THE DESIGN AND ANALYSIS OF SALMONID TAGGING STUDIES IN THE COLUMBIA BASIN

VOLUME XXIV

A Statistical Critique of Estimating Salmon Escapement in the Pacific Northwest

Prepared by:
Amber L. Parsons and John R. Skalski
School of Aquatic and Fishery Sciences
University of Washington
1325 Fourth Avenue, Suite 1820
Seattle, WA 98101-2509

Prepared for:
U.S. Department of Energy
Bonneville Power Administration
Division of Fish and Wildlife
P.O. Box 3621
Portland, OR 97208-3621

Project No. 1989-107-00
Contract No. 00039987

June 2009

The Design and Analysis of Salmonid Tagging Studies in the Columbia Basin

Other Publications in this Series

Volume I: Skalski, J. R., J. A. Perez-Comas, R. L. Townsend, and J. Lady. 1998. Assessment of temporal trends in daily survival estimates of spring chinook, 1994-1996. Technical report submitted to BPA, Project 89-107-00, Contract DE-BI79-90BP02341. 24 pp. plus appendix.

Volume II: Newman, K. 1998. Estimating salmonid survival with combined PIT-CWT tagging. Technical report (DOE/BP-35885-11) to BPA, Project 91-051-00, Contract 87-BI-35885.

Volume III: Newman, K. 1998. Experiment designs and statistical models to estimate the effect of transportation on survival of Columbia River system salmonids. Technical report (DOE/BP-35885-11a) to BPA, Project 91-051-00, Contract 87-BI-35885.

Volume IV: Perez-Comas, J. A., and J. R. Skalski. Submitted. Preliminary assessment of the effects of pulsed flows on smolt migratory behavior. Technical report to BPA, Project 89-107-00, Contract DE-BI79-90BP02341.

Volume V: Perez-Comas, J. A., and J. R. Skalski. Submitted. Analysis of in-river growth for PIT-tagged spring chinook smolt. Technical report to BPA, Project 89-107-00, Contract DE-BI79-90BP02341.

Volume VI: Skalski, J. R., J. A. Perez-Comas, P. Westhagen, and S. G. Smith. 1998. Assessment of season-wild survival of Snake River yearling chinook salmon, 1994-1996. Technical report to BPA, Project 89-107-00, Contract DE-BI79-90BP02341. 23 pp. plus appendix.

Volume VII: Lowther, A. B., and J. R. Skalski. 1998. Monte-Carlo comparison of confidence interval procedures for estimating survival in a release-recapture study, with applications to Snake River salmonids. Technical report (DOE/BP-02341-5) to BPA, Project 89-107-00, Contract 90-BI-02341.

Volume VIII: Lowther, A. B., and J. R. Skalski. 1998. Improved survival and residualization estimates for fall chinook using release-recapture methods. Technical report (DOE/BP-02341-6) to BPA, Project 89-107-00, Contract 90-BI-02341.

Volume IX: Townsend, R. L., and J. R. Skalski. Submitted. A comparison of statistical methods of estimating treatment-control ratios (transportation benefit ratios), based on spring chinook salmon on the Columbia River, 1986-1988. Technical report to BPA, Project 91-051-00, Contract 87-BI-35885.

Volume X: Westhagen, P., and J. R. Skalski. 1998. Instructional guide to using program CaptHist to create SURPH files for survival analysis using PTAGIS data files. Technical report (DOE/BP-02341-4) to BPA, Project 89-107-00, Contract 90-BI-02341.

Volume XI: Skalski, J. R., R. L. Townsend, A. E. Giorgi, and J. R. Stevenson. Submitted. Recommendations on the design and analysis of radiotelemetry studies of salmonid smolts to estimate survival and passage efficiencies. Technical report to BPA, Project 89-107-00, Contract DE-BI79-90BP02341. 33 pp.

Volume XII: Ryding, K. E., and J. R. Skalski. 1999. A multinomial model for estimating ocean survival from salmonid coded wire-tag data. Technical report (DOE/BP-91572-3) to BPA, Project 91-051-00, Contract 96-BI-91572.

Volume XIII: Perez-Comas, J. A., and J. R. Skalski. 2000. Appraisal of system-wide survival estimation of Snake River yearling chinook salmon using PIT-tags recovered from Caspian tern and double-crested cormorant breeding colonies on Rice Island. Technical report to BPA, Project No. 8910700, Contract DE-BI79-90BP02341.

Volume XIV: Perez-Comas, J. A., and J. R. Skalski. 2000. Appraisal of the relationship between tag detection efficiency at Bonneville Dam and the precision in estuarine and marine survival estimates of returning pit-tagged chinook salmon. Technical report to BPA, Project No. 8910700, Contract DE-BI79-90BP02341.

Volume XV: Perez-Comas, J. A., and J. R. Skalski. 2000. Appraisal of the relationship between tag detection efficiency at Bonneville Dam and the precision in-river survival estimates of returning PIT-tagged chinook salmon. Technical report to BPA, Project No. 8910700, Contract DE-BI79-90BP02341.

Volume XVI: Skalski, J. R., and J. A. Perez-Comas. 2000. Alternative designs for future adult PIT-tag detection studies. Technical report to BPA, Project No. 8910700, Contract DEBI79-90BP02341.

Volume XVII: Burgess, C. A., and J. R. Skalski. 2001. Effects of ocean covariates and release timing on first ocean-year survival of fall Chinook salmon from Oregon and Washington coastal hatcheries. Technical report to BPA, Project No. 199105100, Contract 1996BI91572.

Volume XVIII: Burgess, C. A., T. J. Miller, and J. R. Skalski. 2003. Precision and accuracy of the transportation-to-inriver (T/I) ratio estimator of survival benefits to juvenile salmonids transported around the Columbia River Basin dams. Technical report to BPA, Project No. 19910500, Contract 00013690.

Volume XIX: Buchanan, R. A., and J. R. Skalski. 2006. Analysis of fall Chinook salmon PIT-tag data: Estimating transportation effects. Technical report to BPA, Project No. 198910700, Contract 00012494.

Volume XX: J. R. Skalski. 2006. Evaluation and recommendations on alternative hydroacoustic array deployments for the mouth of the Columbia River to provide estimates of salmonid smolt survival and movements. Technical report to BPA, Project No. 198910700, Contract 00012494.

Volume XXI: J. R. Skalski and J. Griswold. 2006. A summary of methods for conducting salmonid fry mark-recapture studies for estimating survival in tributaries. Technical report to BPA, Project No. 198910700, Contract 00012494.

Volume XXII: R. W. Perry and J. R. Skalski. 2008. Evaluating wetland restoration projects in the Columbia River Estuary using hydroacoustic telemetry arrays to estimate movement, survival, and residence times of juvenile salmonids. Technical report to BPA, Project No. 198910700, Contract 00012494.

Volume XXIII: R. W. Perry and J. R. Skalski. 2008. Effects of array configuration on statistical independence of replicated telemetry arrays used in smolt survival studies. Technical report to BPA, Project No. 198910700, Contract 00012494.

Other Publications Related to this Series

Other related publications, reports and papers available through the professional literature or from the Bonneville Power Administration (BPA) Public Information Center - CKPS-1, P.O. Box 3621, Portland, OR 97208.

Skalski, J. R., R. A. Buchanan, and J. Griswold. 2009. Review of marking methods and release-recapture designs for estimating salmonid fry survival in tributary waters. Reviews in Fisheries Science 17(3):391–401.

Buchanan, R. A., and J. R. Skalski. 2007. A migratory life-cycle release-recapture model for salmonid PIT-tag investigations. Journal of Agricultural, Biological, and Environmental Statistics 12(3):325-345.

Buchanan, R. A., J. R. Skalski, and S. G. Smith. 2006. Estimating the effects of smolt transportation from different vantage points and management perspectives. North American Journal of Fisheries Management 26:460-472.

Townsend, R. L., J. R. Skalski, P. Dillingham, and T. W. Steig. 2006. Correcting bias in survival estimation resulting from tag failure in acoustic and radiotelemetry studies. Journal of Agricultural, Biological, and Environmental Statistics 11:183-196.

Johnson, G. E., J. B. Hedgepeth, J. R. Skalski, and A. E. Giorgi. 2004. A Markov chain analysis of fish movements to determine entrainment zones. Fisheries Research 69(3):349-358.

Hedgepeth, J. B., G. E. Johnson, J. R. Skalski, and J. Burczynski. 2002. Active fish sonar (AFTS) for assessing fish behavior. Acta Acoustic United with Acustica 88:739-742.

Skalski, J. R., R. Townsend, J. Lady, A. E. Giorgi, and J. R. Stevenson. 2002. Estimating route-specific passage and survival probabilities at a hydroelectric project from smolt radiotelemetry studies. Canadian Journal of Fisheries and Aquatic Sciences 59: 1385-139.

Skalski, J. R., J. Lady, R. Townsend, A. E. Giorgi, J. R. Stevenson, C. M. Peven, and R. D. McDonald. 2001. Estimating inriver survival of migrating salmonid smolts using radiotelemetry. Canadian Journal of Fisheries and Aquatic Sciences 58:1987-1997.

Lowther, A. B., and J. R. Skalski. 1998. A multinomial likelihood model for estimating survival probabilities and overwintering for fall chinook salmon using release-recapture methods. Journal of Agricultural, Biological, and Environmental Statistics 3: 223- 236.

Skalski, J. R. 1998. Estimating season-wide survival rates of outmigrating salmon smolt in the Snake River, Washington. Canadian Journal of Fisheries and Aquatic Sciences 55: 761-769.

Skalski, J. R., and J. A. Perez-Comas. 1998. Using PIT-tag recapture probabilities to estimate project-wide fish efficiency and spill effectiveness at Lower Granite Dam. School of Fisheries, University of Washington. Report prepared for U.S. Army Corps of Engineers, Contract No. DACW68-96-C0018, Walla Walla District, 201 North Third Street, Walla Walla, WA 99362-9265, 67 pp.

Skalski, J. R., and J. A. Perez-Comas. 1998. Using steelhead and chinook salmon PIT-tag recapture probabilities to estimate FGE and SE at Lower Granite Dam. School of Fisheries, University of Washington. Report prepared for U.S. Army Corps of Engineers, Contract No. DACW68-96-C0018, Walla Walla District, 201 North Third Street, Walla Walla, WA 99362-9265, 44 pp.

Skalski, J. R., S. G. Smith, R. N. Iwamoto, J. G. Williams, and A. Hoffmann. 1998. Use of PIT-tags to estimate survival of migrating juvenile salmonids in the Snake and Columbia Rivers. Canadian Journal of Fisheries and Aquatic Sciences 55:1484-1493.

Newman, K. 1997. Bayesian averaging of generalized linear models for passive integrated tag recoveries from salmonids in the Snake River. North American Journal of Fisheries Management 17: 362-377.

Mathur, D., P. G. Heisey, E. T. Euston, and J. R. Skalski. 1996. Turbine passage survival estimates for chinook salmon smolt (*Oncorhynchus tshawytscha*) at a large dam on the Columbia River. Canadian Journal of Fisheries and Aquatic Sciences 53:542-549.

Skalski, J. R. 1996. Regression of abundance estimates from mark-recapture surveys against environmental variables. Canadian Journal of Fisheries and Aquatic Sciences 53: 196-204.

Skalski, J. R., R. L. Townsend, R. F. Donnelly, and R. W. Hilborn. 1996. The relationship between survival of Columbia River fall chinook salmon and inriver environmental factors: Analysis of historic data for juvenile and adult salmonid production. Final report, Phase II. Technical Report (DOE/BP-35885-10) to BPA, Project 91-051-00, Contract 90-BI-02341.

Smith, S. G., J. R. Skalski, J. R., J. W. Schlechte, A. Hoffmann, and V. Cassen. 1996. Introduction to SURPH.1 analysis of release-recapture data for survival studies. Technical report DOE/BP-02341-3) to BPA, Project 89-107-00, Contract 90-BI-02341.

Newman, K. 1995. Adult salmonid PIT-tag returns to Columbia River's Lower Granite Dam. Technical report (DOE/BP-35885-5) to BPA, Project 91-051-00, Contract 87-BI-35885.

Smith, S. G., J. R. Skalski, J. R., J. W. Schlechte, A. Hoffmann, and V. Cassen. 1994. SURPH.1 Manual: Statistical survival analysis of fish and wildlife tagging studies. Technical report (DOE/BP-02341-2) to BPA, Project 89-107-00, Contract 90-BI-02341.

Dauble, D. D., J. R. Skalski, A. Hoffmann, and A. E. Giorgi. 1993. Evaluation and application of statistical methods for estimating smolt survival. Technical report (DOE/BP-62611-1) to BPA, Project 86-118-00, Contract 90-AI-62611; Project 89-107-00, Contract 90-BI-02341; and Project 91-051-00, Contract 87-BI-35885.

Skalski, J. R., A. Hoffmann, and S. G. Smith. 1993. Development of survival relationships using concomitant variables measured from individual smolt implanted with PIT-tags. Annual report 1990-1991 (DOE/BP-02341-1) to BPA, Project 89-107-00, Contract 90-BI-02341.

Skalski, J. R., and A. E. Giorgi. 1993. Juvenile passage program: A plan for estimating smolt travel time and survival in the Snake and Columbia rivers. Technical report (DOE/BP-35885-3) to BPA, Project 91-051-00, Contract 87-BI-35885.

Smith, S. G., J. R. Skalski, and A. E. Giorgi. 1993. Statistical evaluation of travel time estimation based on data from freeze-branded chinook salmon on the Snake River, 1982-1990. Technical report (DOE/BP-35885-4) to BPA, Project 91-051-00, Contract 87-BI-35885.

Giorgi, A. E. 1990. Mortality of yearling chinook salmon prior to arrival at Lower Granite Dam on the Snake River. Technical report (DOE/BP-16570-1) to BPA, Project 91-051-00, Contract 87-BI-35885.

