Manual of Fisheries Survey Methods II: with periodic updates

Chapter 7: Stream Fish Population Estimates by Mark-and-Recapture and Depletion Methods

Roger N. Lockwood and James C. Schneider

Chapter 7: Stream Fish Population Estimates by Mark-and-Recapture and Depletion Methods

Roger N. Lockwood and James C. Schneider

Estimates of the total number of fish in sections of streams can be made reliably and inexpensively by subsampling a portion of the population. Two basic methods are available, mark-and-recapture and depletion. Either method is appropriate for shallow streams which can be waded and thoroughly sampled with electrofishing gear. Mark-and-recapture methods can also be used in deeper streams (as by electrofishing from a boat) if it can be reasonably assumed all targeted fish are vulnerable because either (a) all parts of the stream can be sampled or (b) marked fish are randomly mixed with unmarked fish.

7.1 Mark-and-recapture estimates

The mark-and recapture method is generally preferred over the depletion method and has been shown to be unbiased when more than 50% of a population is marked (Jensen 1992). The mark-and-recapture method requires the following conditions:

1. Marked and unmarked fish have the same mortality rates;
2. Marked and unmarked fish are equally vulnerable to capture;
3. Marks are retained during the sampling period and all marks on recaptured fish are recognized;
4. Marked fish randomly mix with unmarked fish;
5. There is negligible emigration or immigration during the recapture period.

The general process for estimating a fish population using the mark-and-recapture method entails:

1. Collecting a sample of fish of the target species from a discrete section of stream during an initial "marking run";
2. Giving fish identifying marks, such as a tag or temporary fin clip;
3. Tabulating data by species and size;
4. Releasing fish in good condition back into the same area;
5. Allowing at least 1 day for marked fish to recover and become mixed in the population;
6. Collecting a random sample of fish during a subsequent "recapture run";
7. Noting the ratio of marked to unmarked fish by species and size (e.g., inch group);
8. Calculating for each combination of species and size group (to compensate for gear selectivity) an estimate of abundance by a Petersen equation;
9. Summing the size group estimates by species to obtain an estimate of the total population within the size range actually sampled.

7.1.1 Chapman - Petersen methods

Ricker (1975) discusses the calculation of population estimates in detail. He recommends a slight variation of the Chapman modification of the Petersen equation because it gives a statistically unbiased estimate for finite populations, such as we deal with in inland waters. The Chapman variation is very similar to the Bailey modification of the Petersen equation, which is also widely

used, and both produce estimates slightly less than the simple Petersen equation. Estimation of population N and variance of N, with the Chapman modification, follow as:

$$N = \frac{(M+1)(C+1)}{R+1} \, , \tag{1}$$

$$Variance \ of \ N = \frac{(M+1)^2(C+1)(C-R)}{(R+1)^2(R+2)} = \frac{N^2(C-R)}{(C+1)(R+2)} \, , \tag{2}$$

$$\text{Standard error} = \sqrt{\text{Variance of } N} \, ,$$

$$95\% \text{ confidence limits} = N \pm t(\text{Standard error}) \, ,$$

where,

C = total number of fish caught in second sample (including recaptures),

M = number of fish caught, marked and released in first sample,

N = population estimate,

R = number of recaptures in the second sample (fish marked and released in the first sample),

t = Student's t for C-1 degrees of freedom.

Variance Equation (2) should be used whenever variance estimates are to be combined, for example when summing estimates and variances for two to more size groups to obtain a total population estimate. However, it is not the best estimator of variance for single estimates (Ricker 1975). His recommendation is to use either binomial charts or a Poisson distribution (Table 7.1). These provide low and high ranges for R which are then substituted in Equation (1) to calculate the lower and upper confidence limits. These limits are typically asymmetrical and measure variability more accurately. While 95% confidence limits are often used for research, 68% limits (1 standard error) may be suitable for management purposes.

Ricker (1975) stated that the probability of a systematic statistical bias in the population estimate can be ignored if recaptures number 3-4 or more. Therefore, if necessary, pool data from adjacent size groups to obtain **at least** 3-4 recaptures per estimate.

