# FRED Reports

AN ESTIMATE OF THE SIZE OF A MIGRATING POPULATION OF JUVENILE SALMON USING AN INDEX OF TRAP EFFICIENCY OBTAINED BY DYE MARKING

by

Kit Rawson

Number 28

**Alaska Department of Fish & Game**
Division of Fisheries Rehabilitation,
Enhancement and Development

# AN ESTIMATE OF THE SIZE OF A MIGRATING POPULATION OF JUVENILE SALMON USING AN INDEX OF TRAP EFFICIENCY OBTAINED BY DYE MARKING

by
Kit Rawson
Number 28

Alaska Department of Fish and Game
Division of Fisheries Rehabilitation,
Enhancement and Development

Don W. Collinsworth
Commissioner

Stanley A. Moberly
Director

P.O. Box 3-2000
Juneau, Alaska 99802

April 1984

TABLE OF CONTENTS

## LIST OF TABLES



## LIST OF FIGURES

## ABSTRACT

When juvenile salmon are enumerated during their downstream migration, it is often possible to capture and count only a portion of the run. By means of dye marking or other techniques a sample of fish can be marked and released upstream, and the number of marked fish recovered in the traps can be used as an index of trap efficiency. The estimated trap efficiency can then be used to expand the counts of fish caught in the traps to obtain an estimate of the total number of juvenile salmon in the migration.

The intuitive estimate of the run size based upon trap efficiency is biased, and its variance is not simple to compute. Based upon work by LaPlace, as reviewed by Cochran, the bias and variance of this estimate can be computed. Applying LaPlace's computations to the salmon enumeration problem, I present an estimate for the number of fish in the migration, which corrects for the bias in the intuitive estimate, as well as a formula for computing a confidence interval for this estimate. The practical implementation of this technique is illustrated through three examples from recent FRED Division projects.



KEY WORDS: dye, marking, smolt, fry, enumeration, mark recapture, variance, bias, confidence interval, trap efficiency.


## INTRODUCTION

In several areas of the state the Division of Fisheries Rehabilitation, Enhancement, and Development (FRED) is conducting projects to enumerate juvenile salmon swimming downstream. In general it is impossible to count all of the run; so, traps placed in the river are used to obtain a sample. In order to expand the trap counts to estimate the total number of fish in the migration it is necessary to estimate the proportion of the run that is sampled by the traps, i.e., the trap efficiency. For this purpose, FRED Division personnel have developed a technique based upon the marking, release, and subsequent recapture of some of these migrants. In this method some of the fish caught in the traps are placed in a solution of dye, which colors them and makes them easily distinguishable from undyed fish. The dyed fish are then transported upstream where they are released, and the number of dyed fish subsequently recaptured in the traps is used to estimate the trap efficiency. The purpose of this report is to present a statistically valid method for obtaining an estimate of run size with a confidence interval based on the dye mark and recapture technique.

## THE STATISTICAL MODEL

### Single Release of Dyed Fish

In developing the statistical model for the analysis of data obtained by the dye marking technique, I will first consider the case of a single release of dyed fish. This applies when only one batch of fish is dyed and released during the migration season. It can also apply in cases where several dye tests are performed, and the results are later lumped and considered as coming from a single release. This lumping would be valid only if the estimated trap efficiencies recovered in the several releases were deemed not to be significantly different on the basis of chi-square tests (see Example 1 below). The notation used in this report for the case of a single release of dyed fish is summarized in Table 1.

To begin with, I assume that each fish, dyed or undyed, is caught in the traps independently of the fate of other fish. I also assume that each fish is caught in the traps with the same probability, $\varepsilon$ (standing for trap efficiency). From these assumptions it follows that (see Table 1 for definitions of notation):

1) $d$ is a random variable following a binomial probability distribution with parameters $\varepsilon$ and $D$, and

2) each $n_i$ is a random variable following a binomial probability distribution with parameters $\varepsilon$ and $N_i$.

If we assume further that trap efficiency, $\varepsilon$, remains constant for all $k$ days of the study, then

3) $n$ is a binomial random variable with parameters $\varepsilon$ and $N$.

The more detailed assumptions necessary for 1), 2), and 3) are listed in Table 2.

Given this statistical model, it should be possible to derive formulas to estimate parameters of interest with confidence intervals. For the trap efficiency the formula is simply:

$$\hat{\varepsilon} = d/D. \qquad [1]$$

A confidence interval for $\hat{\varepsilon}$ can be obtained using well-known results from the binomial distribution. When $\hat{\varepsilon}$ is small (less than 0.1), Fleiss (1981, p. 14) suggests using a procedure that yields an asymmetric confidence interval, which is more accurate than the symmetric one that results from applying the usual approximation to a normal distribution. In any case, it is not difficult to estimate $\varepsilon$ with a confidence interval.

Obtaining an estimate of $N_i$ with a confidence interval is not so easy. An intuitive estimate is:

$$\hat{N}_{iR} = n_i/\hat{\varepsilon} = n_i(D/d). \qquad [2]$$

This formula would be a simple "expansion" of the trap counts based on the estimated trap efficiency. The problem with equation [2] is that it involves a ratio of two binomial random variables ($n_i$ and $d$), making the probability distribution of $\hat{N}_{iR}$ difficult to derive directly. However, the large sample bias and variance of $\hat{N}_{iR}$ can be approximated using Bayesian methods and assuming a uniform prior distribution for $\varepsilon$. It turns out that LaPlace used such an analysis in estimating the error in his famous estimate

Table 1. Notation for Dye Marking

| | |
|---|---|
| $D$ | Number of fish dyed and released upstream (single dyed release) |
| $D_i$ | Number of fish dyed and released on night i (daily dyed release) |
| $d$ | Number of dyed fish captured in downstream traps (single dyed release) |
| $d_i$ | Number of dyed fish recaptured from the release on night i (daily dyed release |
| $\varepsilon$ | True trap efficiency = $E[d/D]$ |
| $\hat{\varepsilon}$ | Estimated trap efficiency = $d/D$ |
| $N_i$ | Number of unmarked fish migrating past traps on night i |
| $n_i$ | Number of fish caught in traps on night i |
| $k$ | Number of nights in period of interest |
| $N$ | Total number of fish migrating past traps in k nights = $\sum_{i=1}^{k} N_i$ |
| $n$ | Total number of unmarked fish caught in traps in k nights = $\sum_{i=1}^{k} n_i$ |
| $z_\alpha$ | $1-\alpha/2$ percentage point of standard normal distribution |

Table 2. Assumptions for single dyed release.