PREFACE

Project 1989-107-00 was initiated to develop the statistical theory, methods, and statistical software to design and analyze PIT-tag survival studies. This project developed the initial study designs for the NOAA Fisheries/University of Washington (UW) Snake River survival studies of 1993−present. This project continues to respond to the changing needs of the scientific community in the Pacific Northwest as they face new challenges to extract life-history data from an increasing variety of fish-tagging studies. The project's mission is to help assure tagging studies are designed and analyzed from the onset to extract the best available information using state-of-the-art statistical methods. In so doing, investigators can focus on the management implications of their findings without being distracted by concerns of whether the study's design and analyses are correct.

All studies in the current series, the Design and Analysis of Tagging Studies in the Columbia Basin, were conducted to help maximize the amount of information that can be obtained from fish tagging studies for the purposes of monitoring fish survival and related demographic parameters throughout its life cycle. Volume XXIV of this series provides a comprehensive review of the statistical accuracy and precision of the methods currently employed to estimate salmon escapement in the Pacific Northwest.

ABSTRACT

Accurate and precise estimates of escapement are required for effective management of Pacific Northwest Salmon stocks. Most reviews of escapement estimation techniques are outdated and are logistically, rather than statistically, oriented. A comprehensive review of the statistical accuracy and precision of the methods currently employed to estimate salmon escapement in the Pacific Northwest is needed. In this report, the bias and precision of six different types of escapement estimators are derived. The six types of estimators reviewed are the Area-Under-the-Curve (including both fish-days and stream life components), Mark-Recapture (for open and closed populations), Passage Count (dams, fences, and sonar), Peak Count (live and dead), Redd Count, and Carcass Count Methods. The assumptions supporting each method are discussed with regard to the ease of their achievability and the effect of their failure on the estimators. The results of each review are summarized in method-specific tables. The final section includes an overall summary of the “good,” “bad,” and “ugly”' methods, five issues managers need to take into consideration as they select an escapement method, and six suggestions for future research that could benefit salmon escapement estimation techniques.

TABLE OFCONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Funding for this work came from the Pacific Northwest region's electrical ratepayers through the Columbia River Fish and Wildlife Program administered by the Bonneville Power Administration through Project No. 1989-107-00.

Chapter 1

INTRODUCTION TO ESCAPEMENT ESTIMATION

1.1 Escapement Defined

Anadromous fishes hatch and sometimes rear in fresh water, migrate to the ocean until maturity, and then return to fresh water to reproduce. Pacific salmon (*Oncorhynchus* spp.) are commercially valuable and the annual return is a convenient time for harvest. Fish that escape the commercial, tribal, and sport fisheries are termed the "escapement." It is often desirable to distinguish between the numbers of returning adults from hatchery and wild origins. However, in this report, no distinction is made and all adult salmon arriving at their natal streams (not necessarily spawning) are referred to as the "escapement."

1.2 Need for Reliable Escapement Estimates

Reliable estimates of escapement are required to assess the impacts of management decisions and to evaluate the success (or failure) of conservation efforts. Escapement is the most direct measure of the health of a stock because it is only those fish in the escapement that have the opportunity to reproduce, contributing to future generations. Accurate estimates of escapement are required to establish sustainable harvest limits, forecast future run sizes, and to provide confidence in management strategies (Ames et al., 1974; Beidler and Nickelson, 1980; Perrin and Irvine, 1990; Bue et al., 1998; Manske and Schwarz, 2000).

An estimate of escapement needs to be both accurate and precise in order to be reliable (Paulik and Robson, 1969; Rose et al., 2000). Accuracy refers to the size of the bias of an estimate, $\hat{\theta}$. Bias is defined as the difference between the expected value, $E(\hat{\theta})$, of the estimate and the true value of the parameter (Cochran, 1977, page 13):

$$\text{Bias} = E(\hat{\theta}) - \theta. \tag{1.1}$$

If the difference between the expected (mean) value of the estimator and the true value is zero, then the estimate is said to be "unbiased." If the absolute difference is greater than zero, the estimator is said to be "biased." A distinction must be made between "statistically" biased and "experimentally" or "systematically" biased estimators (Paulik and Robson, 1969). A statistically biased estimator is one whose expected value deviates from the true value of the estimated parameter. An experimentally or systematically biased estimator is one whose value is affected by the sampling process itself. For example, a thermometer may be miss-calibrated to report all temperature measurements hotter than they really are by half a degree. Alternately, some assumption required by the experimental design is not met such that the estimates are biased because of the situation, but not because of the statistical properties of the estimator.

Precision refers to the size of the differences of each estimate from the expected (mean) value of all estimates that would be generated if the sampling procedure were repeated an infinite number of times (Cochran, 1977, page 16). Relative precision is defined as the probability of the estimate falling within $\epsilon \times 100\%$ of the true value, $(1-\alpha)100\%$ of the time:

$$Pr\left(|\frac{\hat{\theta}-\theta}{\theta}| < \epsilon\right) = 1 - \alpha, \tag{1.2}$$

where ϵ and α must be specified in advance. ϵ determines the size of precision. Usually, α is set to 0.05.

The most commonly used measure of sampling uncertainty is the variance of the estimator (Wolter, 1985, page 1). The variance, $V(\hat{\theta})$, of a random variable, in this case the estimator $\hat{\theta}$, is defined as the second central moment of the random variable (Casella and Berger, 2002, page 59):

$$V(\hat{\theta}) = E(\hat{\theta} - E(\hat{\theta}))^2. \tag{1.3}$$

The square root of the variance is the standard error, $SE(\hat{\theta}) = \sqrt{V(\hat{\theta})}$. The "coefficient of variation" (CV) is a measure of the relative size of the standard error to the estimate itself (Cochran, 1977, page 54):

$$CV(\hat{\theta}) = \frac{\sqrt{V(\hat{\theta})}}{\hat{\theta}} = \frac{SE(\hat{\theta})}{\hat{\theta}}. \tag{1.4}$$

The CV is a ratio of the "noise" to the "signal strength" of the estimate. In Equation (1.2) for example, ϵ is defined as the product of the CV and a quantile of the standard Gaussian distribution:

$$\epsilon = Z_{1-\frac{\alpha}{2}} CV(\theta). \tag{1.5}$$

A small CV constrains ϵ to be small so that the probability of the relative difference between the estimate and the true value of the estimator will be small.

The accuracy and precision of an estimate are not unrelated. Cochran (1977, section 1.8) showed how undetected bias can alter the true precision of an estimate depending on the direction and magnitude of the bias. As a general rule, the effect of the bias is considered negligible if the ratio of the absolute bias to standard error is less than 10%:

$$\frac{|\text{Bias}|}{\sigma} < 0.10.$$

The size of the precision does not similarly affect the accuracy (bias) of an estimate. However, some unbiased estimators may be so variable (imprecise) that they are useless for management decisions (Paulik and Robson, 1969; Rose et al., 2000).

1.3 Current Escapement Estimation Methods

Several methods of estimating salmon escapement are currently being used in the Pacific Northwest of the United States. Some methods, such as peak adult counts and spawning redd counts provide only an index of abundance. Other methods, like the area-under-the-curve method or mark-recapture experiments attempt to provide estimates of actual escapement. Still other methods, like dam passage counts and weir counts can provide near censuses of the returning adult population. Each agency concerned with salmon escapement uses their own suite of methods, so an accurate estimate of the proportions of use for each method is difficult to attain. StreamNet is a joint project between the Pacific Northwest's fish and wildlife agencies and tribes, administered by the Pacific States Marine Fisheries Commission, to "provide data and data services in support of the region's Fish and Wildlife Program and other efforts to manage and restore the region's aquatic resources" (Pacific States Marine Fisheries Commission, 2008). The StreamNet database contains escapement estimates from Washington, Oregon, Idaho, California, Alaska, and British Columbia, although the data from Alaska, California, and British Columbia are sparse. Salmon escapement estimates are classified as being generated from spawning grounds, redd counts, peak counts, or dam/wier counts, and are grouped together so that data in sequential years from the same sections of rivers are reported as "trends." Spawning ground trends may be generated from tag-recapture studies, carcass counts, or counts used for the area-under-the-curve method.

The results of a series of StreamNet queries for data from 1997-2007 show fifteen trends for pink salmon (*O. gorbuscha*), all of which were done on the spawning ground (Figure 1.1). Of 213 trends for chum salmon (*O. keta*), over half were also done on the spawning ground, although 81, were based on peak counts and 1 was a redd count (Figure 1.1). Coho salmon (*O. kisutch*) had the most trends, 4277, on record. For coho, 44.0% were on the spawning grounds, 2.8% were from redds, 52.4% were

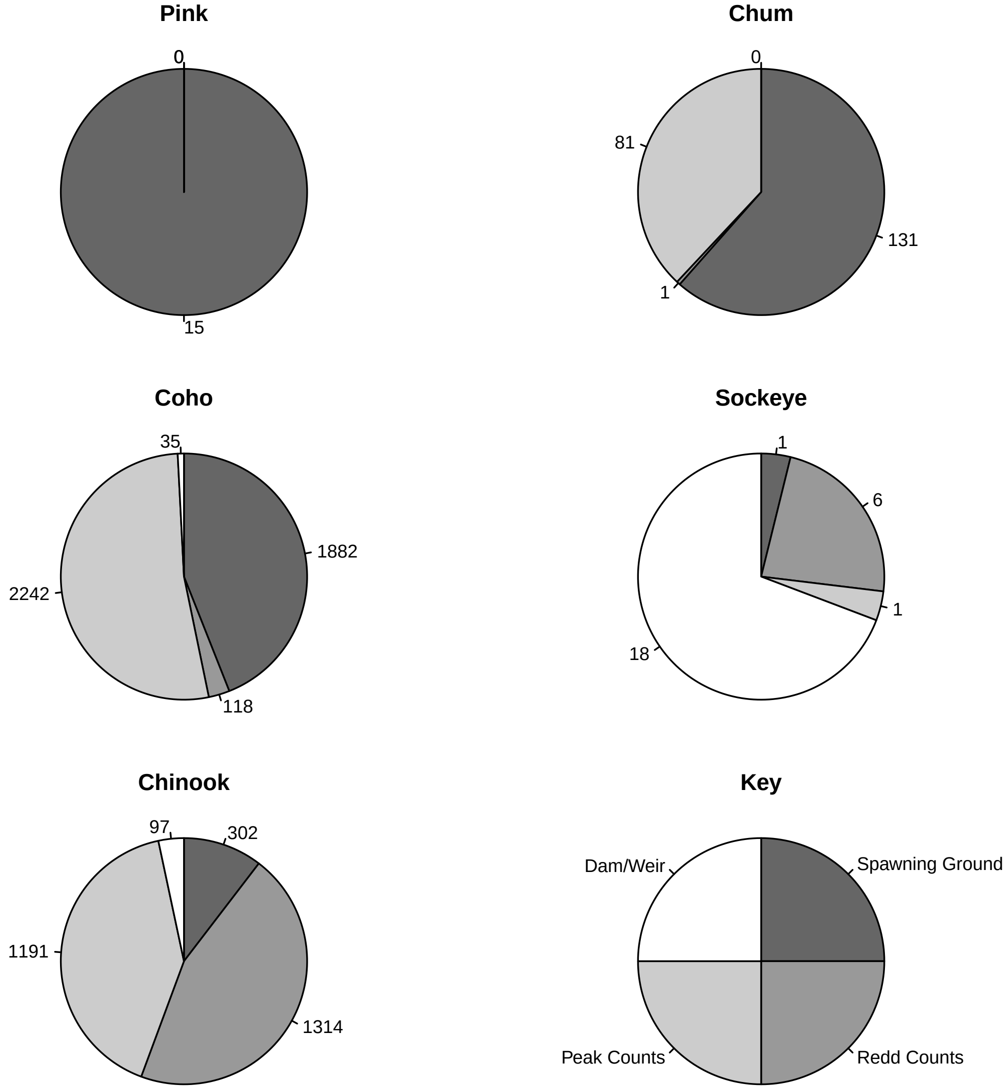

Figure 1.1: Results of StreamNet.org search for Washington, Oregon, Idaho, California, and Alaska salmon (*Onchorhynchus* spp.) escapement estimation efforts for 1997-2007. "Spawning Ground" includes mark-recapture, area-under-the-curve, or carcass count methods. The numbers adjacent to each wedge indicate the number of trends on record using the escapement estimation method represented by that wedge. A trend is a record of the escapement for a specific river section or tributary. The zero at the top of the Pink pie indicates that there were no redd, peak, or dam/weir counts. The zero at the top of the Chum pie shows that no chum escapements were estimated via dam/weir counts.

from peak counts, and the remaining 0.8% were dam or weir counts (Figure 1.1). By contrast, only 3.8% of sockeye (*O. nerka*) trends were from the spawning grounds and most of the trends were from dams or weirs (69.2%) (Figure 1.1). For sockeye, 23.1% of trends from 1997-2007 were from redds and only 3.8% were from peak counts (Figure 1.1). Chinook salmon (*O. tshawytscha*) had the second largest number of trends with 2904. Of these, 10.4% were spawning ground counts, 45.2% were from redd counts, 41.0% were from peak counts, and 3.3% were from dams or weirs (Figure 1.1).

Knudsen (2000) identified 9,430 populations of Pacific salmon in the United States. Of the identified populations, Knudsen (2000) rated the means by which 3,567 of the populations' escapements were estimated. Spawning ground methods including what Knudsen (2000) called "Good", "Fair", "Poor" and "Carcass" indices accounted for 2.5% of the escapement methods (Figure 1.2). The indices were based on area-under-the-curve and carcass count methods, but did not include tag-recapture studies. Tag-recapture studies were included with dam, weir, sonar, and tower counts, which made up 25.1% of the methods (Figure 1.2). Peak count methods were the most commonly used method of estimating escapement and made up 70.9% (2530 of 3567) of the estimates (Figure 1.2). The remaining 1.5% of the populations were estimated using redd count methods (Figure 1.2).

