on M–R studies conducted on Akalura Creek from 1991 to 1996 (e.g., Swanton 1996), the expected trap efficiency, e_h , was 10%. We anticipated that the fraction of the run in each stratum would be 20, 35, 25,

15 and 5%. Therefore, the expected stratum relative error, r_h , was 30% (equation 25), and the estimated number of fish to mark per stratum, M_h , was 465 (equation 22 and Table 2), which we rounded up to 500. Actual sample sizes, which were usually slightly higher than 500, are given in Table 3.

In addition to the M–R trap, a second trap was installed and fitted with leads spanning the remaining portion of the stream (Figure 2). Collectively, the 2 traps captured 100% of the smolt emigration; this total configuration is hereafter referred to as the weir. The total number of smolts passing the weir were recorded daily. The weir census provided an opportunity to evaluate potential bias in the M–R estimate.

Results

Table 3 summarizes the Akalura Lake smolt data showing the stratification scheme. The average number of days required to recapture marked smolts was 1.7 d (all strata combined). A chi-square test of temporal consistency indicated the strata could not be pooled ($P < 0.001$; Table 4). A comparison of smolt abundance estimates and weir counts are given in Table 5. The total trap estimate, either unadjusted or adjusted for mark survival, did not differ significantly from the total weir count of 193,064 ($P = 0.16$ for the unadjusted estimate and $P = 0.40$ for the adjusted estimate). However, abundance estimates in periods 3 and 5 were significantly higher than the corresponding weir counts: for period 3, $P = 0.0021$ (unadjusted) and $P = 0.012$ (adjusted) and for period 5, $P = 0.0003$ (unadjusted) and $P = 0.0006$ (adjusted). On the other hand, a comparison of cumulative abundance curves indicated a fairly high degree of agreement between the weir counts and M–R estimates (Figure 3). Although the true trap efficiencies (trap catch divided by the weir count) exhibited considerable variation within periods, efficiency estimates reflected reasonably well the average true efficiencies within each stratum (Table 6). As expected, the true trap efficiency was underestimated in period 5, and to a lesser extent in period 3. However, only a small portion of the run (1.5%) occurred during period 5; thus the effect on the total abundance estimate was negligible. The normal approximation and the parametric bootstrap method (with and without the adjustment for mark survival) gave very similar precision estimates (Table 7).

DISCUSSION

The M–R model we described, stratified so that trap efficiency trials and capture periods are discretely