1. All of the fish dyed released upstream pass the traps during the period of the study.
2. The probability that a dyed fish enters one of the traps equals $\varepsilon$ for all dyed fish.
3. The probability that an unmarked fish enters one of the traps equals $\varepsilon$ for all unmarked fish.
4. Fish are caught or not caught in the traps independently of the fate of other fish.
5. Trap efficiency $\varepsilon$ is constant for all k days of the study.
6. All of the fish entering the traps throughout the period of the study are counted.

of the population of France based on the ratio of births to total population in certain districts (LaPlace 1820, Engl. Trans. 1951). LaPlace's assumptions are equivalent to assumptions 1) and 2) above, and his results can therefore be applied to the dye marking problem.

## Analysis of the Model using LaPlace's Method

Cochran (1978) describes LaPlace's method for determining the bias and the variance of the ratio estimate in equation [2]. It is possible to make an analogy between the variables of LaPlace's study and the variables in the dye marking situation (Table 3). With the analogy between variables made, it is then clear that statistical assumptions 1) and 2) noted above (Table 2) correspond to LaPlace's statistical assumptions (Cochran 1978):

i) x is binomial (y, p),

ii) X is binomial (Y, p), and

iii) x and X are independent.

Cochran reviews LaPlace's logic leading from these assumptions to the distribution of z, where $z = Y - Y_R$ (in the present case $z = N_i - N_{iR}$).

The result is that for large samples, the distribution of z approximates a normal distribution with

$$E[z] = n_i(D-d)/d^2$$

$$\text{and Var}[z] = n_i(n_i+d)D(D-d)/d^3. \qquad [3]$$

In the above two formulas, I have substituted dye marking notation (Table 1) for Cochran's notation. From equations [3] I derived the approximately unbiased estimator, $N_i$, and the variance and confidence interval formulas in Table 4.

## Daily Release of Dyed Fish

In some circumstances it is appropriate to release marked fish on each day of the sampling period. This would be the case, for example, if water levels fluctuate so frequently that the trap efficiency is likely to change daily by a significant amount. In some cases the opening to a trap must be adjusted in order to function properly as the water level changes. Since such adjustments change the trap efficiency, separate dye tests must be performed at least as often as the trap opening is adjusted. Finally, in some cases, not all of the fish caught in the traps are examined for the presence of a dye. If such subsampling of fish is part of a sampling procedure, then it is imperative that dyed fish be released as often as the traps are sampled.

The statistical analysis of the data obtained when marked fish are released daily is accomplished in two steps. First of all, each day's migration is

Table 3. Correspondence between variable names used by Cochran (1978) and the variables in a dye marking study (Table 1).

| Cochran's Notation | Meaning in LaPlace's Study | Corresponding Dye Marking Notation |
|---|---|---|
| x | Number of births in sampled districts | d |
| y | Total population of sampled districts | D |
| X | Number of births in France | n |
| Y | Population of France | N |
| p | The constant ratio of births to total population in France | ε |

Table 4. Formulas for analyzing dye marking data.

---

1. Single dyed release:

   1a. Daily migration estimates -

   $$\hat{N}_i = n_i \left[\frac{D}{d} + \frac{(D-d)}{d^2}\right] = \frac{n_i}{d}\left[D + \frac{D-d}{d}\right] = \frac{n_i D}{d}\left[1 + \frac{D-d}{Dd}\right]$$

   $$\text{Var}[\hat{N}_i] = n_i\,(n_i + d)\,D\,(D-d)/d^3$$

   $$s = \sqrt{\text{Var}[\hat{N}_i]}$$

   $$(1-\alpha)\text{ C.I. for } N_i = [\hat{N}_i - z_\alpha (s), \hat{N}_i + z_\alpha (s)]$$

   1b. Overall migration estimate - as in 1a., but substitute $\hat{N}$ for $\hat{N}_i$ and $n$ for $n_i$.

2. Daily dyed release:

   2a. Daily migration estimates - as in 1a., but substitute $D_i$ for $D$ and $d_i$ for $d$.

   2b. Overall migration estimate -

   $$\hat{N} = \sum_{i=1}^{k} \hat{N}_i$$

   $$\text{Var}[\hat{N}] = \sum_{i=1}^{k} \text{var}[\hat{N}_i]$$

   $$s = \sqrt{\text{var}[\hat{N}]}$$

---

estimated as a single release of marked fish and a single recovery. The overall seasonal migration estimate is simply the sum of the daily estimates and the variance of the overall estimate is the sum of the daily variances. Using this overall variance, a confidence interval can be calculated for the estimated total number of fish.

The assumptions required for the daily release of dyed fish are summarized in Table 5, special notation is in Table 1, and the appropriate formulas are in Table 4.

## EXAMPLES

The dye marking technique has been used successfully in several FRED Division projects. The following three examples will illustrate the application of the methods discussed in this report to different situations. The raw data and calculated results are presented for each example so the reader may verify the use of the formulas in Table 4.

### Example 1: 1981 Kasilof River Smolt Enumeration (single release of dyed fish)

The sockeye salmon smolt migration in the Kasilof River is monitored as part of the effort to evaluate the success of the fry released in Tustumena Lake (on the Kenai Peninsula) from the Crooked Creek Hatchery near Kasilof (Flagg 1982). During 1981 the traps were operated from 7 May through 1 July, and dyed fish were released on four different days, 1 - 2 weeks apart.

The trap efficiency appeared to decline over time (Figure 1); perhaps, this was due to increasing water flow rates. If this apparent decline had been statistically significant, then it would not have been appropriate to combine all four dye tests for the analysis, and the overall estimate would have had to have been obtained in some other way (e.g., as in Example 3 below). However, a chi-square test showed that the percent recoveries in the four dye tests were not significantly different ($\chi^2 = 3.12$, d.f. = 3); so, in this case, it was acceptable to combine the tests. The possibility of a significant decline in percent recovery over time, however, illustrates the need for performing several dye tests during the period of migration.