1.4 Escapement Estimation Methods Reviews

In the past there have been two major reviews of methods used to estimate salmon escapement in western North America. Cousens et al. (1982) summarized and reviewed the escapement estimation methods employed by the U. S. and Canadian governments in the 1980's. For each method, Cousens et al. (1982) described the necessary procedures, the locations in the Pacific Northwest of its use, what managers considered the best and worst situations in which to use it, and how accurate

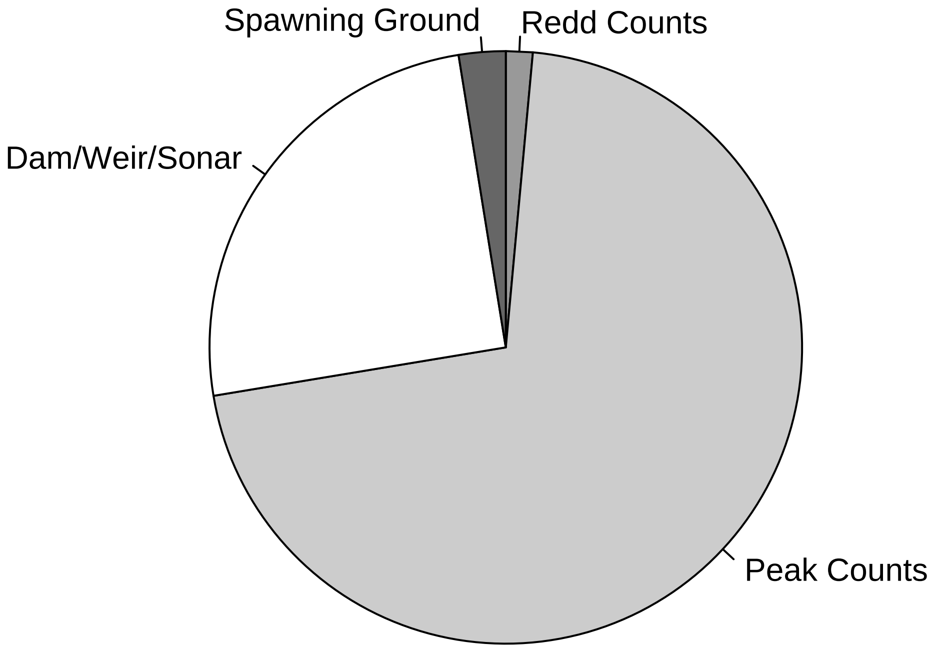

Figure 1.2: Number of United States populations of salmon (*Oncorhynchus* spp.) for which each category of escapement estimation is used. Taken from Table 17.3 of Knudsen (2000), where “Total redds” and “Redd surveys” are combined in Redd Counts (52), “Peak count”, “One count”, and “One count–sporadic” categories are combined in Peak Counts (2530),“Total” and “Total estimate” categories are combined in Dam/Weir/Sonar (894), and “Good”, “Fair”, “Poor” and “Carcass” indices are combined in Spawning Ground (91).

it was compared to other methods. Cousens et al. (1982) did not review the Area-Under-the-Curve Method, which has become widely used in both Canada and the United States (Beidler and Nickelson, 1980; Ames, 1984; Irvine et al., 1992). Cousens et al. (1982) only provided mathematical derivations for mark-recapture methods. Analyses of statistical bias were not included, nor were explicit formulas for variance estimation.

Schubert (1998) reviewed the methods used by the Canadian Department of Fisheries and Oceans to estimate salmon escapement to the Fraser River. He described arrival indices, mark-recapture methods, enumeration fences, and peak count methods, offered suggestions for improvement of operational techniques, and compared the empirical accuracy of different methods. A mathematical expression was provided for only the peak count method. No analysis of estimator bias nor explicit variance formulas were offered. No discussion of the area-under-the-curve method or the use of sonar was included.

There have been numerous reviews comparing empirical estimates of escapement for a single system using two or multiple techniques. Johnston et al. (1986) compared "the reliability, accuracy, precision, and cost of fence counts, mark recapture techniques and a Bendix counter" (a sonar instrument) for pink, chum, and coho salmon in the Keogh River, B.C., Canada. Irvine et al. (1992) compared two different stream sampling schemes used in the area-under-the-curve method to mark-recapture estimates and fence counts in two streams on Vancouver Island, Canada. In Japan, Miyakoshi et al. (2003) compared escapement estimates of masu salmon (*Onchorhynchus masou*) based on carcass counts to those based on tag-recovery techniques. Schubert (2000) did an extensive study on the Stellako River, British Columbia, comparing different mark-recapture estimates with a fence count. He included an empirical comparison of methods and a thorough evaluation and discussion of assumptions, but did not include rigorous statistical analyses of the different estimators. The results of empirical comparisons between methods are specific to the systems in which they were conducted and are judged relative to each other, rather than using statistical theory.

There have been a handful of reviews examining a single technique applied across multiple locations. Perrin and Irvine (1990) compiled stream life time estimates and techniques from salmon fishery managers. They compared actual estimates and empirical performance of methods but did not analyze the statistical properties of the estimators. Hilborn et al. (1999) used data on pink salmon from 18 streams in Prince William Sound, Alaska to compare area-under-the-curve estimates of escapement based on parametric spawner curves with those based on the traditional trapezoidal method. The statistical properties of the parametric spawner curves were made explicit, though they were less accurate in practice than the traditional method. As with comparisons between methods, most empirical comparisons within methods are system-specific and not based on statistical theory.

1.5 Purpose and Scope of Report

The purpose of this report is to offer a statistically-based review of methods currently employed to estimate salmon escapement in the Pacific Northwest. We conducted a StreamNet study and an extensive literature search to find techniques used in Washington, Oregon, California, Alaska, and British Columbia between 1997 and 2007. We chose methods to review based on 1) StreamNet queries, 2) recent salmon management book reviews, 3) more than 280 peer-reviewed journal articles and management agency publications, and 4) personal communication with the Washington Department of Fish and Wildlife (WDFW). We reviewed the Area-Under-the-Curve Methods (Chapters 2 and 3), Mark-Recapture Methods (Chapter 4), Passage Count Methods (Chapter 5), Peak Count Methods (Chapter 6), Redd Count Methods (Chapter 7), and Carcass Count Methods (Chapter 8). For each method reviewed we determined the statistical bias and variance of each estimator, and described bias adjustments that are commonly made. We determined the effects of assumption violations on the estimators and found empirical evidence regarding the relative ease or difficulty of ensuring assumptions were met.

1.6 Report Methods

Each reviewed escapement method, with the exception of the area-under-the-curve method, was presented in its own chapter. The area-under-the-curve method has been divided into two chapters: one for fish-days and one for stream life. In each chapter, we gave a brief introduction, listed the notation used throughout the chapter, and listed the assumptions required to support each method. For each variation or adjustments we gave a general description and then an explicit mathematical derivation of the escapement estimator. We evaluate the direction and magnitude of any statistical bias in each estimator and rate the bias of the estimator under ideal circumstances. This included adjustments to make the estimator less statistically biased, but did not take violations of assumptions into account. The possible ratings for the estimator were "Poor," "Fair," and "Good." The estimator was deemed "Poor" if it was inherently biased even with adjustments. The estimator was deemed "Fair" if it was statistically unbiased to the first order Taylor series expansion, but has been shown to be somewhat unreliable in empirical studies. The estimator was "Good" if it was statistically unbiased to the first order Taylor series expansion and performs well empirically.

For each estimator, we determined if a variance estimator exists, and if not, derived an estimator for the variance, where possible. We rated each method's variance as "None," "Poor," "Good," or "N/A." A "None" indicates that either there was no variance estimate provided or possible. The variance was deemed "Poor" if it failed to incorporate all sources of error or was poorly derived. The variance was deemed "Good" if all sources of error were incorporated correctly into the estimator. When an escapement estimate was based on an exact count, the sampling variance was zero and a variance estimator was no longer applicable, “N/A.”

We referred to pertinent assumptions as they became relevant during the explanation of each estimator and described how violations of the assumptions affected the systematic bias. We rated each of the estimators based on the ease of ensuring the required assumptions. The possible ratings for the assumptions were “Reasonable,” “Unreasonable,” “Conflictual,” and “Unknown.” A “Reasonable” rating implied that the assumptions could generally be ensured in well-designed studies. An “Unreasonable” rating indicates that the assumptions were often biologically or logistically unreasonable. A method with “Conflictual” assumptions was one for which fulfilling one assumption precluded either the fulfillment of another or the rationale for using that particular method. For example, the area-under-the-curve method requires an estimate of stream life to estimate escapement, and some estimates of stream life require an estimate of escapement. Only the Redd Area Method received an “Unknown” assumption rating, because the rationale behind the assignment of redd area to number of females was poorly described and supported.

We rated each estimator’s performance with respect to empirical studies. The comments under the “Empirical” column are case specific and should be self explanatory, but three common entries merit explanation. “Inapplicable” implies that the method does not actually measure the desired quantity, such as stream life. “Index only” is assigned to methods that do not provide an actual estimate of abundance, only an index to proportional abundance. “Not much testing” indicates that a method is uncommon or relatively new and has not been subject to many performance tests in the field.

We gave each estimator a “recommendation” based on the method’s bias, variance, assumptions, and empirical performance. The recommendations were either “Do Not Use,” “Cautionary,” “Good Alternative,” and “Best Choice.” A method was given a “Do Not Use” if the estimator was inherently biased or failed to estimate it’s desired end. A method was given a “Cautionary” if the estimator was statistically unbiased but the empirical evidence suggests that either the experimental bias or the variance is so large that the estimator is rendered practically useless. A method was given a "Good Alternative" if the estimator was theoretically unbiased, but the assumptions required are difficult to attain. A method was only recommended as a "Best Choice" if both the estimator and variance were unbiased and the supporting assumptions were reasonably attainable. It is possible that an escapement category of methods may have multiple or zero "Best Choice"s.

We included a step-by-step example for all "Best Choice" techniques and for two others. Data for the examples was taken from the literature where possible, although some data has been assumed.

The results of my evaluation of each method are summarized in a table at the end of each chapter, along with a brief written summary explaining the recommendation for each method. We conclude the review with a discussion of five issues managers need to take into consideration as they select an escapement method, and six suggestions for future research that could benefit salmon escapement estimation techniques.

Chapter 2

AREA UNDER THE CURVE: FISH DAYS

2.1 Introduction to AUC

The idea supporting the area-under-the-curve (AUC) method is that every fish escaping to the survey area resides there for a certain number of days, i.e. every fish has a certain stream life. If one were to sum the individual stream lives of all the fish and divide the total by the average stream life, one would know the number of fish that had escaped to that stream. In practice, individual stream lives are rarely known outside an artificial spawning channel (e.g. Schroeder, 1973). More frequently, the total number of fish-days is estimated as the number of fish present each day summed over all the days of the run, though reports using daily counts of live fish are rare. Most often the fish are counted periodically, either by foot surveys (Perrin and Irvine, 1990; Irvine et al., 1992; Lady and Skalski, 1998), boat or snorkel surveys (Cousens et al., 1982; Bocking et al., 1988), or aerial surveys (Hill, 1997; Bue et al., 1998; Neilson and Geen, 1981). The total number of fish-days is extrapolated from periodic counts.

The estimate of total fish-days is divided by an estimate of average stream life to obtain the estimate of escapement:

$$\text{Escapement} = \frac{\text{Estimated Total Fish-days}}{\text{Average Residence Time}} \tag{2.1}$$

2.2 *Fish-days Notation*

We will use the following notation throughout the Fish-days chapter. Estimates are denoted with carets, for example, escapement, N, is estimated by $\hat{N}$.

$Cov(x, y)$	The covariance of random variables x and y
$CV(x)$	Coefficient of variation of random variable x
D_z	Number of days in time stratum z
d_z	Number of sampled days in time stratum z
$E_x(Y)$	Expected value of variable Y, taken with respect to x
F	Total fish-days, area under the curve
H	Number of strata in a survey area, indexed by h
K	Number of days in the run
k	Number of survey occasions,($k \leq K$)
κ_i	Number of possible systematic survey start days i.e. number of days between surveys
L_i	Actual number of live individuals at the time of the ith survey
l_i	Observed number of live individuals at the time of the ith survey
N	Escapement, total number of individuals in a run
P_i	Observer efficiency at the time of the ith survey, i.e. the probability of an observer seeing a fish given that a fish is there
R_h	Number of reaches (250 m) in stratum h of a stream
r_h	Number of sampled reaches in stratum h
T	Average stream life, average residence time
t_i	Time of the ith survey
$V_x(Y)$	The variance of variable Y, taken with respect to x
X_i	Estimate of live fish in entire study area at time of ith survey
Y_i	Total number of fish-days represented by counts from the ith survey
Z	Number of time strata in the run, indexed by z

2.3 AUC Fish-days Assumptions

The following assumptions are required in one or more of the Fish-days Estimator Methods reviewed below. Assumptions (A-1) and (A-2) are required by all methods of estimating total fish-days. Possible assumptions include the following:

(A-1). The probability of detection is known or estimable without bias.

(A-2). Variance of the detection probability estimator is estimable.

(A-3). Probability of detection is the same for all fish present during a given survey.

(A-4). Live fish are counted independently of the method of estimating the detection probability.

(A-5). The entire stream is surveyed.

(A-6). Strata constitute a mutually exclusive and exhaustive partition of the study area.

(A-7). Strata are homogeneous.

(A-8). Sampled reaches or days within the strata are selected randomly.

(A-9). All surveys are instantaneous with respect to the arrival and death of individual fish.