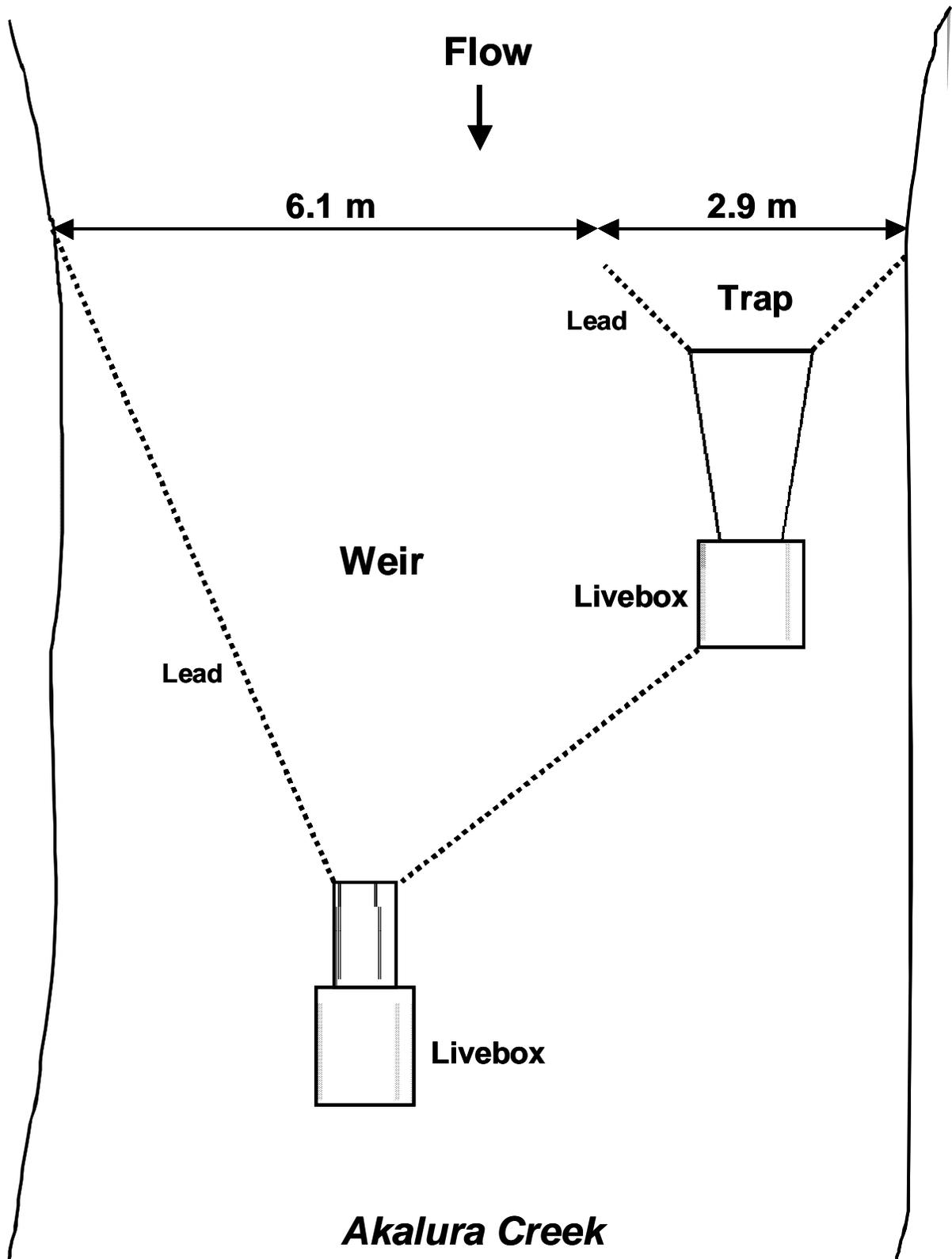


Figure 2. Schematic of the trap and weir configuration for capturing sockeye salmon smolts in Akalura Creek in 1997. The site was located about 5.6 km below the outlet of Akalura Lake.

Table 3. Summary of mark–recapture data of sockeye salmon smolts migrating from Akalura Lake in 1997. Trap efficiency trials were conducted in approximately weekly intervals. One trap site was used, and marked smolts were released 1 km upstream. Adjusted releases are based on a short-term mark–survival experiment.

Release Date	Period	(h) Stratum	Releases (M_h)		(m_h) Marked	Captures (u_h)		(n_h) Total
			Unadjusted	Adjusted		Unmarked		
7 May	4/30–5/11 ^a	1	487	468	100	13,244	13,344	
14 May	5/12–5/19	2	550	550	112	15,209	15,321	
22 May	5/20–5/26	3	519	488	86	6,811	6,897	
28 May	5/27–6/2	4	530	530	97	3,599	3,696	
5 Jun	6/3–6/12	5	520	489	31	497	528	

^a Only 30 smolts were captured from April 30 to May 4.

paired, has some practical and monetary advantages over the 2-sample stratified design (e.g., Darroch 1961) for estimating salmon smolt runs. (1) Because marking occurs at discrete intervals and continuous sampling occurs at only one site, personnel costs are substantially reduced — possibly by as much as half.

(2) Because each release of marked smolts corresponds to one capture period, only one type of mark is needed, which simplifies marking procedures, fish handling, and recapture tallying. (3) When only one capture site is used (1-site experiment), material costs (primarily for traps) are reduced by about half.