Example 7.1–If 100 fish were marked and released from the first run, and the second run contained 80 fish of which 10 were recaptures:

$$N = \frac{(100+1)(80+1)}{10+1} = 744 \, \text{fish} \, ,$$

$$Variance \ of \ N = \frac{(100+1)^2(80+1)(80-10)}{(10+1)^2(10+2)} = 39,834 \, ,$$

$$\text{Standard error} = \sqrt{39834} = 200 = 68\% \text{ confidence limit} \, ,$$

and approximate 95% confidence limits $= N \pm 2(200) = N \pm 400$,

i.e., Lower limit (N_L) = 344 and Upper limit (N_U) = 1,144.

Continued on next page.

Example 7.1–Continued.

Better 95% confidence limits from the equation given in Table 7.1 are:

$$R + 1.92 \pm 1.960\sqrt{R + 1.0} = 11.92 \pm 6.501,$$

$$\text{or } R_L = 5.42 \quad \text{and} \quad R_U = 18.42.$$

When substituted for R in Equation (1):

$$N_L = \frac{(100 + 1)(80 + 1)}{18.42 + 1} = 421,$$

and

$$N_U = \frac{(100 + 1)(80 + 1)}{5.42 + 1} = 1,274.$$

We conclude from this example that the population contains about 744 fish, but the statistical error is relatively large and with 95% certainty the true number lies between 421 and 1,274. Note that this measurement of error is for random error only, and any systematic error (e.g., avoidance of recapture by marked fish) is unknown.

In many studies the investigator may desire to add several population estimates and have confidence limits for the total population ($\hat{N}$). For example, one might have separate population estimates for trout in inch groups 8, 9, and ≥10, which when added give the total number of catchable-size trout. The appropriate equation for computing a total variance for j inch groups is:

$$Var(\hat{N}) = \sum_{i=1}^{j} Var(N)_i \tag{3}$$

Example 7.2–Inch group 8 has an estimated population of 357 fish with a variance of 20,392; inch group 9 has an estimated population of 293 fish with a variance of 12,100; and inch group 10+ has an estimated population of 153 with a variance of 3,935:

$$\hat{N} = 357 + 293 + 153 = 803 \text{ fish},$$

$$\text{Variance of } \hat{N} = 20,392 + 12,100 + 3,935 = 36,427,$$

$$\text{Standard error of } \hat{N} = \sqrt{36,427} = 191,$$

and approximate 95% confidence limits $= \hat{N} \pm 2(191) = \hat{N} \pm 382$,

i.e., $N_L = 421$ and $N_U = 1,185$.

7.2 Depletion estimates

The depletion method (also known as the "Zippin" method, see Zippin 1958) is satisfactory if the stream is very small, it is expedient to collect all data within a short time period such as one day, and the population being estimated is relatively small (roughly less than 2,000 individuals). If fish are likely to migrate in or out of a study section soon (say in less than 1 week), the depletion method is superior to the mark-and-recapture method due to a shorter sampling time period. This method

requires that an adequate number of fish be removed on each sampling pass so that measurably fewer fish are available for capture and removal on a subsequent pass. Two types of depletion methods are used, two-pass and multiple-pass. Because of differences in gear selectivity, partitioning estimates by species and size groups is recommended. For both two-pass and multiple-pass methods, size group estimates and their variances are summed, as with mark-and-recapture methods, to provide total population estimates. The following conditions must be met for accurate depletion method estimates:

1. Emigration and immigration by fish during the sampling period must be negligible;

2. All fish within a specified sample group must be equally vulnerable to capture during a pass;

3. Vulnerability to capture of fish in a specified sample group must remain constant for each pass (e.g., fish do not become more wary of capture);

4. Collection effort and conditions which affect collection efficiency, such as water clarity, must remain constant.

Depletion estimates are made following the general process:

1. Remove (or mark to simulate removal) fish within a discrete section of stream;

2. Record number of fish removed (or marked) by species and size group;

3. Repeat steps 1 and 2;

4. If steps 1 and 2 were completed twice, calculate population estimates using two-pass equations;

5. If steps 1 and 2 were completed more than twice, calculate population estimates using multiple-pass equations.

7.2.1 Two-pass depletion methods

For two-pass depletion estimates, fish are captured and removed during two capture sessions. Equations provided here are described in greater detail in Seber and Le Cren (1967). Population estimate N and variance of N are calculated as:

$$p = \frac{C_1 - C_2}{C_1} ,$$ (4)

$$N = \frac{C_1^2}{(C_1 - C_2)} ,$$ (5)

$$\text{Variance of } N = \frac{C_1^2 C_2^2 (C_1 + C_2)}{(C_1 - C_2)^4} ,$$ (6)

$$\text{Standard error of } N = \sqrt{\text{Variance of } N} ,$$

where,

C_1 = number of fish removed in first sample,

C_2 = number of fish removed in second sample,

N = population estimate,

p = probability of capture,

Two-pass depletion estimates are unbiased when $p \geq 0.80$ and quite unreliable when $p \leq 0.20$ (i.e., when less than 20% of the population is caught per pass).