(A-10). Once in the survey, individuals do not leave the study site until death.

(A-11). Arrivals and deaths are distributed uniformly across a survey interval.

(A-12). The count for the first survey is zero.

(A-13). The count for the last survey is zero.

(A-14). Stream is surveyed every day of the run.

(A-15). The k survey dates are a random sample of the K total run days.

(A-16). Sampled reaches within the strata are selected based on ease of access or suspicion of salmon habitat.

(A-17). Sampled reaches are selected at random without stratification.

(A-18). Stream is surveyed every κ days.

(A-19). The underlying pattern of fish abundance is independent of survey times, i.e. there is no periodicity in the daily abundance of salmon if surveys are systematic.

(A-20). There exists an unbiased estimate of stream life.

(A-21). Arrival times are distributed according to a Gaussian distribution.

(A-22). The probability of detection is 100%.

(A-23). Observer efficiency remains constant over the season.

(A-24). Observer efficiency is known without error.

(A-25). The variance of the stream life estimate is estimable.

(A-26). Estimates of fish-days and stream life are independent.

The above assumptions will be discussed as they arise in the development of the estimators.

2.4 Observer Efficiency Estimates

Counts of live fish often require an estimate of observer efficiency (i.e. probability of detection) to account for the fish present that were not detected. Estimates are made by assuming 100% observer efficiency (Parken et al., 2003; Seiler et al., 2004), by subjective estimates of visibility (Ames, 1984; Bocking et al., 1988), by follow-up electrofishing (Bocking et al., 1988; Irvine et al., 1992), by removal estimates (Solazzi et al., 1983), or by comparisons between count estimates of different methods (Bocking et al., 1988; Irvine et al., 1993; Bue et al., 1998; Jones III et al., 1998; Visser et al., 2002). The details of observer efficiency estimates are beyond the scope of this report. For the remainder of the report, any inclusion of "observer efficiency" or "detection probability" will assume that there is an unbiased estimate available (A-1) which includes an unbiased estimate of its variance (A-2).

2.5 Binomial Sighting Process

For every live individual in the survey area on a given survey occasion, surveyors will either see it, or they will not. If the probability of detection is the same for all fish (A-3) then the number of fish detected is a realization from a binomial trial (Liao, 1994). The number of live fish seen by surveyors, l_i, is a random variable dependent on the number of live fish in the stream, L_i, and the probability of the surveyors seeing a fish at the ith survey occasion, such that:

$$l_i \sim \text{Binomial}(L_i, P_i).$$

The expected value and variance of the number of individuals detected on the ith survey are:

$$E(l_i) = L_i P_i \qquad \text{and} \qquad V(l_i) = L_i P_i (1 - P_i)$$

The number of life fish present can be estimated from the number of live fish seen,

l_i, and an estimate of the observer efficiency (A-1):

$$\hat{L}_i = \frac{l_i}{\hat{P}_i}. \tag{2.2}$$

Equation (2.2) is unbiased only to the first and second orders Taylor series expansion. The expected value of Equation (2.2) is:

$$\begin{aligned}
E(L_i) &= E(\tfrac{l_i}{\hat{P}_i}) \\
&= \underbrace{\frac{E(l_i)}{E(\hat{P}_i)}}_{\text{First}} - \underbrace{\frac{E(l_i)}{E(P_i^2)} E\left(\frac{l_i}{P_i} - \frac{l_i}{P_i}\right)}_{\text{Second}\quad\text{order}} + \underbrace{\frac{E(l_i)}{E(P_i^3)} E\left(\left(\frac{l_i}{P_i} - \frac{l_i}{P_i}\right)^2\right)}_{\text{Third}\quad\text{order}} + \underbrace{\begin{matrix}\text{Negligible}\\ \text{Remainder}\end{matrix}}_{\text{Fourth+}} \\
&= \tfrac{L_i P_i}{P_i} - 0 + \left(\tfrac{L_i P_i}{P_i^3}\right) V(P_i) + \text{Negligible} \quad \text{Remainder} \\
&= L_i + \left(\tfrac{L_i}{P_i^2}\right) V(P_i) \\
&= L_i(1 + CV^2(P_i))
\end{aligned} \tag{2.3}$$

From Equation (2.3) it is clear that Equation (2.2) is positively biased unless the observer efficiency is known without error. It is also clear that the smaller the $CV(P_i)$, the smaller the bias in the estimate of fish present. Generally, Equation (2.2) is considered first order unbiased for biological purposes.

The variance of the live fish estimate given in Equation (2.2) is derived by conditioning on the two-stage sampling process:

$$\begin{aligned}
V(\hat{L}_i) &= V\left(\tfrac{l_i}{\hat{P}_i}\right) \\
&= E_2(V_1\left(\tfrac{l_i}{\hat{P}_i}|2\right)) + V_2(E_1\left(\tfrac{l_i}{\hat{P}_i}|2\right)) \\
&= E_2(\tfrac{(l_i)^2 V(\hat{P}_i)}{P^4}) + V_2\left(\tfrac{l_i}{P_i}\right) \\
&= \tfrac{V(\hat{P}_i)}{P^4}[V(l_i) + E^2(l_i)] + \tfrac{1}{(P_i)^2} V_2(l_i) \\
&= \tfrac{V(\hat{P}_i)}{P^4}[L_i P_i(1 - P_i) + L_i^2 P_i^2] + \tfrac{L_i P_i(1-P_i)}{(P_i)^2} \\
&= CV^2(\hat{P}_i)[\tfrac{L_i(1-P_i)}{(P_i)} + L_i^2] + \tfrac{L_i(1-P_i)}{(P_i)}
\end{aligned} \tag{2.4}$$

Equation (2.4) is estimated by:

$$\hat{V}(\hat{L}_i) = \frac{l_i(1 - \hat{P}_i)}{(\hat{P}_i)^2} + \hat{CV}^2(\hat{P}_i)\left[\frac{l_i(1 - \hat{P}_i)}{(\hat{P}_i)^2} + \left(\frac{l_i}{\hat{P}_i}\right)\right], \tag{2.5}$$

which is first order unbiased for the variance of Equation (2.2).

2.6 Variability of Fish-days Estimators

There are four sources of variability contributing to estimates of total fish-days. The first source of variability comes from estimating observer efficiency by $\hat{P}$, i.e. $V(\hat{P}|P)$ (A-1). The second source of variability is associated with the binomial sighting process and depends on observer efficiency, P, as shown in Equation (2.4). The third source of variability is associated with the selection of stream sections (reaches) to be surveyed. Frequently, only parts of the study streams are surveyed (Perrin and Irvine, 1990; Irvine et al., 1992) and the counts are expanded to represent the entire study area. The fourth source of variability is associated with the selection of days on which to perform to surveys. A few studies have performed surveys on a daily basis (Killick, 1955; Crone and Bond, 1974), but more often surveys are conducted every few days or one to two weeks apart (Hill, 1997; Bue et al., 1998). In order to derive an accurate estimate of the variance of fish-days, the variances from all four sources must be included. The general form of the estimate of variance for fish-days is:

$$V(\hat{F}) = \underbrace{E_4E_3E_2V_1(\hat{F})}_{\text{Source 1}} + \underbrace{E_4E_3V_2E_1(\hat{F})}_{\text{Source 2}} + \underbrace{E_4V_3E_2E_1(\hat{F})}_{\text{Source 3}} + \underbrace{V_4E_3E_2E_1(\hat{F})}_{\text{Source 4}} \qquad (2.6)$$

where E_i is the expected value and V_i is the variance of the estimator taken with respect to variability source i (1 = observer efficiency, 2 = sighting, 3 = spatial, 4 = temporal). The variance contributed to the fish-days estimator by estimating observer efficiency is Source 1. Source 2 comes from the binomial sighting process described in Section 2.5. Sources 3 and 4 come from the spatial and temporal variability, respectively.

No published technique includes a variance estimator incorporating all four sources of variance. The Ames (1984) method, described in section 2.2.3, does not allow for a variance estimate of any kind because it ignores the variance of the estimates of observer efficiency, treats the survey dates as fixed, and uses a subjective technique to draw the spawner abundance curve. Studies have estimated observer efficiency (Irvine et al., 1992; Liao, 1994), but only Liao (1994) incorporated the variance of the observer efficiency estimate (Source 1) into the overall variance. Liao (1994) is also the only study to derive a variance estimate for Stage 2 (Source 2) based on the binomial sighting process. Authors using the trapezoidal method English et al. (1992); Irvine et al. (1992, 1993); Parken et al. (2003), describe in Section 2.7.2, estimate the spatial variability (Source 3), but ignore the temporal variability (Source 4). The Liao (1994) method (Section 2.7.4) includes the sighting and temporal variability (Sources 2 and 4), but assumes that the entire study stream is surveyed (A-5) such that spatial (Source 3) variability is zero.

2.7 Total Fish-days Estimation Methods

The area-under-the-curve method requires two pieces of information to estimate escapement: total fish-days and average stream residence time. Total fish-days can be derived two ways: as the sum, over all fish, of the number of days each fish is alive in the spawning area, $F = \sum_{j=1}^{N} T_j$, or as the number of fish alive in the study area each day summed over all days, $F = \sum_{i=1}^{K} L_i$. In practice, the stream life of even a few individual spawners is rarely known, let alone the entire population. And if we knew how many salmon to sum over, we would not need to estimate fish-days. Because total fish days is also the number of fish alive in the study area each day summed over all days, it can be estimated by making periodic counts of the salmon in the study area. The raw periodic counts can be adjusted by estimates of observer efficiency, $\hat{P}_i$, to yield the corrected counts, $\hat{L}_i = \frac{l_i}{\hat{P}_i}$, estimating the true number of adult salmon alive in the stream, L_i, on survey occasion i. A curve is fitted to the corrected periodic counts, which Ames (1984) called the “spawner abundance curve.” The area under the spawner abundance curve represents the total fish-days in the run, i.e. the number of fish that there would be in the run if every fish lived exactly one day (Neilson and Geen, 1981). Most salmon live in the stream longer than one day, so an estimate of stream life is required to complete the area-under-the-curve method. The techniques and assumptions for estimating stream life are discussed in Chapter 3.

2.7.1 The Ames (1984) Method

The earliest peer-reviewed publication describing a "spawner abundance curve" was authored by William J. McNeil (1964*a*). He was interested in how the density of female pink salmon in two streams in Alaska affected egg deposition and recruitment. Since McNeil (1964*b*) was only interested in the number of females, he only counted females, fit a curve to the points by hand, and calculated "female days" as the area under the fitted curve. He also used female redd residence time instead of average stream life. McNeil's (1964*a*, *b*) methods pertained only to the number of female spawners. They can only be used to make inference on the males if there is an known or estimated male-to-female ratio.

Crone and Bond (1974) extended McNeil's (1964, 1966) method by counting both males and females, but applied the residence time of females to both counts. Ames (1984) is credited with being the first to document the use of periodic counts of both male and female salmon in the entire study stream and to use an average stream life rather than redd residence time. For the Ames Method, counts are made of both males and females, along with subjective estimates of visibility for the days surveyed. Both the actual counts and counts adjusted for visibility are plotted on graph paper, and a spawner abundance curve is drawn by hand between the two, as shown in Figure 2.1. The area beneath the curve is measured in square inches using a compensating polar planimeter and multiplied by the number of fish-days per square inch.

Because the fish-days estimate is based on heuristic estimates of observer efficiency and a subjective curve is imposed on the data, the accuracy and precision of the

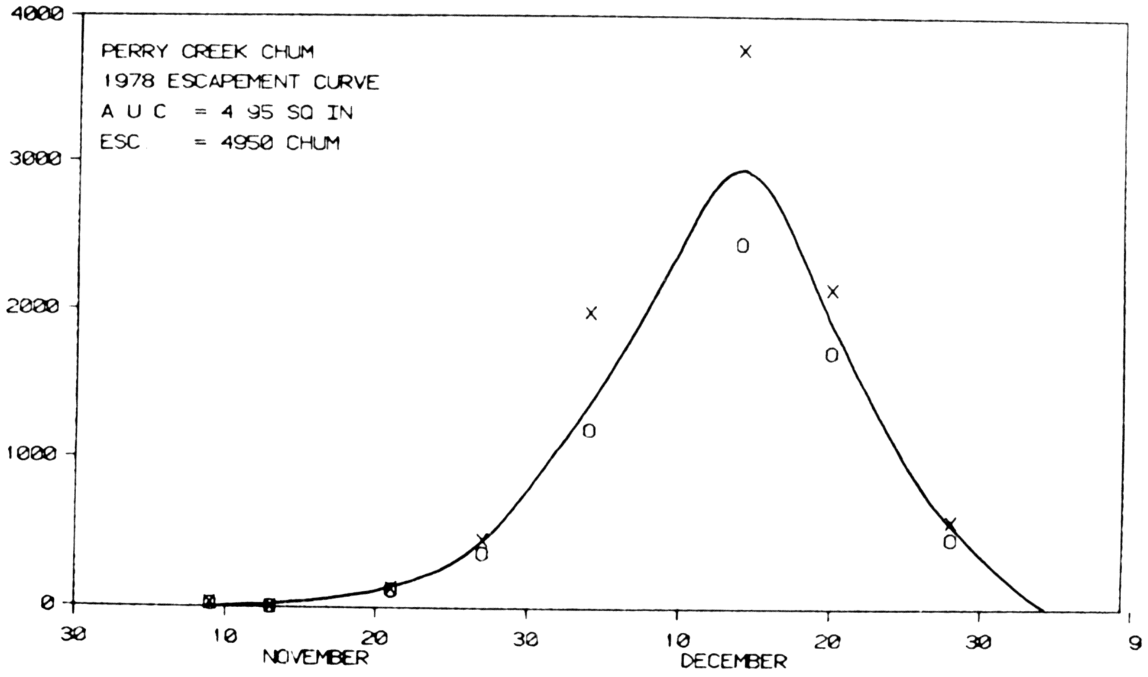

Figure 2.1: The Ames Method of estimating fish-days. Actual counts of Perry Creek chum salmon are shown with open circles (o). Expanded counts are shown with x's (x) and are based on subjective estimates of observer efficiency. The fish-days curve is drawn by hand between the two types of counts. Image taken from Figure 3 of Ames (1984).

estimate are unknown. The bias could be in either direction and there is no variance estimator for this approach.