Because the results from all four dye tests could be combined in this case, this is an example of a single release of dyed fish. Therefore, the data are analyzed as if a single dye test had been performed. The total number of fish dyed and released upstream was 2,560, and 176 of these were recaptured. A total of 155,596 unmarked smolts were captured in the traps during the season. Substituting these numbers into the formulas in Table 4 gives an estimate of $2.276 \times 10^6$ for the total smolt migration and a 95% confidence interval of $[1.953 \times 10^6, 2.599 \times 10^6]$.

Besides yielding an estimate with confidence intervals of the total smolt migration, this method allows one to estimate each day's outmigration with confidence intervals. Figure 2 shows the upper and lower 95% confidence

Table 5. Assumptions for daily release of dyed fish.

---

1. All marked fish released upstream on a given night are either caught that night or pass the traps that night. No marked fish from a given night's release are caught on subsequent nights.

2. On a given night, i, the probability that a marked fish is caught equals $\varepsilon_i$ for each marked fish released.

3. On a given night, i, the probability that an unmarked fish is caught equals $\varepsilon_i$ for each unmarked fish migrating that night.

4. A given fish is caught or not caught independently of the fate of other fish.

5. If only a portion of the fish in the traps are examined for marks rather than the entire catch then this sample is selected randomly with respect to marked and unmarked fish.

6. Marked or unmarked fish are never in double jeopardy of being caught either during a given night or on two different nights.

---

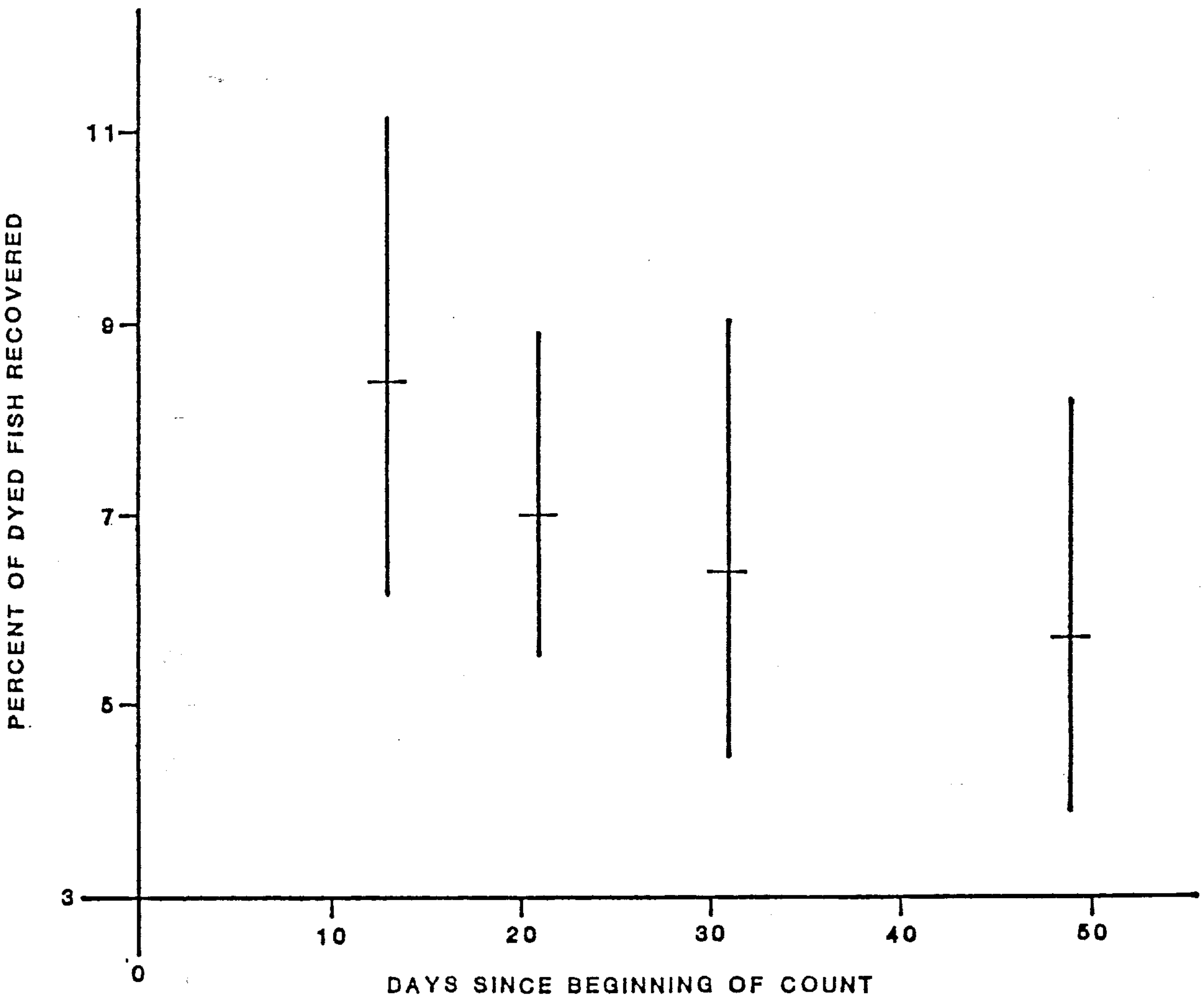


FIGURE 1. Estimated trap efficiencies and 95% confidence intervals for the four dye tests, Kasilof River, 1981.