Example 7.3–On the first pass 200 fish were collected and on the second pass 95 fish were collected. Estimated population and confidence limits are calculated as:

$$p = \frac{200 - 95}{200} = 0.525 ,$$

$$N = \frac{200^2}{(200 - 95)} = 381 .$$

$$\text{Variance of } N = \frac{(200^2)(95^2)(200 + 95)}{(200 - 95)^4} = 876 ,$$

$$\text{Standard error of } N = \sqrt{876} = 30 ,$$

and approximate 95% confidence limits $= N \pm 2(30) = N \pm 60 $,

i.e., $N_L = 321$ and $N_U = 441$.

7.2.2 Multiple-pass depletion methods

This method requires three or more passes on the selected stream section and involves additional calculations to estimate the population. The multiple pass depletion method relies heavily upon consistent catchability ($p_1 = p_2 = p_3 = ... p_s = p$). Further description of these equations are found in Carle and Strub (1978). Estimation steps are as follows:

$$T = \sum_{i=1}^{k} C_i , \tag{7}$$

$$X = \sum_{i=1}^{k} (k - i) C_i , \tag{8}$$

where,

i = pass number,

k = number of removals (passes),

C_i = number of fish caught in i^{th} sample,

X = an intermediate statistic used below,

T = total number of fish caught in all passes.

The maximum likelihood estimate of N is determined by an iterative process by substituting values for n until:

$$\left[\frac{n+1}{n-T+1}\right] \prod_{i=1}^{k}\left[\frac{kn-X-T+1+(k-i)}{kn-X+2+(k-i)}\right]_i \leq 1.0, \tag{9}$$

where n is the smallest integer satisfying Equation (9). Note that results of Equation (9) are rounded to one decimal place. Probability of capture, $p,$ and variance of N are then estimated by:

$$p = \frac{T}{kN-X}, \tag{10}$$

$$\text{Variance of } N = \frac{N(N-T)T}{T^2 - N(N-T)\left[\frac{(kp)^2}{(1-p)}\right]}, \tag{11}$$

$$\text{Standard error of } N = \sqrt{\text{Variance of } N}.$$

Since estimating N is an iterative process, a suggested initial value for n is T, and subsequent selections for n should progressively increase from T. These Equations should be setup in a spreadsheet to facilitate selection of n, and estimates of N and variance of N.

Example 7.4–On the first pass 300 fish were removed, 130 on the second, and 69 on the third:

$$T = 300 + 130 + 69 = 499,$$

$$X = [(3\text{-}1) * 300] + [(3\text{-}2) * 130] + [(3\text{-}3) * 69] = 600 + 130 + 0 = 730.$$

We know there were at least 499 fish (T) in the population, so for our first n let's try 499 in Equation (9):

$$\left[\frac{499+1}{499-499+1}\right]\left[\frac{(3*499)-730-499+1+(3-1)}{(3*499)-730+2+(3-1)}\right]\left[\frac{(3*499)-730-499+1+(3-2)}{(3*499)-730+2+(3-2)}\right]\left[\frac{(3*499)-730-499+1+(3-3)}{(3*499)-730+2+(3-3)}\right]$$

$$= (500.0)(0.3515)(0.3506)(0.3498) = 21.5540 \text{ which is rounded to } 21.6.$$

Since 21.6>1.0 we must select another number for n greater than 499. For our second try let's use $n = 520$:

$$\left[\frac{520+1}{520-499+1}\right]\left[\frac{(3*520)-730-499+1+(3-1)}{(3*520)-730+2+(3-1)}\right]\left[\frac{(3*520)-730-499+1+(3-2)}{(3*520)-730+2+(3-2)}\right]\left[\frac{(3*520)-730-499+1+(3-3)}{(3*520)-730+2+(3-3)}\right]$$

$$= (23.6818)(0.4005)(0.3998)(0.3990) = 1.5.$$

Continued on next page.

Example 7.4–Continued.