2.7.2 The Trapezoidal Method (Irvine et al., 1992)

Irvine et al. (1992) estimated the escapement of coho salmon in French Creek and Black Creek on Vancouver Island, British Columbia. Rather than survey the entirety of both creeks, they partitioned the creeks into exhaustive and mutually exclusive strata (A-6) based on tributary confluences and stream gradients as shown in Figure 2.2. Two or three 250-meter reaches (river sections) were sampled from each stratum. The length of 250-meters was chosen as the optimal survey length based on numerical simulations of fish density and distribution patterns Bocking et al. (1988).

They adjusted the counts from each reach by observer efficiency (A-1), then expanded the reach counts to estimate the total number of live fish in each strata. The counts from each strata were summed to estimate the total number of live individuals present in the study area for each survey occasion:

$$\hat{L_i} = \sum_{h=1}^{H} \frac{R_h}{r_h} \sum_{j=1}^{r_h} \frac{l_{ihj}}{\hat{P_{ihj}}} \tag{2.7}$$

where H is the number of strata, R_h is the number of 250-m reaches in strata h, r_h is the number of reaches in strata h that were sampled on survey occasion i. Equation (2.7) is unbiased for the total number of individuals alive in the study area on survey occasion i as long as the 250-meter reaches are selected randomly (A-8) (Cochran, 1977, page 91), surveys are conducted instantaneously with regard to arrivals and deaths of individual fish (A-9), and fish do not leave the study area once they have entered (A-10).

Irvine et al. (1992) standardized the curve in the AUC method by connecting the periodic counts from Equation (2.7) with straight lines (Figure 2.3). This assumes that arrivals and deaths happen uniformly between surveys (A-11).

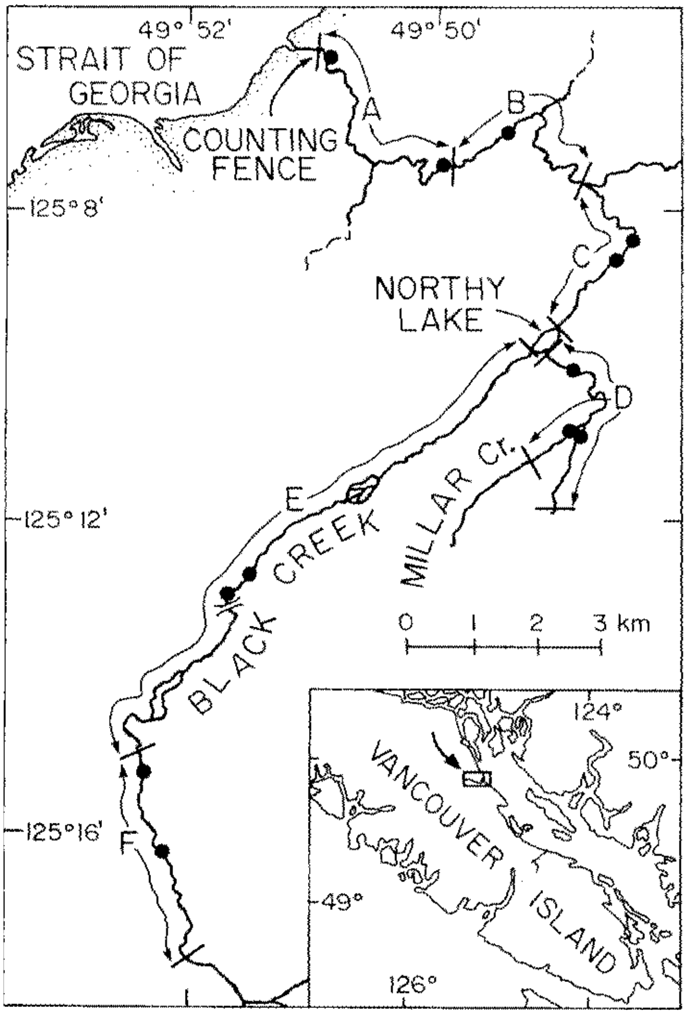

Figure 2.2: A map of Black Creek on Vancouver Island showing six strata (A-F) and examples of 250-meter sampling reaches. Taken from Figure 1 of Irvine et al. (1992).

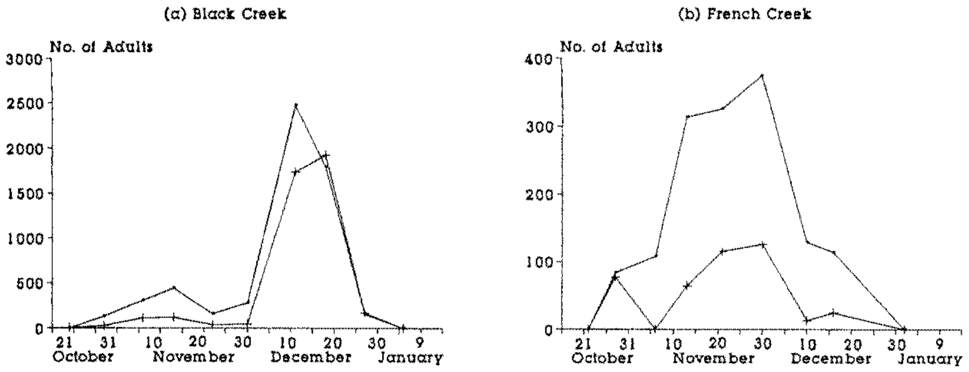

Figure 2.3: Linear interpolation between estimated numbers of adult coho in Black and French Creek during 1989 for the Trapezoidal Method. Two different spatial sub-sampling schemes were used: stratified random sampling (+) and stratified index sampling (●). Taken from Figure 4 of Irvine et al. (1992)

The resulting shape is a polygon and why this is sometimes referred to as the "polygon method". The area underneath this "curve" is calculated by partitioning the polygon into trapezoids and summing the area of each trapezoid:

$$\hat{F} = \frac{1}{2}\sum_{i=1}^{k-1}(t_{i+1} - t_i)(\widehat{L_{i+1}} + \hat{L}_i). \tag{2.8}$$

With the constraint that $\hat{L_1} = \hat{L_k} = 0$ (A-12) and (A-13), the trapezoidal estimator can be re-written as (Appendix A):

$$\hat{F} = \frac{1}{2}\sum_{i=2}^{k-1}(t_{i+1} - t_{i-1})\hat{L}_i. \tag{2.9}$$

The first and last counts can be assumed to be zero (A-12) and (A-13) if surveys commence prior to the onset of the run and continue until no more live fish are detected. If the first or last count is not zero, then Equation (2.9) will be negatively biased for fish-days.

The expected value of the estimator, $\hat{F}$, as described in Equation 2.9, is (Liao, 1994):

$$\begin{aligned} E(\hat{F}) &= E(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}(t_{i+1}-t_{i-1})\hat{L}_i) \\ &= \tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}(t_{i+1}-t_{i-1})E(\hat{L}_i) \\ &= \tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}(t_{i+1}-t_{i-1})L_i \\ &= \textstyle\sum_{i=2}^{k-1}L_i \quad \text{iff } (t_{i+1}-t_{i-1})=2 \quad \text{for all } i \text{ and} \\ &= \textstyle\sum_{i=1}^{k}L_i \quad \text{iff } L_0, L_k = 0 \\ &= \textstyle\sum_{i=1}^{K}L_i \quad \text{iff } k=K. \end{aligned}$$

Therefore, the trapezoidal estimator is unbiased only if $k = K$ and the spawning area is surveyed every day of the run (A-14).

Irvine et al. (1992) divided the survey streams into strata based on river confluences and stream gradient (A-6). The variance of the fish-days estimator depends on the sampling scheme used to select the surveyed reaches of each strata. Irvine et al. (1992) selected sampling reaches using both a stratified random design (STR) and a stratified index design (SIS). For the stratified index design, reaches were purposefully chosen based on ease of access or the suspected presence of spawning habitat (A-16). No valid inference can be drawn from the estimator under the SIS design because the reaches were not selected randomly. Nor is there a valid variance estimate for the SIS estimator. For the stratified random design, subunits (250 meter reaches) of each strata were randomly selected to be surveyed (A-8). The variance of the periodic counts using STR is derived from Equation (5.6) of Cochran (1977, page 92). It includes the variance between reaches within each stratum for each survey occasion, which is given by:

$$S_h^2 = \frac{1}{R_h - 1}\sum_{j=1}^{R_h}(L_{jh} - \bar{L_h})^2 \qquad \text{where} \quad \bar{L_h} = \frac{1}{R_h}\sum_{h=1}^{R_h} L_h \tag{2.10}$$

and is estimated by:

$$s_h^2 = \frac{1}{r_h - 1}\sum_{j=1}^{r_h}(\hat{L}_{jh} - \bar{\hat{L}}_h)^2 \qquad \text{where} \quad \bar{\hat{L}}_h = \frac{1}{r_h}\sum_{h=1}^{r_h} \frac{l_h}{\hat{P}_h}, \tag{2.11}$$

where R_h is the number of reaches in stratum h, r_h is the number of surveyed reaches in stratum h, and $\hat{L}_{jh}$ is the estimated number of fish present in the jth sampled reach of stratum h. Note we have momentarily dropped the survey date subscript i to aid readability. The variance for the estimated number of live individuals between strata on survey occasion i is given by Equation 5.10 in Cochran (1977, page 93) as:

$$V(\hat{L}_i) = \sum_{h=1}^{H} R_h^2 \left(1 - \frac{r_h}{R_h}\right) \frac{S_h^2}{R_h} \tag{2.12}$$

and is estimated by

$$\hat{V}(\hat{L}_i) = \sum_{h=1}^{H} R_h^2 \left(1 - \frac{r_h}{R_h}\right) \frac{s_h^2}{r_h}. \tag{2.13}$$

The variance for the STR trapezoidal fish-days estimate given by Irvine et al. (1992) is:

$$V(\hat{F}) = 0.25 \sum_{i=2}^{k-1} V(\hat{L}_i)(t_{i+1} - t_{i-1})^2 \tag{2.14}$$

where $V(\hat{L}_i)$ is given by Equation 2.12. It underestimates the true variance because it does not include the variability associated with estimating observer efficiency, the sighting process, or the selection of survey dates.

Using the four-stage variance scheme described in Section 2.6, the variance from the Irvine et al. (1992) version of the trapezoidal estimator should be (Appendix B):

$$V(\hat{F}) = \frac{k}{4K} \sum_{i=2}^{K-1} \left[(t_{i+1} - t_{i-1})^2 \sum_{h=1}^{H} \frac{R_{hi}}{r_{hi}} \sum_{j=1}^{R_{hi}} \left\{ \frac{V(\hat{P}_i)}{P_i^4} [L_i P_i (1 - P_i) + (L_i P_i)^2] \right\} \right] \tag{2.15}$$

$$+\frac{k}{4K} \sum_{i=2}^{K-1} \left[(t_{i+1} - t_{i-1})^2 \sum_{h=1}^{H} \frac{R_{hi}}{r_{hi}} \sum_{j=1}^{R_{hi}} \left\{ \frac{L_i (1 - P_i)}{P_i} \right\} \right] \tag{2.16}$$

$$+E_4 \left[\frac{1}{4} \sum_{i=2}^{k-1} (t_{i+1} - t_{i-1})^2 V_3 \left(\underbrace{\sum_{h=1}^{H} \frac{R_{hi}}{r_{hi}} \sum_{j=1}^{r_{hi}} L_{ihj}}_{X_i} \right) \right] \tag{2.17}$$

$$+\frac{1}{4}V_4\left(\underbrace{\sum_{i=2}^{k-1}\left\{(t_{i+1}-t_{i-1})\sum_{h=1}^{H}\sum_{j=1}^{R_{hi}}L_{ihj}\right\}}_{Y_i}\right) \tag{2.18}$$

where K, k, H, h, and t_i are all known constants.

The $V_3(X_i)$ part of Line (2.17) depends on the sampling scheme used to select the reaches in the stream strata. If the reaches are chosen at random (A-8), then the variance of the quantity in the parentheses, $V_3(X_i)$, is (Cochran, 1977, page 93):

$$\sum_{h=1}^{H}R_h(R_h-R_h)\frac{S_h^2}{r_h}=\sum_{h=1}^{H}R_h^2\left(1-\frac{r_h}{R_h}\right)\frac{S_h^2}{r_h},$$

where S_h^2 is given by Equation (2.10). Therefore, Line (2.17) becomes

$$+\frac{k}{4K}\sum_{i=2}^{K-1}\left[(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}R_h^2\left(1-\frac{r_h}{R_h}\right)\frac{S_h^2}{r_h}\right]. \tag{2.19}$$

If the sampled reaches are chosen completely at random without stratification (A-17), then H=1 and Line (2.17) becomes (Cochran, 1977, page 24):

$$+\frac{k}{4K}\sum_{i=2}^{K-1}\left[(t_{i+1}-t_{i-1})^2\left\{\frac{R_h^2S_h^2}{r_h}\left(1-\frac{r_h}{R_h}\right)\right\}\right],$$

where R_h is the number of reaches in the stream and r_h is total number of reaches sampled. If the entire stream is sampled (A-5), Line (2.17) becomes zero.