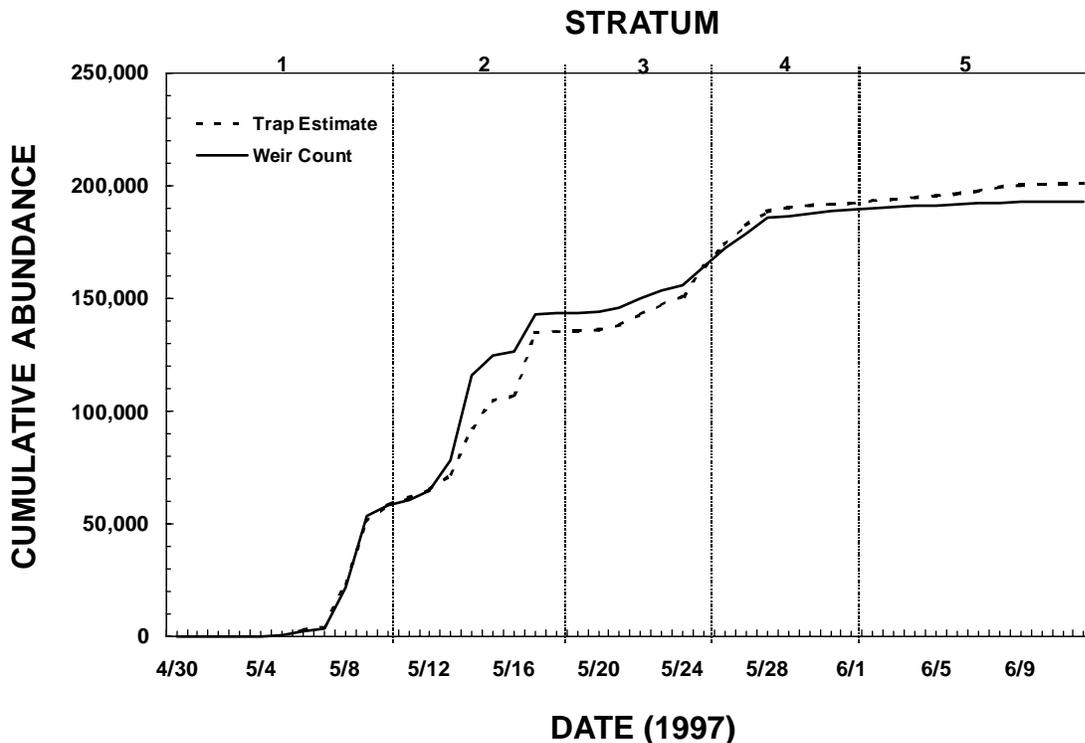


Figure 3. Comparison of cumulative frequency distribution curves of weir counts and mark–recapture abundance estimates of a sockeye salmon smolts migrating from Akalura Lake in 1997. Vertical lines indicate the temporal strata.

Table 4. Contingency table used to test for consistency in the capture probability of sockeye salmon smolts migrating from Akalura Lake in 1997; m_h is the number of marked smolts recaptured and $M_h - m_h$ is the number not recaptured. A chi-square test of homogeneity indicated significant temporal change in capture probability ($P < 0.001$); therefore, the strata were not pooled. Trap efficiency estimates are also given.

Period	Recapture m_h	Status $M_h - m_h$	Total (M_h)	Estimated Efficiency (%)
(1) 4/30–5/11	100	387	487	20.5
(2) 5/12–5/19	112	438	550	20.4
(3) 5/20–5/26	86	433	519	16.6
(4) 5/27–6/2	97	433	530	18.3
(5) 6/3–6/12	31	489	520	6.0
Total	426	2,180	2,606	

Violating the assumption of within-stratum consistency in capture probability is a potential drawback of the simple stratified design. However, the technique accounts for major changes and trends in capture probability, which in most cases will substantially reduce bias, compared with a nonstratified Peterson estimate. Minimizing stratum length while maintaining discrete periods will help alleviate this potential problem. The 1997 Akalura Lake M–R study indicated that the technique can be applied successfully using periods of approximately 1 week. A similar study conducted at Red Lake, Kodiak Island (Barrett et al. 1993), also produced no detectable bias in the estimate of the 1992 sockeye smolt emigration. In that study, the weir count was 1.314 million smolts and the M–R estimate was 1.210 million smolts (95% confidence interval: 0.908–1.513 million).