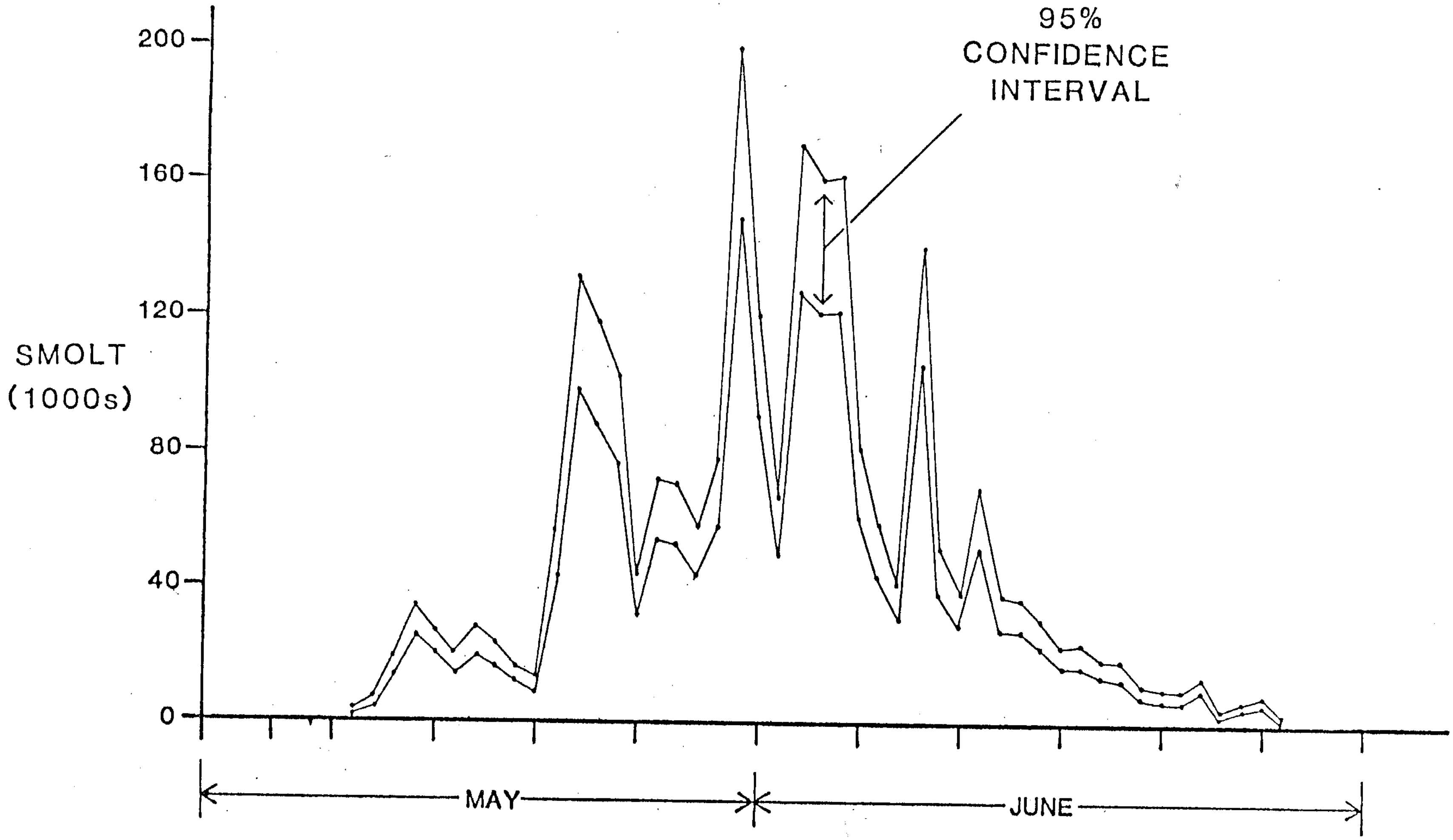


FIGURE 2. Estimated daily smolt migration numbers with 95% confidence intervals, Kasilof River, 1981.

bounds for the smolt outmigration for each day of the 1981 run. These confidence limits are obtained by applying the formulas in Table 4 to the daily smolt counts, as indicated in Table 6. The most striking conclusion from Figure 2 is that the observed spikes in the magnitude of the run are "real" within the bounds of the 95% confidence intervals. Therefore, the sampling effort used in this case was adequate to monitor the day-to-day timing of the smolt migration.

Example 2: 1982 Upper Thumb River Fry Enumeration (daily release of dyed fish)

On the Upper Thumb River, above Karluk Lake on Kodiak Island sockeye, fry are enumerated each year to determine the emergence rate from the eyed eggs that were previously planted upstream. This project is fully described in a report by L. White (in prep.). In the spring of 1982, a large Canadian fan trap with wings was placed in a shallow part of the river several miles downstream from the egg planting site. The fish swimming into the trap were held in a live box in the stream behind the trap. Each morning the live box was emptied and a sample of the fish was examined for marks. Counts were kept of the total numbers of marked and unmarked fish in the sample. The other fish were emptied from the live box and held for release the next evening. A number of the newly caught, unmarked fish were placed in Bismarck Brown dye and after being colored were held in a live box several hundred yards upstream of the trap. Just after dark each evening, the marked fish were released into the stream (White 1983).

The Thumb River fry enumeration is clearly a daily release of dyed fish. The proportion of the total run examined for marks changes from day to day for three reasons: (1) the rates of stream flow fluctuate daily in this relatively small river, (2) the trap may be adjusted daily to prevent crowding in the live box, and (3) an unknown fraction of the fish caught are examined for marks. Therefore, it is necessary to mark and release fish every day in order to separately monitor each day's trap efficiency.

There were several logistical complications in the Thumb River situation. Early in the season, but after the fry migration had started, there were several periods of freezing when the trap could not be operated. Late in the migration period, there were a few days when the traps could not be fished because of high stream flow rates. Using the methods discussed in this report it is impossible to estimate the number of fry moving out during such periods.

The raw data for the 1982 Upper Thumb evaluation are presented in Table 7, and the estimates calculated from these are in Table 8. Each day's migration is estimated as a separate single release and recovery of dyed fish, and the confidence interval for each daily estimate is calculated from that same day's data. No estimates are made for days when there was no fishing. Similarly no estimate is made for the days when no marked fish were released. In this situation, it is not possible to use a previous day's estimated trap efficiency to estimate the fraction of the run examined on a day when no marks were released. This is so because of the three factors cited above, which cause unknown daily fluctuations in the trap efficiency.