The $V_4(\sum_{i=2}^{k-1}Y_i)$ part of Line (2.18) depends on the sampling scheme used to select the dates surveyed. Let Y_i be the value associated with each survey as calculated in Line (2.18). If surveys are conducted daily such that $k=K$ (A-14), then there is no uncertainty contributed from the survey date selection and $V_4(Y)=0$. If the days are chosen at random from the K days of the run (A-8), then Line (2.18) becomes (Cochran, 1977, page 23):

$$+\frac{1}{4}\frac{K^2S_Y^2}{k}\left(1-\frac{k}{K}\right)\quad\text{where } S_Y^2=\frac{\sum_{i=1}^{K}(Y_i-\bar{Y})^2}{K-1}.$$

In order to randomly select days for surveys, the length of the run must be known in advance. More commonly, surveys are conducted systematically until the run is over. In systematic sampling (SYS) surveys are conducted every κ days (A-18) so there are κ possible sample sets, one starting on each of the first κ days of the run. SYS is more precise than SRS when the variation within the selected sample set is larger than the variation in the population overall (Cochran, 1977, page 208). However, there is no unbiased estimator of the variance for systematically sampled days because the selected sample set constitutes a single draw from the κ possible sets. The reader is referred to Wolter (1985) for more details including a discussion of eight different variance estimators for the mean of systematic samples.

In stratified random sampling (STR), the spawning season is divided into Z strata and d_z survey dates are randomly selected out of the possible D_z dates from each strata. Each observation Y_{ij} is the quantity in the brackets in Line (2.18) associated with the ith observation from the jth stratum, so Line (2.18) should be rewritten as:

$$\begin{aligned} & +\tfrac{1}{4}V_4\left(\textstyle\sum_{z=1}^{Z}\frac{D_z}{d_z}\sum_{i=1}^{d_z}Y_{iz}\right) \\ = & +\tfrac{1}{4}\textstyle\sum_{z=1}^{Z}D_z^2\left(1-\frac{d_z}{D_z}\right)\frac{S_z^2}{d_z}, \end{aligned} \tag{2.20}$$

where $S_z^2 = \frac{\sum_{i=1}^{D_z}(Y_{iz}-\bar{Y}_z)^2}{D_z-1}$ and $\bar{Y}_z = \frac{1}{D_z}\sum_{i=1}^{D_z}Y_{iz}$.

The variance estimate of Equations (2.15)-(2.18) is derived by substituting parameters with their estimated values and the variances with their estimates. For example, the variance estimator for fish-days when both stream reaches and are survey dates are chosen randomly within strata (STR) is:

$$\begin{aligned} \hat{V}(\hat{F}) = & \tfrac{k}{4K}\textstyle\sum_{i=2}^{k-1}\left[(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}\frac{R_{hi}}{r_{hi}}\sum_{j=1}^{r_{hi}}\left\{\frac{\hat{V}(\hat{P}_{ijh})}{\hat{P}_{ijh}^4}[l_{ijh}(1-\hat{P}_{ijh})+(l_{ijh})^2]\right\}\right] \\ & +\tfrac{k}{4K}\textstyle\sum_{i=2}^{k-1}\left[(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}\frac{R_{hi}}{r_{hi}}\sum_{j=1}^{r_{hi}}\left\{\frac{l_{ijh}(1-\hat{P}_{ijh})}{\hat{P}_{ijh}^2}\right\}\right] \\ & +\tfrac{k}{4K}\textstyle\sum_{i=2}^{k-1}\left[(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}R_h(R_h-r_h)\frac{s_h^2}{r_h}\right] \\ & +\tfrac{1}{4}\textstyle\sum_{z=1}^{Z}D_z^2\left(1-\frac{d_z}{D_z}\right)\frac{s_z^2}{d_z} \end{aligned} \tag{2.21}$$

where $s_z^2 = \frac{\sum_{i=1}^{d_z}(Y_{iz}-\hat{\bar{Y}}_z)^2}{d_z-1}$ and $\hat{\bar{Y}}_z = \frac{1}{d_z}\sum_{i=1}^{d_z}Y_{iz}$.

Non-Zero Tail Count Adjustment

The classic trapezoidal method requires that the first and last counts, i.e. the tails, are equal to zero ((A-12) and (A-13)). This may not be the case. Often the onset of the run begins earlier than anticipated, or lasts longer than the surveying effort. One way to resolve non-zero tail counts is to *a posteriori* select dates when the counts are suspected to have been zero. Irvine et al. (1992) drew the spawner abundance curve to zero between 5 and 15 days after the final survey day based on the condition of the fish on the final survey day. This approach is subjective, and the direction and magnitude of any resulting bias is unknown. There is no estimator for the variance.

Bue et al. (1998) extended the curve to zero at the time of one half of the average stream life value (A-20). This adjustment requires that the distribution of salmon arrivals before the first survey be uniform (A-11), which is not unreasonable. Given that a fish arrived before the first survey and survived until the survey, any arrival time within one stream life is equally likely. On average, the amount of time spent in the stream before the first survey would be half the stream life. Similarly, given that a fish is still alive at the end of the study, it will have, on average, half its stream life left. Even though the exact times of the first arrival and the last death cannot be measured, the Bue et al. (1998) adjustment is a biologically reasonable way to adjust for non-zero tail counts.

Trapezoidal Example with Non-zero Tail Counts

The data for this example are sockeye salmon in the Okanogan River in 1951 taken from Table 5 of Gangmark and Fulton (1952). The original data were not stratified, but for the purpose of this example, we assume that the numbers reported are the sum of two reaches randomly sampled from each of three strata (2.1). We assume that the first two strata have five possible reaches and that the third strata has four possible reaches. We also assume an observer efficiency unique to each day (equal within and among strata). The time surrounding sampling days is calculated as the number of days from the start of the previously sampled day to the start of the following sampling day (2.1). Since neither the first nor the last surveys had counts of zero, the tails are extrapolated to half a stream life before the first survey and after the last survey. We used the stream life of $\hat{T} = 11.85$ days calculated from the non-parametric timer tag method described in Section 3.4.15. Survey dates are also assumed to be randomly selected from two strata comprising of the first and last halves of the season.

Table 2.1: Data for trapezoidal fish-days example. Taken from Gangmark and Fulton (1952, Table 5) with dummy estimates of observer efficiency (P).

Date	Survey (i)	Time ($t_{i+1} - t_{i-1}$)	Strata 1 ($l_{i1.}$)	Strata 2 ($l_{i2.}$)	Strata 3 ($l_{i3.}$)	Obs. Ef. ($\hat{P}_i$)	Estimate $\hat{L}_i$
9/19	N/A	N/A	0	0	0	N/A	0
9/25	1	15	489	17	7	0.90	1,421.111
10/4	2	15	4,172	376	112	0.85	13,640
10/10	3	13	1,795	1,065	354	0.80	9,822.5
10/17	4	13	865	450	64	0.85	4,018.235
10/23	N/A	N/A	0	0	0	N/A	0

Using Equation (2.7), the estimate for the number of sockeye present during the first survey is:

$$\hat{L}_1 = \left(\frac{5}{2}\right)\frac{489}{.90} + \left(\frac{5}{2}\right)\frac{17}{.90} + \left(\frac{4}{2}\right)\frac{7}{.90} = 1,421.111, \tag{2.22}$$

and the season total estimate for escapement of sockeye is (see Equation (2.9)):

$$\begin{aligned} \hat{N} &= \tfrac{1}{2}[15(1,421.111) + 15(13,640) + 13(9,822.5) + 13(4,018.235)] \\ &= \tfrac{1}{2}[21316.67 + 204600 + 127692.5 + 52237.06] \\ &= 202,923.1 \end{aligned} \tag{2.23}$$

The uncertainty of the trapezoidal estimator of escapement arises from four sources: estimating observer efficiency; the binomial sighting process; selecting surveyed reach sites; and selecting survey dates. Steps used to estimate the variance contributed by each source for the first actual survey are presented along with the variance estimate from the season using Equation (2.21).

To estimate the variance contributed by estimating observer efficiency, the individual reach counts, not merely the strata total for each day, must be known. Strata totals are assumed as shown in Table 2.2.

Table 2.2: Individual reach counts and estimated strata variance. Data taken from Gangmark and Fulton (1952, Table 5) with dummy estimates of observer efficiency (P).

Date	Survey	Strata 1	Strata 2	Strata 3
	(i)	($l_{i1.}$)	($l_{i2.}$)	($l_{i3.}$)
9/19	N/A	0,0	0,0	0,0
9/25	1	196, 293	7, 10	3, 4
10/4	2	1669, 2503	150, 226	45, 67
10/10	3	718, 1077	426, 639	142, 212
10/17	4	346,519	180, 270	26, 38
10/23	N/A	0,0	0,0	0,0

Since observer efficiencies are the same across reaches and strata for each day, the variance contributed to the trapezoidal estimate of escapement can be written as (c.f.Equation (2.15)):

$$\frac{k}{4K}\sum_{i=2}^{k-1}\left[(t_{i+1}-t_{i-1})^2\frac{\hat{V}(\hat{P}_i)}{\hat{P}_i^4}\sum_{h=1}^{H}\frac{R_{hi}}{r_{hi}}\sum_{j=1}^{r_{hi}}[l_{ijh}(1-\hat{P}_i)+(l_{ijh})^2]\right]. \qquad (2.24)$$

For the first survey day ($i = 1$), the quantity in the large brackets is:

$$\begin{aligned}
& 15^2 \tfrac{0.05}{0.90^4} \quad \left\{ \tfrac{5}{2}(196(0.1) + 196^2 + 293(0.1) + 293^2) \right. \\
& \qquad\qquad + \tfrac{5}{2}(7(0.1) + 7^2 + 10(0.1) + 10^2) \\
& \qquad\qquad \left. + \ \tfrac{4}{2}(3(0.1) + 3^2 + 4(0.1) + 4^2) \right\} \qquad (2.25) \\
& = 15^2 \tfrac{0.05}{0.90^4} \quad \{310,784.75 + 376.75 + 51.40\} \\
& = 3,890,161
\end{aligned}$$

The uncertainty contributed to the trapezoidal estimate of escapement from the binomial sighting process is (c.f.Equation (2.16)):

$$\frac{k}{4K} \sum_{i=2}^{k-1} \left[(t_{i+1} - t_{i-1})^2 \sum_{h=1}^{H} \frac{R_{hi}}{r_{hi}} \sum_{j=1}^{r_{hi}} \left\{ \frac{l_{ijh}(1 - \hat{P}_i)}{\hat{P}_i^2} \right\} \right]. \qquad (2.26)$$

For the first survey day ($i = 1$), the quantity in the large brackets is:

$$\begin{aligned}
& 15^2 \left\{ \tfrac{5}{2} \left(\tfrac{489(0.1)}{(0.9)^2} \right) + \tfrac{5}{2} \left(\tfrac{17(0.1)}{(0.9)^2} \right) + \tfrac{4}{2} \left(\tfrac{7(0.1)}{(0.9)^2} \right) \right\} \\
& = 15^2 \{150.93 + 5.25 + 1.73\} \qquad (2.27) \\
& = 35527.78
\end{aligned}$$

The uncertainty contributed to the trapezoidal estimate of escapement from the binomial sighting process is (c.f.Equation 2.17):

$$\frac{k}{4K} \sum_{i=2}^{k-1} \left[(t_{i+1} - t_{i-1})^2 \sum_{h=1}^{H} R_h (R_h - r_h) \frac{s_{ih}^2}{r_h} \right]. \qquad (2.28)$$

where the within strata sampling variance s_{ih}^2 is estimated as:

$$s_h^2 = \frac{1}{r_h - 1} \sum_{j=1}^{r_h} \left(\frac{l_{ijh}}{\hat{P}_i} - \bar{\hat{L}}_{ih} \right)^2 \qquad \text{where} \quad \bar{\hat{L}}_h = \frac{1}{r_h} \sum_{h=1}^{r_h} \frac{l_h}{\hat{P}_h}. \qquad (2.29)$$

For instance, the variance within the first strata is:

$$s_{11}^2 = \frac{1}{2 - 1} \left[\left(\frac{196}{0.9} - \frac{489}{2(0.9)} \right)^2 + \left(\frac{293}{0.9} - \frac{489}{2(0.9)} \right)^2 \right] = 5808.025 \qquad (2.30)$$

For the first survey day ($i = 1$), the quantity in the large brackets is:

$$\begin{aligned}
&15^2 \left\{5(5-2)\tfrac{s_{11}^2}{2} + 5(5-2)\tfrac{s_{12}^2}{2} + 4(4-2)\tfrac{s_{13}^2}{2}\right\} \\
&= 15^2\{43,560.185 + 41.667 + 2.469\} \\
&= 9,810,972.
\end{aligned} \tag{2.31}$$

The uncertainty contributed to the trapezoidal estimator due to the selection of survey days is (c.f.Equation (2.21)):

$$\frac{1}{4}\{D_1^2\left(1-\frac{d_1}{D_1}\right)\frac{s_{d1}^2}{d_1} + D_2^2\left(1-\frac{d_2}{D_2}\right)\frac{s_{d2}^2}{d_2}\}, \tag{2.32}$$

when the survey dates have been randomly selected within the first and second halves of the run. The within-time-strata variability for the first half of the run is given by:

$$\begin{aligned}
s_{d1}^2 &= (0-75305.56)^2 + (15(1421.11)-75305.56)^2 + (15(13640)-75305.56)^2 \\
&= 12,651,390,093
\end{aligned} \tag{2.33}$$

so the uncertainty contributed to the example trapezoidal estimator by date selection is:

$$\begin{aligned}
&\tfrac{1}{4}\left\{(19)^2\left(1-\tfrac{3}{19}\right)\tfrac{12,651,390,093}{3} + (16)^2\left(1-\tfrac{3}{16}\right)\tfrac{4,121,268,079}{3}\right\} \\
&= 391,937,195,715
\end{aligned} \tag{2.34}$$

The variance estimate for the entire season is:

$$\begin{aligned}
\hat{V}(\hat{F}) &= \tfrac{6}{4(34)}\left[3890161 + 302097337 + 61208764 + 12333966\right] \\
&\quad + \tfrac{6}{4(34)}\left[35527.78 + 541588.2 + 415000.6 + 119838.0\right] \\
&\quad + \tfrac{6}{4(34)}\left[9810972 + 819331038 + 175136152 + 33424984\right] \\
&\quad + 391,937,195,715 \\
&= 391,999,769,773.
\end{aligned} \tag{2.35}$$

The standard error is $SE(\hat{N}) = \sqrt{391,999,769,773} = 626098.8$ salmon and the coefficient of variation is:

$$\hat{CV}(\hat{N}) = \frac{626098.8}{202,923.1} = 3.085. \tag{2.36}$$

In this example, the uncertainty in the escapement estimate renders the estimate useless. Not only is the estimate unbiased when surveys are conducted daily, but uncertainty is reduced by five orders of magnitude.