Table 5. Mark–recapture abundance estimates of sockeye salmon smolts migrating from Akalura Lake in 1997. A 1-site experiment was used, and estimates unadjusted and adjusted for mark survival are given. Standard errors (in parentheses) were estimated using a parametric bootstrap method based on the hypergeometric distribution.

Period	Smolt Abundance Estimates		Weir Count ¹
	Unadjusted	Adjusted	
(1) 4/30–5/11	63,991 (5560)	61,499 (5585)	60,511
(2) 5/12–5/19	74,161 (6227)	74,161 (6227)	83,265
(3) 5/20–5/26	40,709 (3973)	38,283 (3894)	28,504*
(4) 5/27–6/2	19,501 (1817)	19,501 (1817)	17,978
(5) 6/3–6/12	8,092 (1470)	7,610 (1398)	2,806*
Total	206,454 (9441)	201,054 (9518)	193,064

¹Weir counts indicated with an asterisk differ significantly ($P < 0.05$) from the mark–recapture estimate.

Most smolt M–R projects will require at least one season of development before final implementation. First, the type of trap(s) must be selected. Inclined-plane traps, for example, have been used on a number of smolt projects in Alaska and appear to work well under various conditions, including high-velocity flows and streams with heavy debris loads (e.g., Todd 1994). Thedinga et al. (1994) used rotary-screw traps to catch 4 salmonid species in Situk River in southeastern Alaska. Swanton et al. (1996) used Canadian fan traps to capture sockeye smolts in 2 Kodiak Island streams, although this type of trap may be susceptible to debris plugging (Todd 1994). Other capture techniques that should be considered include partial weirs (e.g., Dempson and Stansbury 1991) and various types of nets. If a 2-site experiment is planned, we recommend using different upstream and downstream trapping

Table 6. Comparison of trap efficiency estimates based on mark–recapture trials and “true” trap efficiencies calculated as the trap catch divided by the weir count. Relative smolt abundance, based on weir counts, is shown to indicate the importance of each sampling period. Estimated efficiency is given for unadjusted and mark-survival-adjusted releases. The range in true efficiency is given to indicate variation within each stratum.

Period	Relative Abundance (%)	Estimated Efficiency (%)		True Efficiency (%)	
		Unadjusted	Adjusted	Mean	Range
(1) 4/30–5/11	31.3	20.5	21.4	21.9	10.3–35.4
(2) 5/12–5/19	43.1	20.4	20.4	18.3	10.3–34.8
(3) 5/20–5/26	14.8	16.6	17.6	23.9	19.0–29.5
(4) 5/27–6/2	9.3	18.3	18.3	20.0	10.2–25.4
(5) 6/3–6/12	1.5	6.0	6.3	17.7	9.3–34.5

Table 7. Comparison of precision estimates using the normal approximation and parametric bootstrapping for mark–recapture estimates of the sockeye salmon smolt emigration in Akalura Creek in 1997. Unadjusted and mark–survival–adjusted estimates are given for the bootstrap technique. Bootstrapping was based on the hypergeometric distribution and consisted of 1,000 iterations; the percentile method was used to obtain confidence intervals.

Technique	Variance	Standard Error	95% Confidence Limits	
			Lower	Upper
Normal Approximation	9.180E+07	9,581	187,675	225,233
Bootstrap (unadjusted)	8.914E+07	9,441	188,865	225,030
Bootstrap (adjusted)	9.058E+07	9,518	183,361	220,318

methods to help reduce possible within-stratum bias (Seber 1982: 85–87).