Subsequent trials with n equal to 530, 540, 545 and 546 yield estimates for N of 1.2, 1.1, 1.1 and finally 1.0. Since our goal is to come close to 1.0 without going over, the maximum likelihood estimate of the population (N) is 546. We can now estimate probability of capture (p) and error statistics:

$$p = \frac{499}{(3*546)-730} = 0.5496 \,,$$

$$\text{Variance of } N = \frac{546*(546-499)*499}{499^2 - 546*(546-499)*\left[\dfrac{(3*0.5496)^2}{(1-0.5496)}\right]} = \frac{12,805,338}{94,109} = 136 \,,$$

$$\text{Standard error of } N = \sqrt{136} = 11.7 \,,$$

approximate 95% confidence limit $= N \pm 2(11.7) = N \pm 23$,

i.e., $N_L = 523$ and $N_U = 569$.

As previously mentioned, the depletion method assumes that all s capture probabilities for k passes are equivalent, that is $p_1 = p_2 = \ldots p_s = p$. When using the two-pass method and fish are removed from the section on the first pass, there is no way to verify this assumption. However, when fish are marked and released back into the section (to simulate removal) number of fish captured on each pass are expected to be similar. Consider the data presented in Example 7.1. On the first pass, 100 fish (C_1) were captured and marked. On the second pass, 80 fish were captured of which 70 (C_2) were unmarked. Treating these data as depletion data gives $C_1 = 100$, $C_2 = 70$ with $N = 333$ (using Equation 5). This result is substantially different from the mark-and-recapture estimate of 744 using Equation (1) and is caused by unequal catch probabilities. If capture probabilities had been equal, 100 fish would have been captured on the second pass and Equations (1) and (5) would have given similar results. This verification technique may only be used when 1 or more days elapse between passes.

Example 7.5–On the first pass 120 fish are marked and released back into the section. On the second pass 60 marked fish and 60 unmarked fish are captured:

Using mark-and-recapture Equation (1):

$$N = \frac{(120+1)(120+1)}{60+1} = 240 \,.$$

Using depletion Equation (5):

$$N = \frac{120^2}{120-60} = 240 \,.$$

When more than two passes are made equality of p's can be verified in numerous ways. Seber and Le Cren (1967) suggest simply plotting each catch against the sum of all previous catches:

$$\sum_{i=0}^{k-1} C_i \ .$$

(12)

Example 7.6–Using data from example 7.4 where C_1=300, C_2=130 and C_3=69, the following values would be plotted:

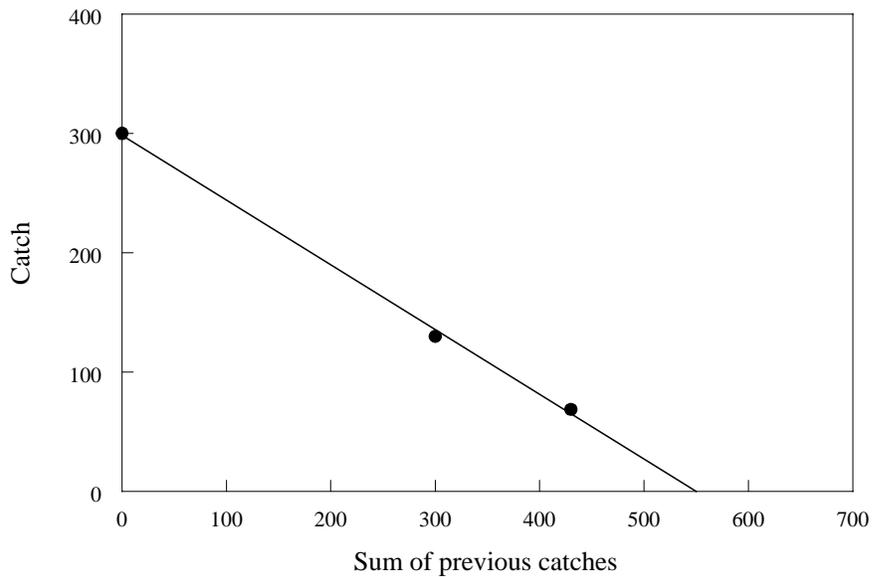| Pass	Sum of previous catches	Catch
1	0	300
2	300	130
3	430	69

Figure 7.1–Individual catches plotted against the sum of previous catches from example 7.4.