2.7.3 Parametric Fish-days

One of the difficulties with the trapezoidal method is that it requires the first and last counts to be zero. Often surveys are terminated after the peak of the run or after certain escapement goals have been met so that end-of-the-run data are not available (Su et al., 2001). Hilborn et al. (1999) developed a likelihood model which assumed that the arrivals and deaths followed Gaussian distributions. Su et al. (2001) applied a Bayesian model to the parametric model of Hilborn et al. (1999) for 24 years of data for Kadashan Creek in Alaska. The estimates for years without end of the run data "borrowed strength" from other years. Adkison and Su (2001) performed a simulation study to compare the estimates from the annual likelihood models with the estimates from the heirarchial Bayesian model of Su et al. (2001). They found that for years with little or no post-peak counts, the Bayesian model estimates were less biased and more precise than estimates for each year individually.

The difficulty with a likelihood model for fish-days is that one must assume an arrival distribution. Hilborn et al. (1999), Su et al. (2001), and Adkison and Su (2001) all assumed that salmon arrival times were distributed according to a Gaussian distribution (A-21). Dam passage and weir counts show that salmon often arrive in irregular pulses (e.g. Davidson, 1966). The direction and magnitude of any bias in the estimates based on a likelihood model are unknown due to the unknown lack-of-fit to the fitted likelihood distribution. In fact, while borrowing data from data rich years to supplement data from data-sparse years may improve the estimates of the sparse years (Su et al., 2001) the trapezoidal estimator was generally more accurate than the parametric fish-days estimator when compared year by year (Hilborn et al., 1999).

2.7.4 Average Spawner Method

The traditional Trapezoidal Method treats the survey dates as though they were fixed, and uncertainty in the estimate due to survey date selection as zero. This underestimates the true variance if surveys are conducted at less than daily intervals (Liao, 1994). The Average Spawner Method, originally developed by Gangmark and Fulton (1952) treats the periodic live counts of spawners as a representative sample from the total number of days in the run. The average of the periodic counts is multiplied by the total number of days in the run, then divided by stream life. Gangmark and Fulton (1952) found that the ratio of the duration of the run of Columbia River "bluebacks" (sockeye) divided by the estimated stream life was $35/7 = 5$. So the average periodic count was multiplied by 5 and the method was dubbed "The Factor 5 Method." The Factor 5 Method was miss-applied in Washington State. Rather than multiplying the average count by the run length, a single count, often the peak count was used (Gangmark and Fulton, 1952; Ames and Phinney, 1977). Multiplying a single count by five was found to be wildly inaccurate and the Factor 5 Method was abandoned in favor of current AUC methods (Ames and Phinney, 1977). The original technique of Gangmark and Fulton (1952) remains valid, however, and a variance estimate has since been developed by Liao (1994).

The average abundance, calculated from periodic counts of live salmon is $\bar{l} = \frac{1}{k}\sum_{i=1}^{k} l_i$. Note that Gangmark and Fulton (1952) denoted average abundance as "M," and the total length of the run as "D," whereas we have used K. In my notation, the Gangmark and Fulton (1952) estimator of escapement is:

$$\hat{N} = \bar{l}\frac{K}{T} = \frac{\bar{l}K}{T}, \tag{2.37}$$

such that the estimator of fish days is

$$\hat{F} = \bar{l}K = \frac{K}{k}\sum_{i=1}^{k} l_i. \tag{2.38}$$

By omitting any adjustment for observer efficiency, Gangmark and Fulton (1952) assumed that the probability of detection was 100% (A-22). Visual counts of salmon depend on weather and water conditions (e.g. Gangmark and Fulton, 1952) and are usually an undercount of the true number of salmon present. Therefore Equation (2.38) is likely negatively biased for total fish days to an unknown degree.

Liao (1994) independently developed an estimator similar to Equation (2.38) based on finite population sampling which included an estimate of detection probability (A-1):

$$\hat{F} = \frac{K}{k}\sum_{i=1}^{k}\hat{L}_i = \frac{K}{k}\sum_{i=1}^{k}\frac{l_i}{\hat{P}_i}. \tag{2.39}$$

Liao (1994, page 35) showed that the estimator of fish days given in Equation (2.39) is first order Taylor series unbiased for any number of survey dates by conditioning on the selection of survey days (1) and the binomial sighting process (2):

$$\begin{aligned} E(\hat{F}) &= E\left(\tfrac{K}{k}\textstyle\sum_{i=1}^{k}\tfrac{l_i}{P_i}\right) = E_1\left[E_2\left(\tfrac{K}{k}\textstyle\sum_{i=1}^{k}\tfrac{l_i}{P_i}|k\right)\right] \\ &= E_1\left[\tfrac{K}{k}\textstyle\sum_{i=1}^{k}E_2\left(\tfrac{l_i}{P_i}|k\right)\right] \\ &= E_1\left(\tfrac{K}{k}\textstyle\sum_{i=1}^{k}L_i|k\right) \\ &= \textstyle\sum_{i=1}^{K}L_i = F. \end{aligned} \tag{2.40}$$

Gangmark and Fulton (1952) did not provide an estimate of variance for their estimator. When deriving the estimate of variance for his estimator of fish-days, Liao (1994) assumed that the entire stream was surveyed (A-5), eliminating Source 3 of the variance from Equation (2.6). He gave two formulas for the variance of his estimator for each of two survey-date sampling schemes. One formula incorporated the variance associated with the survey-date selection process (Source 4), the binomial sighting process (Source 2), and estimating observer efficiency (Source 1), but assumed that the observer efficiencies were the same for the entire season (A-23). The other formula assumed that the observer efficiencies were known without error (A-24), which is never really the case in practice.

Liao (1994) gave the variance of fish-days, $\hat{F}$, derived from a simple random sample (SRS) of survey dates with known observer efficiency as:

$$V_{SRS}(\hat{F}) = \underbrace{\frac{K}{k}\sum_{i=1}^{K}\frac{L_i(1-P_i)}{P_i}}_{\text{Source 2}} + \underbrace{\frac{K^2}{K}\left[1-\frac{k}{K}\right]\left(\frac{1}{K-1}\sum_{i=1}^{K}(L_i-\bar{L})^2\right)}_{\text{Source 4}} \qquad (2.41)$$

where $\bar{L}=\frac{1}{K}\sum_{i=1}^{K} L_i$. Equation (2.41) does not incorporate the variance associated with estimating observer efficiency, Source 1. The true variance also includes the variance associated with estimating P_i:

$$\text{Source 1 variance } = \left(\frac{K}{k}\right)^2 \sum_{i=1}^{K}\frac{V(\hat{P}_i)}{P_i^4}\left[L_iP_i(1-P_i)+(L_iP_i)^2.\right] \qquad (2.42)$$

Using an SRS date selection scheme, the variance is the sum of Equations (2.41) and (2.42) and is estimated by replacing parameters with their estimated values:

$$\begin{aligned} \hat{V}_{SRS}(\hat{F}) &= \left(\tfrac{K}{k}\right)^2 \textstyle\sum_{i=1}^{k}\frac{\hat{V}(\hat{P}_i)}{\hat{P_i}^4}\left[l_i(1-\hat{P}_i)+(l_i)^2\right] \\ &+ \tfrac{K}{k}\textstyle\sum_{i=1}^{k}\frac{l_i(1-\hat{P}_i)}{\hat{P_i}^2} \\ &+ \tfrac{K^2}{k}\left[1-\tfrac{k}{K}\right]\left(\tfrac{1}{k-1}\textstyle\sum_{i=1}^{k}(\hat{L}_i-\bar{\hat{L}})^2\right), \end{aligned} \qquad (2.43)$$

where $\bar{\hat{L}}=\frac{1}{k}\sum_{i=1}^{k}\frac{l_i}{\hat{P}_i}$. Equation (2.43) is unbiased for the variance of the fish-days estimate when the k survey dates are randomly selected from the K run days (A-8). Unfortunately, the number of days in the run is not usually known in advance and there may be days at the end of the run that have a smaller probability of being chosen for a survey than days at the beginning of the run.

The systematic random sampling (SYS) scheme is the most commonly used approach to selecting survey days. For SYS, surveys are conducted once every κ days with κ chosen *a priori* (A-18). SYS requires less effort and is therefore less expensive than daily surveys, and it affords a predictable surveying schedule over the season. Also, systematic sampling allows surveyors to add survey days to the run as necessary so that the total number of run days, K, does not have to be known in advance, as it does with SRS. The major drawback to SYS survey schemes is that there is no unbiased estimator for the variance (Wolter, 1985).

A better way to select survey days is to stratify the run by time and randomly select two or more days in each time strata to survey. This is called Stratified Random Sampling (STR) and it allows for random selection of survey days without knowing how many days there will be *a priori* because strata can be added or removed from the surveying schedule as the length of the run dictates. Under STR the estimate of fish-days, $\hat{F}$, is (Cochran, 1977, page 89):

$$\hat{F} = \sum_{z=1}^{Z} \frac{D_z}{d_z} \sum_{i=1}^{d_z} \frac{l_{zi}}{P_{zi}}, \tag{2.44}$$

where D_z is the number of days in stratum z, d_z is the number of days surveyed in stratum z, and Z is the total number of strata. If, for example, a run of 5 weeks ($Z = 5$) was stratified by week ($D_z = 7$) and two surveys were performed each week ($d_z = 2$), Equation (2.44) would be:

$$\hat{F} = \sum_{z=1}^{5} \frac{7}{2} \left(\frac{l_{z1}}{P_{z1}} + \frac{l_{z2}}{P_{z2}} \right). \tag{2.45}$$

As long as the survey dates are randomly selected within each stratum (A-8), Equation (2.44) is unbiased for total fish-days.

The variance of the estimate in Equation (2.44) depends on the variability within each stratum and is given by Cochran (1977, page 93) as:

$$V(\hat{F}) = \sum_{z=1}^{Z} D_z^2 (1 - \frac{d_z}{D_z}) \frac{S_z^2}{d_z} \tag{2.46}$$

where $S_z^2 = \frac{\sum_{i=1}^{D_z} (L_{zi} - \bar{L}_z)^2}{D_z - 1}$ is the true variance within stratum z. The variance of Equation (2.44) is estimated by substituting the sampling variance for each stratum, s_z^2, in place of the true within-stratum variance in Equation (2.46). The within-stratum sampling variance is:

$$s_z^2 = \frac{\sum_{i=1}^{d_z} (\hat{L}_{zi} - \bar{\hat{L}}_z)^2}{d_z - 1}. \tag{2.47}$$

Average Spawner Example

The data for this example are taken from Gangmark and Fulton (1952, Table 5). The data are from four systematic surveys conducted between the fourth and fifth bridges on the Okanogan River to count live sockeye salmon in 1947 (Table 2.3). Gangmark and Fulton (1952) assumed that observer efficiency was 100%. For this example, we have made up four "estimates" of observer efficiency, each with a variance estimate of 0.05 ($\hat{V}(\hat{P}_i|P_i) = 0.05$ for $i = 1, 2, 3, 4$). The run was judged to be 35 days long ($K = 35$).

Table 2.3: Data for average spawner example. Taken from Gangmark and Fulton (1952, Table 5) with dummy estimates of observer efficiency (P).

Date	Survey	Count	Obs. Ef.	Adjusted Count
	(i)	(l_i)	($\hat{P}_i$)	($\hat{L}_i = l_i/\hat{P}_i$)
10/1	1	254	0.85	298.82
10/8	2	1171	0.90	1301.10
10/15	3	1474	0.89	1656.20
10/24	4	424	0.82	517.07
Average				943.30

Using Equation (2.39) the average spawner estimate of fish-days is:

$$\hat{F} = \frac{35}{4}[298.82 + 1301.10 + 1656.20 + 517.07] = 33015. \tag{2.48}$$

The survey days were chosen systematically, and no unbiased estimate of the variance exists. Alternatively, assuming the survey days were chosen randomly for this example, the variance follows Equation (2.43):

$$\begin{aligned}
\hat{V}_{SRS}(\hat{F}) &= \left(\tfrac{35}{4}\right)^2 \Big[\tfrac{0.05}{0.85^4}(254(.15)+254^2)+\tfrac{0.05}{0.90^4}(1171(.10)+1171^2) \\
&\qquad +\tfrac{0.05}{0.89^4}(1474(.11)+1474^2)+\tfrac{0.05}{0.82^4}(424(.18)+424)^2\Big] \\
&+ \tfrac{35}{4}\Big[\tfrac{254(1-0.85)}{0.85^2}+\tfrac{1171(1-0.90)}{0.90^2}+\tfrac{1474(1-0.89)}{0.89^2}+\tfrac{424(1-0.72)}{0.82^2}\Big] \\
&+ \tfrac{35^2}{4}\left[1-\tfrac{4}{35}\right]\left(\tfrac{1}{4-1}\right)[(298.82-929.25)^2+(1301.10-929.25)^2 \qquad (2.49) \\
&\qquad +(1656.20-929.25)^2+(460.87-929.25)^2] \\
\hat{V}_{SRS}(\hat{F}) &= 23,254,888+4,510.64+111,505,956 \\
&= 134,765,355.
\end{aligned}$$

The standard error is estimated as:

$$S\hat{E}(\hat{F}) = \sqrt{134,765,355} = 11,609 \tag{2.50}$$

and the coefficient of variation is estimated as:

$$C\hat{V}(\hat{F}) = \frac{\sqrt{134,765,355}}{33,015} = 0.35162 \tag{2.51}$$

Survey date selection makes a contribution to the uncertainty which is an order of magnitude larger than the next largest contributor, estimating observer efficiency. As the number of surveys increases, the uncertainty contributed by the survey date selection decreases, thereby decreasing total variance.