Table 6. Daily smolt counts at the Kasilof River traps and estimated daily smolt outmigration with 95% confidence intervals.

| Date | Smolts Counted | Estimated Total Run | 95% Confidence Interval From | To |
|---|---|---|---|---|
| May 7 | 25 | 360 | 220 | 510 |
| 8 | 35 | 510 | 330 | 690 |
| 9 | 59 | 860 | 620 | 1,100 |
| 10 | 150 | 2,180 | 1,730 | 2,640 |
| 11 | 188 | 2,730 | 2,200 | 3,270 |
| 12 | 386 | 5,610 | 4,660 | 6,570 |
| 13 | 1,139 | 16,560 | 14,040 | 19,080 |
| 14 | 2,151 | 31,300 | 26,700 | 35,900 |
| 15 | 1,620 | 23,600 | 20,000 | 27,100 |
| 16 | 1,293 | 18,800 | 16,000 | 21,600 |
| 17 | 1,682 | 24,500 | 20,800 | 28,100 |
| 18 | 1,411 | 20,500 | 17,440 | 23,600 |
| 19 | 977 | 14,200 | 12,020 | 16,400 |
| 20 | 807 | 11,700 | 9,900 | 13,600 |
| 21 | 3,464 | 50,400 | 43,100 | 57,700 |
| 22 | 8,012 | 116,500 | 99,800 | 133,100 |
| 23 | 7,169 | 104,200 | 89,300 | 119,200 |
| 24 | 6,219 | 90,400 | 77,500 | 103,400 |
| 25 | 2,648 | 38,500 | 32,900 | 44,100 |
| 26 | 4,438 | 64,500 | 55,200 | 73,800 |
| 27 | 4,305 | 62,600 | 53,600 | 71,600 |
| 28 | 3,550 | 51,600 | 44,100 | 59,100 |
| 29 | 4,752 | 69,100 | 59,100 | 79,000 |
| 30 | 12,075 | 175,600 | 150,600 | 201,000 |
| 31 | 7,333 | 106,600 | 91,400 | 121,900 |
| June 1 | 4,069 | 59,200 | 50,600 | 67,700 |
| 2 | 10,313 | 150,000 | 128,600 | 171,300 |
| 3 | 9,743 | 141,700 | 121,400 | 161,900 |
| 4 | 9,811 | 142,700 | 122,300 | 163,000 |
| 5 | 4,961 | 72,100 | 61,800 | 82,500 |
| 6 | 3,561 | 51,800 | 44,300 | 59,300 |
| 7 | 2,511 | 36,500 | 31,200 | 41,800 |
| 8 | 8,516 | 123,800 | 106,100 | 141,500 |
| 9 | 3,167 | 46,000 | 39,400 | 52,700 |
| 10 | 2,374 | 34,500 | 29,500 | 39,600 |
| 11 | 4,223 | 61,400 | 52,500 | 70,300 |
| 12 | 2,281 | 33,200 | 28,300 | 38,000 |
| 13 | 2,244 | 32,600 | 27,800 | 37,400 |
| 14 | 1,861 | 27,100 | 23,100 | 31,100 |
| 15 | 1,382 | 20,100 | 17,080 | 23,100 |

- Continued -

Table 6 Continued.

| Date | Smolts Counted | Estimated Total Run | 95% Confidence Interval From | To |
|---|---|---|---|---|
| 16 | 1,451 | 21,100 | 17,940 | 24,300 |
| 17 | 1,181 | 17,200 | 14,570 | 19,770 |
| 18 | 1,094 | 16,000 | 13,480 | 18,330 |
| 19 | 678 | 9,900 | 8,300 | 11,420 |
| 20 | 637 | 9,300 | 7,780 | 10,740 |
| 21 | 598 | 8,700 | 7,300 | 10,090 |
| 22 | 889 | 12,900 | 10,920 | 14,930 |
| 23 | 258 | 3,750 | 3,060 | 4,440 |
| 24 | 452 | 6,570 | 5,480 | 7,670 |
| 25 | 541 | 7,870 | 6,590 | 9,150 |
| 26 | 206 | 3,000 | 2,420 | 3,570 |
| 27 | 141 | 2,050 | 1,620 | 2,490 |
| 28 | 146 | 2,120 | 1,680 | 2,570 |
| 29 | 254 | 3,700 | 3,010 | 4,370 |
| 30 | 102 | 1,490 | 1,140 | 1,830 |
| July 1 | 43 | 620 | 430 | 830 |
| Overall | 155,596 | 2,276,000 | 1,953,000 | 2,599,000 |

Table 7. Upper Thumb River dye marking - 1982 (basic data)

| Date | Total Fry Captured (n+d) | Marked Fry Released (D) | Marked Fry Recaptured (d) | Unmarked Fry Captured (n) | % of Marks Recaptured (100 d/D) |
|---|---|---|---|---|---|
| 19 Mar | 154 | 140 | 5 | 149 | 3.6 |
| 20 Mar | 141 | 143 | 0 | 141 | 0.0 |
| | 108 | 108 | 10 | 98 | 9.3 |
| | 87 | 83 | 9 | 78 | 10.8 |
| | 21 | 65 | 2 | 19 | 3.1 |
| | 0 | 0 | 0 | 0 | 0.0 |
| 25 Mar | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 118 | 0 | 0 | 118 | 0.0 |
| | 180 | 110 | 8 | 172 | 7.3 |
| | 41 | 99 | 6 | 35 | 6.1 |
| 30 Mar | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 26 | 0 | 0 | 0.0 |
| 01 Apr | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| 05 Apr | 146 | 0 | 0 | 146 | 0.0 |
| | 155 | 143 | 23 | 132 | 16.1 |
| | 264 | 129 | 26 | 238 | 20.2 |
| | 596 | 232 | 57 | 539 | 24.6 |
| | 1,137 | 524 | 217 | 920 | 41.4 |
| 10 Apr | 1,012 | 998 | 248 | 764 | 24.8 |
| | 245 | 735 | 128 | 117 | 17.4 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| 15 Apr | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| 20 Apr | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 1,333 | 1,079 | 97 | 1,236 | 9.0 |
| | 2,010 | 2,028 | 104 | 1,906 | 5.1 |
| | 3,070 | 4,048 | 201 | 2,869 | 5.0 |
| 25 Apr | 4,564 | 3,793 | 376 | 4,188 | 9.9 |
| | 3,007 | 4,463 | 341 | 2,666 | 7.6 |
| | 2,843 | 2,968 | 185 | 2,658 | 6.2 |
| | 1,226 | 2,614 | 212 | 1,014 | 8.1 |
| | 3,131 | 1,191 | 173 | 2,958 | 14.5 |
| 30 Apr | 5,505 | 2,907 | 682 | 4,823 | 23.5 |