The type of mark or tag will also need to be selected. Bismark Brown Y dye, for example, has been used in numerous sockeye smolt M–R studies (Todd 1994). Thedinga et al. (1994) used a Panjet instrument and 3 different inks to tattoo-mark 4 species of salmonid smolts. Dempson and Stansbury (1991) used Floy streamer tags to mark Atlantic salmon smolts. Macdonald and Smith (1980) used aluminum staple tags to mark sockeye smolts. After selecting a marking technique, marked fish should be observed for obvious behavioral effects. Short-term mark survival is simple to evaluate, can be easily incorporated into the analysis, and will provide valuable information about marking and handling effects.

It is often difficult to obtain adequate sample sizes for marking during the early and late segments of the migration. A simple solution to this problem is to release smolts over a few successive days and pool the release data for that period. However, it is most important to obtain adequate sample sizes for marking around the peak(s) of the emigration — those periods that account for the majority of the total smolt abundance. Achieving a trap efficiency of 10–20% should be adequate for most studies (Figure 1) and will require experimenting with the placement, number, and configuration of the trap(s).

For most types of gear, catchability varies with the size of the fish (Ricker 1975). For a migrating smolt population, traps that exhibit size-selectivity have unequal capture probability, which will bias the smolt estimate. The temporal stratification described in this paper attempts to account for variation in capture probability related to migration timing. For species with multiple age classes, such as sockeye salmon, run timing is generally related to age, and thus the major smolt size groups (e.g., Swanton et al. 1996). However, size-selectivity can still occur within an age class and run timing may overlap considerably. One potential rem-

edy for this problem is to stratify the estimate using size groups of the population that have constant catchability (Seber 1982: 81–82). Groups can be established based on known size differences between age classes or substocks of the emigrating smolt. The sampling method is an extension of the simple stratified design and only requires indexing size groups by using an additional subscript in the equations presented in this paper. To test the hypothesis of no size selectivity (H_0) in the recovery (downstream) sample, Seber (1982: 82) recommends a technique described by Robson (1969). The method consists of measuring (e.g., fork length) and marking each fish individually. After recapture sampling, the size distributions of the recaptured and not-recaptured smolts are compared using the Kolmogorov–Smirnov statistic; failure to reject H_0 indicates no size selectivity. Unfortunately, this technique involves added handling and individually tagging each fish, which may be impractical and costly. One alternative is to measure all marked and recaptured fish and censor identical measurements observed in both samples from the marked sample. Another option is to compare size distributions of all marked smolts and the recaptured smolts, although this approach violates (albeit modestly) the assumption of independent samples. It should be noted that, because of large sample sizes, the power of these tests is often so great that inconsequential size-selectivity will result in a statistically significant test.

Age (scale samples) and size (weight and length) data are usually collected concurrently in smolt enumeration studies to develop brood tables and to estimate the average size of smolts in the predominant age classes. We recommend sampling in proportion to the number of captures so that the within-stratum assumption of random sampling is satisfied reasonably well. In practice, it may be necessary to base the daily sample size on the previous day's catch, unless a technique of sampling proportional to the catch, such as a systematic method, is devised.