A goodness of fit test may also be used to evaluate equality of capture probability (White et al. 1982). This follows the χ^2 test form $(\text{Observed - Expected})^2 / \text{Expected}$. The first step is to calculate the expected number of fish collected on each pass using the estimated population $N,$ from Equation (9), and estimated probability of capture $p,$ from Equation (10):

$$E(C_1) = Np ,$$ (13)

and for $i > 1$.

$$E(C_i) = N(1-p)^{i-1} p$$ (14)

Calculated χ^2 then is:

$$\chi^2 = \frac{[C_1 - E(C_1)]^2}{E(C_1)} + \frac{[C_2 - E(C_2)]^2}{E(C_2)} + \cdots + \frac{[C_i - E(C_k)]^2}{E(C_k)} .$$ (15)

The test statistic χ^2 from Equation (15) is compared with $\chi^2_{0.95}$, Table 2, with k-2 degrees of freedom (df). Note that two degrees of freedom are lost because N is estimated (Snedecor and Cochran 1991:77). If $\chi^2 < \chi^2_{0.95}$, probability of capture did not differ significantly (at the 95% level of certainty) between passes; if $\chi^2 \geq \chi^2_{0.95}$, then probability of capture was significantly different with 95% certainty.

Example 7.7–Using the data from example 7.4: C_1=300, C_2=130, C_3=69, N=546 and p=0.5496:

$$E(C_1) = 546(0.5496) = 300 ,$$

$$E(C_2) = 546(1-0.5496)^1(0.5496) = 135 ,$$

$$E(C_3) = 546(1-0.5496)^2(0.5496) = 61 ,$$

$$\chi^2 = \frac{(300-300)^2}{300} + \frac{(135-130)^2}{135} + \frac{(61-69)^2}{61} = 0.000 + 0.185 + 1.049 = 1.234 ,$$

where df = 3-2=1, and $\chi^2_{0.95} = 3.841$.

Since $\chi^2 < \chi^2_{0.95}$ we show no significant difference (95% certainty) between capture probabilities and Equations (7-11) are appropriate.

Variation in capture probability hinges on numerous factors. Of particular concern is increased wariness when fish are exposed to electrofishing. Heggberget and Hesthagen (1979) used the 2-pass depletion method to estimate Atlantic salmon *Salmo salar* and brown trout S*almo trutta* in two small Norway streams and suggested that populations were underestimated by as much as 50% due to electrical current avoidance on the second pass. On the other hand, Peterson and Cederholm (1984) found that probability of capture for shocked juvenile coho salmon *Oncorhynchus kisutch* was similar to previous capture rates after a minimum of 1 hour recovery time. To minimize error, the amount of effort used on each pass should be as constant as possible and estimates should be stratified by species and size group to avoid gear selectivity. Appropriate steps should be taken to minimize immigration and emigration of fish, as by using blocking nets on small streams to greatly reduce fish movement. When sampling streams where blocking nets are not practical, effect of fish movement on population estimates can be reduced by sampling longer sections.

When variation in capture probability is severe, and estimates of N and variance of N using Equations (7-11) are invalid, more computationally intense methods such as those given by Schnute (1983) or White et al (1982) are necessary. A copy of CAPTURE (White et al. 1982) is

available at the Institute for Fisheries Research. This interactive software calculates an appropriate estimate of *N* and variance of *N* when capture probabilities are different.

The assumptions of the depletion method are rigorous regarding constant fish catchability for each sample and that more than 20% (better, >30%) of the population be captured in each sample. The assumptions of the depletion method are most suspect in large streams (more likely to have refuges) and for large fish (more likely to be agile or wary). Therefore, it is wise to design sampling procedures to retain the option of computing mark-and-recapture estimates while conducting depletion sampling. The option is left open by marking and releasing fish after the first pass, noting their recapture in subsequent passes, and ignoring marked fish for depletion method analysis or counting them as "recaptures" for mark-and-recapture method analysis. Note however, if more than one pass is made per day, the marked fish may not have recovered and become mixed, thereby violating the basic assumption of the mark-and-recapture method.

Mark-and-recapture methods also have essential constraints. If marked fish are more easily captured than unmarked fish on the second pass, the population will be underestimated. If marked fish are more difficult to capture than unmarked fish on the second pass, the population will be overestimated. Smallmouth bass *Micropterous dolomieu* in particular are difficult to recapture and mark-and-recapture methods are not recommended (Lyons and Kanehl 1993).