- Continued -

Table 7 Continued.

| Date | Total Fry Captured (n+d) | Marked Fry Released (D) | Marked Fry Recaptured (d) | Unmarked Fry Captured (n) | % of Marks Recaptured (100 d/D) |
|---|---|---|---|---|---|
| 01 May | 4,863 | 3,804 | 536 | 4,327 | 14.1 |
| | 2,480 | 4,277 | 604 | 1,876 | 14.1 |
| | 2,339 | 1,849 | 215 | 2,124 | 11.6 |
| | 3,629 | 2,104 | 213 | 3,416 | 10.1 |
| 05 May | 3,565 | 3,390 | 146 | 3,419 | 4.3 |
| | 4,547 | 5,980 | 299 | 4,248 | 5.0 |
| | 6,059 | 6,030 | 251 | 5,808 | 4.2 |
| | 8,114 | 6,027 | 264 | 7,850 | 4.4 |
| | 1,283 | 7,463 | 401 | 882 | 5.4 |
| 10 May | 1,369 | 1,251 | 46 | 1,323 | 3.7 |
| | 4,004 | 3,091 | 236 | 3,768 | 7.6 |
| | 3,973 | 3,965 | 208 | 3,765 | 5.2 |
| | 4,595 | 6,703 | 444 | 4,151 | 6.6 |
| | 3,949 | 4,543 | 244 | 3,705 | 5.4 |
| 15 May | 6,638 | 3,889 | 318 | 6,320 | 8.2 |
| | 3,718 | 6,571 | 467 | 3,251 | 7.1 |
| | 3,190 | 5,198 | 382 | 2,808 | 7.3 |
| | 4,156 | 3,171 | 188 | 3,968 | 5.9 |
| | 4,433 | 4,798 | 689 | 3,744 | 14.4 |
| 20 May | 2,873 | 4,429 | 219 | 2,654 | 4.9 |
| | 3,324 | 2,867 | 260 | 3,064 | 9.1 |
| | 4,082 | 3,314 | 315 | 3,767 | 9.5 |
| | 2,354 | 4,080 | 485 | 1,869 | 11.9 |
| | 969 | 2,837 | 227 | 742 | 8.0 |
| 25 May | 1,466 | 942 | 34 | 1,432 | 3.6 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| | 0 | 0 | 0 | 0 | 0.0 |
| 29 May | 112 | 1,455 | 52 | 60 | 3.6 |

Table 8. Upper Thumb River dye marking - 1982 (daily migration estimates)

| Date | Estimated Daily Migration | 95% Conf. Limits Lower | 95% Conf. Limits Upper | Estimated Cumulative Migration | 95% Conf. Limits Lower | 95% Conf. Limits Upper |
|---|---|---|---|---|---|---|
| 19 Mar | 4,980 | 1,330 | 8,630 | 4,980 | 1,330 | 8,630 |
| 20 Mar | 0 | 0 | 0 | 4,980 | 1,330 | 8,630 |
| | 1,150 | 500 | 1,810 | 6,130 | 2,420 | 9,840 |
| | 790 | 320 | 1,260 | 6,920 | 3,190 | 10,660 |
| | 920 | 30 | 1,800 | 7,840 | 4,000 | 11,680 |
| | 0 | 0 | 0 | 7,840 | 4,000 | 11,680 |
| 25 Mar | 0 | 0 | 0 | 7,840 | 4,000 | 11,680 |
| | 0 | 0 | 0 | 7,840 | 4,000 | 11,680 |
| | 0 | 0 | 0 | 7,840 | 4,000 | 11,680 |
| | 2,640 | 1,020 | 4,250 | 10,480 | 6,310 | 14,650 |
| | 670 | 180 | 1,150 | 11,150 | 6,950 | 15,340 |
| 30 Mar | 0 | 0 | 0 | 11,150 | 6,950 | 15,340 |
| | 0 | 0 | 0 | 11,150 | 6,950 | 15,340 |
| 01 Apr | 0 | 0 | 0 | 11,150 | 6,950 | 15,340 |
| | 0 | 0 | 0 | 11,150 | 6,950 | 15,340 |
| | 0 | 0 | 0 | 11,150 | 6,950 | 15,340 |
| | 0 | 0 | 0 | 11,150 | 6,950 | 15,340 |
| 05 Apr | 0 | 0 | 0 | 11,150 | 6,950 | 15,340 |
| | 850 | 520 | 1,180 | 12,000 | 7,790 | 16,200 |
| | 1,220 | 790 | 1,640 | 13,210 | 8,980 | 17,440 |
| | 2,220 | 1,700 | 2,740 | 15,440 | 11,170 | 19,700 |
| | 2,230 | 1,980 | 2,480 | 17,660 | 13,390 | 21,900 |
| 10 Apr | 3,080 | 2,700 | 3,470 | 20,700 | 16,500 | 25,000 |
| | 680 | 520 | 830 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| 15 Apr | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| 20 Apr | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 0 | 0 | 0 | 21,400 | 17,130 | 25,700 |
| | 13,880 | 11,170 | 16,590 | 35,300 | 30,200 | 40,400 |
| | 37,500 | 30,400 | 44,700 | 72,800 | 64,000 | 81,600 |
| | 58,100 | 50,000 | 66,100 | 130,900 | 119,000 | 142,800 |
| 25 Apr | 42,300 | 38,100 | 46,600 | 173,200 | 160,600 | 185,800 |
| | 35,000 | 31,200 | 38,800 | 208,000 | 195,000 | 221,000 |
| | 42,900 | 36,700 | 49,000 | 251,000 | 237,000 | 266,000 |
| | 12,560 | 10,780 | 14,330 | 264,000 | 249,000 | 278,000 |
| | 20,500 | 17,580 | 23,400 | 284,000 | 269,000 | 299,000 |
| 30 Apr | 20,600 | 19,140 | 22,000 | 305,000 | 290,000 | 320,000 |