REFERENCES

- Bailey, N. T. J. 1951. On estimating the size of mobile populations from capture–recapture data. *Biometrika* 38:293–306.
- Bailey, N. T. J. 1952. Improvements in the interpretation of recapture data. *Journal of Animal Ecology* 21:120–127.
- Barrett, B. M., P. A. Roche, and C. O. Swanton. 1993. Sockeye salmon *Oncorhynchus nerka* smolt investigations at Red, Akalura, and Upper Station Lakes conducted in response to the 1989 M/V Exxon Valdez oil spill, 1990–1992. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 4K93-1, Kodiak.
- Buckland, S. T., and P. H. Garthwaite. 1991. Quantifying precision of mark–recapture estimates using the bootstrap and related methods. *Biometrics* 47:255–268.
- Chapman, D. G. 1948. A mathematical study of confidence limits of salmon populations calculated from sample tag ratios. *International Pacific Salmon Fisheries Commission Bulletin* 2:69–85.
- Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. *University of California Publications in Statistics* 1:131–160.
- Chapman, D. G., and C. O. Junge. 1956. The estimation of the size of a stratified animal population. *Annals of Mathematics and Statistics* 27:375–389.
- Cochran, W. G. 1977. *Sampling techniques*, 3rd edition. John Wiley and Sons, New York.
- Cochran, W. G. 1978. Laplace's ratio estimator. Pages 3–10 in H. A. David, editor. *Contributions to survey sampling and applied statistics*. Academic Press, New York.
- Coggins, L. G., Jr., and N. Sagalkin. 1998. Akalura Lake sockeye salmon restoration. Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, Exxon Valdez Oil Spill Restoration Final Report 97251-CLO, Kodiak.
- Darroch, J. N. 1961. The two-sample capture–recapture census when tagging and sampling are stratified. *Biometrika* 48:241–260.
- Dempson, J. B., and D. E. Stansbury. 1991. Using partial counting fences and a two-sample stratified design for mark–recapture estimation of an Atlantic salmon smolt population. *North American Journal of Fisheries Management* 11:27–37.
- Edmundson, J. A., G. B. Kyle, S. R. Carlson, and P. A. Shields. 1997. Trophic-level responses to nutrient treatment of meromictic and glacially influenced Coghill Lake. *Alaska Fishery Research Bulletin* 4(2):136–153.
- Efron, B., and R. J. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman and Hall, New York.
- Foerster, R. E. 1954. On the relation of adult sockeye salmon (*Oncorhynchus nerka*) returns to known smolt seaward migration. *Journal of the Fisheries Research Board of Canada* 11:339–350.
- Fried, S. M., and H. J. Yuen. 1987. Forecasting sockeye salmon (*Oncorhynchus nerka*) returns to Bristol Bay, Alaska: a review and critique of methods. Pages 273–279 in H. D. Smith, L. Margolis, and C. C. Wood, editors. Sockeye salmon (*Oncorhynchus nerka*) population biology and future management. Canadian Special Publications of Fisheries and Aquatic Sciences 96.
- Ginetz, R. M. J. 1977. A review of the Babine Lake development project 1961–1976. Environment Canada, Fish and Marine Service Technical Report Pac-T-77-6.
- Hartman, W. L., W. R. Heard, and B. Drucker. 1967. Migratory behavior of sockeye salmon fry and smolts. *Journal of the Fisheries Research Board of Canada* 24:2069–2099.
- Hoar, W. S. 1951. The behavior of chum, pink, and coho salmon in relation to their seaward migration. *Journal of the Fisheries Research Board of Canada* 8:241–263.
- Hume, J. M. B., K. S. Shortreed, and K. F. Morton. 1996. Juvenile sockeye rearing capacity of three lakes in the Frazer River system. *Canadian Journal of Fisheries and Aquatic Sciences* 53:3–17.
- Hyatt, K. D., and J. G. Stockner. 1985. Responses of sockeye salmon (*Oncorhynchus nerka*) to fertilization of British Columbia coastal lakes. *Canadian Journal of Fisheries and Aquatic Sciences* 42:320–331.
- Koenings, J. P., and R. D. Burkett. 1987a. An aquatic Rubic's cube: restoration of the Karluk Lake sockeye salmon (*Oncorhynchus nerka*). Pages 419–434 in H. D. Smith, L. Margolis, and C. C. Wood, editors. Sockeye salmon (*Oncorhynchus nerka*) population biology and future management. Canadian Special Publications of Fisheries and Aquatic Sciences 96.
- Koenings, J. P., and R. D. Burkett. 1987b. Population characteristics of sockeye salmon (*Oncorhynchus nerka*) smolts relative to temperature regimes, euphotic volume, fry density, and forage base within Alaskan lakes. Pages 219–234 in H. D. Smith, L. Margolis, and C. C. Wood, editors. Sockeye salmon (*Oncorhynchus nerka*) population biology and future management. Canadian Special Publications of Fisheries and Aquatic Sciences 96.
- Kyle, G. B., J. P. Koenings, and B. M. Barrett. 1988. Density-dependent, trophic level responses to an introduced run of sockeye salmon (*Oncorhynchus nerka*) at Frazer Lake, Kodiak Island, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences* 45:856–867.
- Macdonald, P. D. M., and H. D. Smith. 1980. Mark–recapture estimation of salmon smolt runs. *Biometrics* 36:401–417.
- Macdonald, P. D. M., H. D. Smith, and L. Jantz. 1987. The utility of Babine smolt enumerations in management of Babine and other Skeena River sockeye salmon (*Oncorhynchus nerka*) stocks. Pages 280–295 in H. D. Smith, L. Margolis, and C. C. Wood, editors. Sockeye salmon (*Oncorhynchus nerka*) population biology and future management. Canadian Special Publications of Fisheries and Aquatic Sciences 96.
- Peterman, R. M. 1981. Form of random variation in salmon smolt-to-adult relations and its influence on production estimates. *Canadian Journal of Fisheries and Aquatic Sciences* 38:1113–1119.
- Rawson, K. 1984. An estimate of the size of a migrating population of juvenile salmon using an index of trap efficiency obtained by dye marking. Alaska Department of Fish