Table 7.1–Poisson distribution of lower and upper 95% confidence limit coefficients[a] for number of recaptures (R), and Student's 95% confidence t values for number of degrees of freedom (df).

Poisson distribution							Student's t value	
R	Lower	Upper	R	Lower	Upper		df	t_{95}
0	0.0	3.7	26	17.0	38.0		1	12.706
1	0.1	5.6	27	17.8	39.2		2	4.303
2	0.2	7.2	28	18.6	40.4		3	3.182
3	0.6	8.8	29	19.4	41.6		4	2.776
4	1.0	10.2	30	20.2	42.8		5	2.571
5	1.6	11.7	31	21.0	44.0		6	2.447
6	2.2	13.1	32	21.8	45.1		7	2.365
7	2.8	14.4	33	22.7	46.3		8	2.306
8	3.4	15.8	34	23.5	47.5		9	2.262
9	4.0	17.1	35	24.3	48.7		10	2.228
10	4.7	18.4	36	25.1	49.8		11	2.201
11	5.4	19.7	37	26.0	51.0		12	2.179
12	6.2	21.0	38	26.8	52.2		13	2.160
13	6.9	22.3	39	27.7	53.3		14	2.145
14	7.7	23.5	40	28.6	54.5		15	2.131
15	8.4	24.8	41	29.4	55.6		16	2.120
16	9.2	26.0	42	30.3	56.8		17	2.110
17	9.9	27.2	43	31.1	57.9		18	2.101
18	10.7	28.4	44	32.0	59.0		19	2.093
19	11.5	29.6	45	32.8	60.2		20	2.086
20	12.2	30.8	46	33.6	61.3		21	2.080
21	13.0	32.0	47	34.5	62.5		22	2.074
22	13.8	33.2	48	35.3	63.6		23	2.069
23	14.6	34.4	49	36.1	64.8		24	2.064
24	15.4	35.6	50	37.0	65.9		60	2.000
25	16.2	36.8					∞	1.960

[a] Substitute the coefficients for R in Formula (1). For larger values of R, use the following equation (Ricker 1975) for 95% limit coefficients: $R + 1.92 \pm 1.96\sqrt{R + 1.0}$.

Table 7.2–Percentiles of the χ^2 distribution for 70% and 95% certainty[a].

Degrees of freedom	$\chi^2_{0.70}$	$\chi^2_{0.95}$
1	1.074	3.841
2	2.408	5.991
3	3.665	7.815
4	4.878	9.488
5	6.064	11.070
6	7.231	12.592
7	8.383	14.067
8	9.524	15.507
9	10.656	16.919
10	11.781	18.307

[a] For additional degrees of freedom or alternate levels of certainty, see χ^2 tables in texts such as Snedecor and Cochran (1991).

7.3 References

Carle, F. L., and M. R. Strub. 1978. A new method for estimating population size from removal data. Biometrics 34: 621-630.

Heggberget, T. G., and Hesthagen, T. 1979. Population estimates of young Atlantic Salmon, *Salmo salar* L., and brown trout, *Salmo trutta* L., by electrofishing in two small streams in north Norway. Rep. Inst. Freshwater Res., Drottningholm, 58: 27-33.

Jensen, A. L. 1992. Integrated area sampling and mark-recapture experiments for sampling fish populations. Biometrics 48: 1201-1205.

Lyons, J., and P. Kanehl. 1993. A comparison of four electroshocking procedures for assessing the abundance of smallmouth bass in Wisconsin streams. U. S. Department of Agriculture, Forest Service, General Technical Report NC-159, St. Paul, Minnesota.

Peterson, N. P., and Cederholm, C. J. 1984. A comparison of the removal and mark-recapture methods of population estimation for juvenile coho salmon in small stream. North American Journal of Fisheries Management 4:99-102.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada, Bulletin 191.

Seber, G. A. F., and E. D. Le Cren. 1967. Estimating population parameters from catches large relative to the population. Journal of Animal Ecology 36: 631-643.

Schnute, J. 1983. A new approach to estimating populations by the removal method. Canadian Journal of Fisheries Aquatic Sciences 40: 2153-2169.

Snedecor, G. W., and W. G. Cochran. 1991. Statistical methods, eighth edition. Iowa State University Press, Ames, Iowa.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos: Los Alamos National Laboratory.

Zippin, C. 1958. The removal method of population estimation. Journal of Wildlife Management 22: 82-90.