- Continued -

Table 8 Continued.

| Date | Estimated Daily Migration | 95% Conf. Limits Lower | Upper | Estimated Cumulative Migration | 95% Conf. Limits Lower | Upper |
|---|---|---|---|---|---|---|
| 01 May | 30,800 | 28,200 | 33,300 | 335,000 | 320,000 | 351,000 |
| | 13,300 | 12,170 | 14,400 | 349,000 | 333,000 | 364,000 |
| | 18,340 | 15,930 | 20,800 | 367,000 | 352,000 | 383,000 |
| | 33,900 | 29,500 | 38,300 | 401,000 | 385,000 | 417,000 |
| 05 May | 79,900 | 67,000 | 92,800 | 481,000 | 460,000 | 501,000 |
| | 85,200 | 75,500 | 94,900 | 566,000 | 543,000 | 589,000 |
| | 140,100 | 122,800 | 157,300 | 706,000 | 678,000 | 735,000 |
| | 179,900 | 158,400 | 201,000 | 886,000 | 850,000 | 922,000 |
| | 16,450 | 14,570 | 18,340 | 902,000 | 867,000 | 938,000 |
| 10 May | 36,700 | 26,400 | 47,100 | 939,000 | 902,000 | 976,000 |
| | 49,500 | 43,300 | 55,800 | 989,000 | 951,000 | 1,027,000 |
| | 72,100 | 62,300 | 81,900 | 1,061,000 | 1,022,000 | 1,100,000 |
| | 62,800 | 56,900 | 68,700 | 1,124,000 | 1,084,000 | 1,163,000 |
| | 69,300 | 60,600 | 77,900 | 1,193,000 | 1,152,000 | 1,233,000 |
| 15 May | 77,500 | 69,200 | 85,900 | 1,270,000 | 1,229,000 | 1,312,000 |
| | 45,800 | 41,600 | 50,100 | 1,316,000 | 1,275,000 | 1,358,000 |
| | 38,300 | 34,400 | 42,200 | 1,255,000 | 1,313,000 | 1,396,000 |
| | 67,300 | 57,800 | 76,800 | 1,422,000 | 1,279,000 | 1,465,000 |
| | 26,100 | 24,100 | 28,100 | 1,448,000 | 1,405,000 | 1,491,000 |
| 20 May | 53,900 | 46,700 | 61,100 | 1,502,000 | 1,458,000 | 1,545,000 |
| | 33,900 | 29,800 | 38,000 | 1,536,000 | 1,492,000 | 1,579,000 |
| | 39,700 | 35,400 | 44,100 | 1,575,000 | 1,532,000 | 1,619,000 |
| | 15,750 | 14,280 | 17,230 | 1,591,000 | 1,547,000 | 1,635,000 |
| | 9,310 | 7,990 | 10,630 | 1,601,000 | 1,557,000 | 1,644,000 |
| 25 May | 0 | 0 | 0 | 1,601,000 | 1,557,000 | 1,644,000 |
| | 40,800 | 27,600 | 54,000 | 1,641,000 | 1,596,000 | 1,687,600 |
| | 0 | 0 | 0 | 1,641,000 | 1,596,000 | 1,687,600 |
| | 0 | 0 | 0 | 1,641,000 | 1,596,000 | 1,687,600 |
| 29 May | 1,710 | 1,100 | 2,320 | 1,643,000 | 1,597,000 | 1,689,000 |

Besides the daily estimates, estimates of the cumulative numbers of migrating fish with confidence intervals may be obtained for the 1982 Upper Thumb data (Table 8). The estimated cumulative total is simply the sum of the daily estimates for all previous days. The confidence interval for the cumulative estimates is determined by assuming that the variance of the estimated cumulative total is the sum of the variances of all previous nights' estimates (Table 4). The relative width of the confidence interval is much smaller for the cumulative estimate than it is for the individual daily estimates (Table 8).

Example 3: 1981 Crescent Lake Smolt Enumeration (hybrid)

The sockeye salmon smolt migration from Crescent Lake, on the west side of Cook Inlet, is monitored as part of a comprehensive pre-fertilization evaluation. The 1981 study is fully described in a report by Kyle and Koenings (1982).

In this case, migrating smolts were captured in two fyke nets placed in the Crescent River. At least once each day, the fish caught in the traps were counted. Releases of dyed fish were made weekly, upstream of the traps at the lake outlet. The percent of dyed fish recovered changed weekly (Table 9), and a chi-square test showed them to be significantly different ($\chi^2 = 448$, $P < 0.005$). Therefore, unlike Example 1, it is not possible to lump all of the releases and analyze the data as a single release of dyed fish. Instead, the data were analyzed by week in a manner similar to the procedure for daily releases of dyed fish, with each week being a separate unit instead of each day as in the case of Example 2. This approach is a hybrid that lies between the single release and daily release of dyed fish: the data are analyzed similarly to a daily release of dyed fish , but it is important that the trap efficiency and sampling fraction remain constant within each week. In this case, the sampling fraction was 100% for the entire study, so the only assumption necessary is that the trap efficiency was constant during each week.

The weekly estimates and overall estimate with confidence intervals are shown in Table 10. The method of combining the weekly estimates to derive an overall estimate is the same one used in Example 2: the overall estimate is obtained by summing the weekly estimates, and the overall variance is obtained by summing the variances of the weekly estimates.

## DISCUSSION

The methods outlined in this report provide a way of obtaining an estimate and a confidence interval for the number of salmon fry or smolts migrating down a river. If the necessary assumptions hold, then the method is statistically sound. The process of dyeing fry and smolts has been shown to work, and it has been established that these marked fish can be recognized upon recapture. The examples discussed in this report illustrate the kinds of situations where the method is likely to work.