- and Game, Division of Fisheries Rehabilitation, Enhancement and Development Report 28, Juneau.
- Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada, Bulletin 191.
- Robson, D. S. 1969. Mark-recapture methods of population estimation. Pages 120–140 in N. L. Johnson and H. Smith Jr., editors. New developments in survey sampling. Wiley Interscience, John Wiley and Sons, New York.
- Robson, D. S., and H. A. Regier. 1964. Sample size in Peterson mark-recapture experiments. Transactions of the American Fisheries Society 93:215–226.
- Schwartz, C. J., and J. B. Dempson. 1994. Mark-recapture estimation of a salmon smolt population. Biometrics 50:98–108.
- Seber, G. A. F. 1970. The effects of tag response on tag-recapture estimates. Biometrika 26:12–22.
- Seber, G. A. F. 1982. The estimation of animal abundance and related parameters, 2nd edition. Charles Griffin and Company Ltd., London.
- Sitter, R. R. 1992. Comparing three bootstrap methods for survey data. The Canadian Journal of Statistics 20:135–154.
- Solomon, D. J. 1981. Smolt migration in Atlantic salmon (*Salmo salar* L.) and sea trout (*Salmo trutta* L.). Pages 196–203 in E. L. Brannon and E. O. Salo, editors. Proceedings of the salmon and trout migratory behavior symposium. School of Fisheries, University of Washington, Seattle.
- Swanton, C. O., P. A. Nelson, and L. G. Coggins. 1996. Sockeye smolt population estimates, outmigration timing, and size at age characteristics for Red, Akalura, and Frazer Lakes, 1995. Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, Regional Information Report 4K96-16, Kodiak.
- Thedinga, J. F., M. L. Murphy, S. W. Johnson, J. M. Lorenz, and K. V. Koski. 1994. Determination of salmonid smolt yield with rotary-screw traps in the Situk River, Alaska, to predict effects of glacial flooding. North American Journal of Fisheries Management 14:837–851.
- Thompson, S. K. 1992. Sampling. John Wiley and Sons, New York.
- Todd, G. L. 1994. A lightweight, inclined-plane trap for sampling salmon smolts in rivers. Alaska Fishery Research Bulletin 1(2):168–175.

The Alaska Department of Fish and Game administers all programs and activities free from discrimination on the bases of race, religion, color, national origin, age, sex, marital status, pregnancy, parenthood, or disability. For information on alternative formats for this and other department publications, please contact the department ADA Coordinator at (voice) 907-465-4120, (TDD) 1-800-478-3648, or FAX 907-465-6078. Any person who believes she/he has been discriminated against should write to: ADF&G, P.O. Box 25526, Juneau, AK 99802-5526 or O.E.O., U.S. Department of the Interior, Washington, DC 20240.