Table 9. Summary of releases and recoveries of dyed fish, Crescent River, 1981 (modified from Table 1 of Kyle and Koenings 1982).

| Date Marked | Recovery Period | No. Marked and Released | Number Recovered | Percent Recovered |
|---|---|---|---|---|
| 6/04 | 6/04 - 6/08 | 218 | 6 | 2.8 |
| 6/12 | 6/12 - 6/15 | 297 | 12 | 4.0 |
| 6/16 | 6/17 - 6/20 | 304 | 2 | 0.7 |
| 6/25 | 6/25 - 6/28 | 298 | 13 | 4.4 |
| 7/03[a/] | 7/04 - 7/07 | 68 | 12 | 17.6 |

[a/] The dye used and conditions of this release were slightly different than the other releases. See Kyle and Koenings (1982).

Table 10. Weekly and overall estimates of smolt migration Crescent River, 1981 (modified from Table 2 of Kyle and Koenings 1982).

| Period | Unmarked Smolts Caught | Estimated Total Migration (1000s) | 95% Confidence Interval Lower | 95% Confidence Interval Upper |
|---|---|---|---|---|
| 5/20 - 5/26 | 4 | | | |
| 5/27 - 6/02 | 778 | | | |
| 6/03 - 6/09 | 5,849 | 246.7 | 78.8 | 414.7 |
| 6/10 - 6/16 | 10,695 | 286.2 | 139.2 | 433.2 |
| 6/17 - 6/23 | 5,954 | 1,358.0 | 107.2 | 2,608.7 |
| 6/24 - 6/30 | 4,825 | 119.0 | 60.0 | 178.1 |
| 7/01 - 7/07 | 301 | 1.9 | 1.0 | 2.8 |
| Overall | - | 2,011.9 | 740.5 | 3,283.4 |

The approach discussed in this report is not the same thing as the Petersen method (Seber 1973). It is wrong to attempt to apply the usual statistical methods used in capture-recapture work (e.g., Seber 1973) to the problem of enumerating migrating smolts. The Petersen approach was designed for a different kind of situation. For example, suppose $n_1$ fish are marked and released in a lake and later $m_2$ of these are recovered in a total sample of $n_2$. The Petersen approach says that the estimate of the lake's total population, $N_1$, equals $n_1n_2/m_2$. Here the total population of the lake includes the marked fish. Now, suppose that D fish are marked and released upstream and d of these are recovered along with n unmarked fish. The D marked fish have already been caught in the traps on previous days and should not be included twice in the population estimate. Therefore, the intuitive estimate of the total run strength, $\hat{N}_2$, is nD/d.

The Petersen approach, however, would be to estimate the population by $\hat{N}_1 = (n+d)D/d$. Notice that $N_1 = N_2 + D$. Therefore, to enumerate the migrating population using the Petersen approach is to erroneously include the marked fish in the estimate. A similar argument would establish that the variance formulas for a Petersen estimate (Seber 1973) will not work for smolt enumeration.

The statistical approach presented here is of course valid for any method of marking fish that does not cause excess mortality, that does not affect the behavior of the fish, and that makes marked fish easily distinguishable from unmarked ones. Several different dyes have been tried by FRED Division personnel, but Bismarck Brown is the only one that meets all three of these criteria. Alternatives to dye marking include fluorescent pigment marking (Gray et al. 1978) and aluminum staple tag marking (Jordan and Smith 1968). The latter method enables one to identify fish with a different mark for each day of marking. MacDonald and Smith (1980) have developed a fairly complex statistical methodology for analyzing data from daily releases of salmon smolts marked with aluminum staple tags. Their approach would be appropriate in cases where the assumptions of the method outlined in the present report do not apply. For example, in some FRED Division studies, marked sockeye salmon smolts have been released in the lake instead of in the river upstream of the traps. In such cases, it has been found that the smolts often do not migrate again quickly. This situation makes it difficult to monitor trap efficiency on a day-by-day or week-by-week basis. In situations where marked smolts must be released in the lake, the aluminum staple tag marking method and associated statistical analysis may be more appropriate than dye marking. However, the logistical problems in such an approach may be difficult.

In summary, the statistical methods described in this report provide a way of using dye marking data to estimate migrant juvenile salmon populations with a confidence interval.

ACKNOWLEDGEMENTS

My understanding of some of the logistical problems associated with dye marking was made possible by visits to several projects and on site discussions with field personnel, especially R. Neterer, C. Neterer, M. Rowse, and D. Waite. Several discussions with L. Flagg and L. White in large part contributed to my conception of the kinds of problems for which statistical solutions were needed. The basic statistical model was formulated during a conversation with D. Reed. I am grateful to M. Kaill and W. Hauser for sharing ideas on the problems discussed here and the encouragement to complete this report. Careful reviews by K. Leon, M. Kaill, W. Hauser, and A. Howe helped to clarify the presentation of this material.

# REFERENCES


Cochran, W. G. 1978. LaPlace's ratio estimator. In H. A. David, ed., contributions to survey sampling and applied statistics. Academic Press, New York, pp. 3-10.

Flagg, Loren. 1982. Sockeye salmon smolt studies Kasilof River, Alaska. ADF&G, FRED draft report.

Fleiss, J. L. 1981. Statistical methods for rates and proportions. John Wiley & Sons.

Gray, P. L., K. R. Florey, J. F. Koerner, and R. P. Marriott. 1978. Coho salmon fluorescent pigment mark-recovery program for the Taku, Berners, and Chilkat Rivers in Southeastern Alaska (1972-1974). ADF&G Informational Leaflet #176, 75 pp.

Jordan, F. P. and H. D. Smith. 1968. An aluminum staple tag for population estimates of salmon smolts. Prog. Fish. Cult. 30:230-234.

Kyle, G. B. and J. P. Koenings. 1982. Crescent Lake Sockeye Salmon Smolt Enumeration and Sampling. ADF&G, FRED Report 27 pp.

LaPlace, P. S. 1820. A philosophical essay on probabilities. English translation, Dover, New York, 1951 (cited by Cochran, 1978).

MacDonald, P. P. M. and H. D. Smith. 1980. Mark-recapture estimation of salmon smolt runs. Biometrics 36:401-417.

Seber, G. A. F. 1973. Estimation of animal abundance and related parameters. Griffin, London.

White, L. 1983. Karluk Lake Sockeye rehabilitation. 1982. ADF&G, FRED Division Draft Report, 42pp.