2.8 Combining Fish-days and Average Stream Life

As stated in Section 2.7 the area-under-the-curve (AUC) estimate of escapement requires both an estimate of total fish-days, $\hat{F}$, and an estimate of average stream life, $\hat{T}$ (A-20):

$$\hat{N} = \frac{\hat{F}}{\hat{T}}, \tag{2.52}$$

which is unbiased to the first and second terms of a Taylor series expansion. In the third term of a Taylor series expansion Equation (2.52) is positively biased in the same manner as the estimate of live fish present in Equation (2.2). A more accurate estimate of escapement is (Rawding et al., 2008):

$$\hat{N}^{*} = \frac{\hat{F}}{\hat{T}}(1 - \frac{\hat{V}(\hat{T})}{\hat{T}^2}) = \hat{N}(1 - \hat{CV}(\hat{T})^2), \tag{2.53}$$

where an estimate of the variance of the stream life estimate must be available (A-25).

The variance of the overall estimate of escapement is usually derived using the delta method and is estimated as:

$$\hat{V}(\hat{N}) = \hat{V}(\hat{F})\left(\frac{1}{\hat{T}}\right)^2 + \hat{V}(\hat{T})\left(\frac{\hat{F}}{\hat{T}^2}\right)^2 + 2\hat{Cov}(\hat{F},\hat{T})\left(\frac{1}{\hat{T}}\right)\left(\frac{\hat{F}}{\hat{T}^2}\right), \tag{2.54}$$

where the covariance of fish-days and average stream life must be estimated.

Often the estimates of fish-days, $\hat{F}$, and average stream life, $\hat{T}$, are made independently (A-26), in which case the variance estimator simplifies to:

$$\begin{aligned} \hat{V}(\hat{N}) &= \hat{V}(\hat{F})\left(\tfrac{1}{\hat{T}}\right)^2 + \hat{V}(\hat{T})\left(\tfrac{\hat{F}}{\hat{T}^2}\right)^2 \\ &= \tfrac{\hat{F}^2}{\hat{T}^2}\left[\tfrac{\hat{V}(\hat{F})}{\hat{F}^2} + \tfrac{\hat{V}(\hat{T})}{\hat{T}^2}\right] \\ &= \hat{N}^2[\hat{CV}^2(\hat{F}) + \hat{CV}^2(\hat{T})] \end{aligned} \tag{2.55}$$

However, when the estimates of fish-days and stream life are independent Goodman's formula can be used with the Delta Method to provide a more accurate expression for variance:

$$\begin{aligned} V(\hat{N}) &= \tfrac{F^2}{T^4}V(\hat{T}) + \tfrac{V(\hat{F})}{T^2} + \tfrac{V(\hat{F})V(\hat{T})}{T^4} \\ &= N^2[CV^2(\hat{F}) + CV^2(\hat{T})] + CV^2(\hat{T})\tfrac{V(\hat{F})}{T^2}, \end{aligned} \tag{2.56}$$

which is estimated by:

$$\hat{V}(\hat{N}) = \hat{N}^2[\hat{CV}^2(\hat{F}) + \hat{CV}^2(\hat{T})] - \hat{CV}^2(\hat{T})\frac{\hat{V}(\hat{F})}{\hat{T}^2}. \tag{2.57}$$

Combining Fish-days and Average Stream Life Example

The estimate of fish-days used in this example, and its variance, were derived using the average spawner method in Section 2.7.4. The estimates of stream life and stream life variance were derived using the non-parametric timer tag method described in Section 3.4.15.

The area-under-the-curve estimate of salmon escapement is:

$$\hat{N} = \frac{\hat{F}}{\hat{T}} = \frac{33,015}{11.85} = 2,786.1 \tag{2.58}$$

Methods used to estimate fish-days and stream life are assumed independent so that the variance estimate of $\hat{N}$ based on the delta method is given by Equation (2.55) as:

$$\begin{aligned} \hat{V}(\hat{N}) &= (2,786.1)^2[(0.35162)^2 + (0.03131)^2] \\ &= 7,762,353[0.12364 + 0.000980312] \\ &= 967,320 \end{aligned} \tag{2.59}$$

and the standard error is:

$$\hat{SE}(\hat{N}) = \sqrt{967,320} = 983.52 \tag{2.60}$$

so the coefficient of variation is:

$$\hat{CV}(\hat{N}) = \frac{\sqrt{967,320}}{2,786.1} = 0.35301 \tag{2.61}$$

The variance estimate of $\hat{N}$ based on Goodman's formula as given by Equation (2.57) as:

$$\begin{aligned} \hat{V}(\hat{N}) &= (2,786.1)^2[(0.35162)^2 + (0.03131)^2] - \tfrac{134,765,355}{(11.85)^2}(0.03131)^2 \\ &= 7,762,353[0.12364 + 0.000980312] - (959713.4)(0.000980312) \\ &= 966,410 \end{aligned} \tag{2.62}$$

and the standard error is:

$$\hat{SE}(\hat{N}) = \sqrt{966,410} = 983.06 \tag{2.63}$$

so the coefficient of variation is:

$$\hat{CV}(\hat{N}) = \frac{\sqrt{966,410}}{2,786.1} = 0.35285. \tag{2.64}$$

2.9 Discussion of Fish-days Techniques

The Ames Method produces an estimator of unknown bias, and has no variance estimator (Table 2.4). This method served as the conceptual framework upon which standardized, unbiased estimators could be built, but should not be used in salmon fisheries management.

The trapezoidal method of AUC is widely used because of its simplicity in both comprehension and computation. Unfortunately, the trapezoidal estimator and estimate of variance are unbiased only if the stream is surveyed every day of the run. While the trapezoidal estimator can provide unbiased estimates of escapement, it requires strict assumptions and conditions to do so, and so should be considered only when daily surveys are possible (Table 2.4).

The Likelihood Area-Under-the-Curve method is mathematically elegant, but not applicable to salmon escapement (Table 2.4). The arrival patterns of Pacific Northwest salmon have been shown to be different than assumed arrival patterns used in the parametric models. Hilborn et al. (1999) showed that the likelihood-based estimates were less accurate than the trapezoidal-based estimates. Since the applicability of the likelihood model for a given year is not known and the estimate are less accurate than other more applicable methods, the likelihood method of estimating fish-days is not recommended for regular management of salmon escapement.

The average spawner method provides unbiased estimates of both fish-days and variance, and requires design-based assumptions that are more attainable than the other methods (Table 2.4). Using a stratified random selection of days on which to perform surveys allows managers to add and drop survey strata (i.e. weeks), while still ensuring an unbiased estimate of the variance. The Average Spawner Method is theoretically the best method available to estimate fish-days for use in the AUC Method. The only drawback is that it is infrequently used in practice so there is little empirical evidence to evaluate its efficacy and assumption achievability.

Table 2.4: Evaluation of fish-days estimation methods for use in the area-under-the-curve method

Name	Estimator	Variance	Assumptions	Empirical Experience	Recommendation
Ames	Poor	None	Unreasonable	Imprecise	Do Not Use
Trapezoidal	Fair	Good	Reasonable	Can be accurate	Good Alternative
Parametric	Unknown	Good	Unreasonable	Less accurate than trapezoidal	Do Not Use
Average Spawner	Good	Good	Reasonable	Not widely used	Best Choice

Chapter 3

AREA UNDER THE CURVE: STREAM LIFE

3.1 Introduction to AUC Stream Life

In the previous chapter we addressed the different methods employed to estimate fish-days and how to combine fish-days with an estimate of stream life according to the area-under-the-curve method. This chapter examines different ways of estimating the average length of time of each salmon is alive in the spawning area. This quantity is most often referred to as stream residence time (Irvine et al., 1992; Hill, 1997; Lady and Skalski, 1998; Manske and Schwarz, 2000) but is also called survey life (Ames, 1984; Perrin and Irvine, 1990) or stream life (Bue et al., 1998). We use the term "stream life" for simplicity with the assumption that it applies only to regions of the stream being surveyed.

Bias in the estimate of average stream life has the potential to greatly bias the estimate of escapement (Perrin and Irvine, 1990; Su et al., 2001). Because stream life is in the denominator of the escapement estimate, an over-estimate of the average stream life will result in an under-estimate of total escapement. Conversely, an under-estimate in average stream life will result in an over-estimate in total escapement.

Perrin and Irvine (1990) reviewed nine techniques used to estimate "survey life." Six of the reviewed techniques (RT1-RT6) estimated live residence time in a specified natural survey area, one technique estimated redd residence time (RTR), one technique was applied to hatchery life (FLH), and the last was "true" survey life (TSL). None of the six estimators for natural streams were found to be without bias (Perrin and Irvine, 1990; Liao, 1994). In the following sections we describe the six stream life estimation techniques discussed by Perrin and Irvine (1990), and nine other methods used to estimate average stream life.

3.2 *Stream Life Notation*

Notation used throughout this chapter is listed below. All parameter estimates are denoted with carets, for example, stream life, T, is estimated by $\hat{T}$. Expected values, variances, and covariances are denoted $E()$, $V()$, and $Cov()$, respectively.

$\alpha=$ Rate parameter of the Weibull distribution.

$\beta=$ Shape parameter of the Weibull distribution.

$B_i=$ Number of individuals arriving after survey $i-1$ and surviving until the time of survey i, i.e. "births".

$b_i=$ Number of newly tagged individuals released at the time of the i survey.

$c_i=$ Number of marked individuals seen for the last time at time of the i survey.

$D_i=$ Actual number of dead individuals present at the time of the i survey.

$d_i=$ Observed number of dead individuals at the time of the i survey.

$F_T(q)=$ The cumulative distribution of random variable T up to quantile q.

$F_T^{-1}=$ The time at which the cumulative distribution of T reaches quantile q.

$f(x)=$ The Probability distribution function of random variable x.

$K=$ Number of days in the run.

$k=$ Number of survey occasions, $(k \leq K)$.

λ= The survival and detection probabilities of the final survey occasion.

L_i= Actual number of live individuals at the time of the i survey.

l_i= Observed number of live individuals at the time of the i survey.

M_i= Total number of tagged individuals alive at the time of the i survey.

m_i= Number of previously marked individuals resighted at the time of the i survey.

N= Escapement, total number of individuals in a run.

N_h= Number of units in a strata.

n_h= Number of units sampled in strata h.

P_D= Probability of detecting a dead fish given a carcass is present.

P_L= Probability of detecting a live fish given a live fish is present.

ϕ_i= Probability of survival from the time of survey i to $i+1$, conditional on being alive at the time of the ith survey.

q= Quantile.

r_i= Death rate for the interval from survey occasion i to $i+1$.

s_i= Number of individuals seen in on survey occasion i.

T= Average stream life, average residence time.

T_j= Stream life of individual j.

t_i= Time of the ith survey.

$\tau_i =$ $m_i + b_i - c_i$, Total number of marked individuals seen at the time of the i survey and later.

$\theta =$ theta.

$u_h =$ Number of individuals with a given capture history, where h is a series of 1's and 0's.

$v_i =$ Number of marked individuals known to be alive after the time of the ith survey.

$\chi_i =$ Probability of being alive at the time of the ith survey and not seen again after.

$w_i =$ Weight given to the count on the ith survey.

$Z =$ Total number of time strata.

3.3 AUC Stream Life Assumptions

Assumptions used by one or more of the stream life estimation methods are listed below. Not all assumptions are required by any one method, and only (A-1) and (A-2) are applicable to all methods. The assumptions are discussed in more detail as they arise in descriptions of individual methods.

(A-1) Stream life refers to the duration of time between entrance into the spawning area and death.

(A-2) All fish entering the study stream remain until death, i.e. no fish swim out of the spawning area once they have entered.

(A-3) Average stream life for a given species remains constant from year to year.

(A-4) Average stream life for a given species remains constant from stream to stream.

(A-5) Tagged fish constitute a representative random sample of the fish in the survey area over the course of a run; i.e. all fish have an equal probability of being tagged.

(A-6) Handling and tagging have no impact on the survival of the fish.

(A-7) Each carcass is counted at most one time

(A-8) All fish have the same stream life.

(A-9) All carcasses are detected immediately following their death.

(A-10) The cumulative abundance curve can be used as a proxy to estimate the arrival curve of migrating salmon.

(A-11) Records of carcasses are kept for the entire stream.

(A-12) The detection probability of carcasses for each survey is estimable.

(A-13) The processes governing the arrival and survival of live fish are the same as those governing the death and decay processes of carcasses.

(A-14) Detection probability is 100%.

(A-15) Live fish are counted only once, for instance at a weir.

(A-16) An unbiased estimate of escapement is available.

(A-17) A variance estimate for the escapement estimate is available.

(A-18) An unbiased estimate of fish-days is available.

(A-19) The estimates of fish-days and escapement pertain to the same population.

(A-20) A variance estimate for the fish-days estimator is available.

(A-21) The estimates of fish-days and escapement are made independently.

(A-22) Strata constitute a mutually exclusive and exhaustive partition of the study area.

(A-23) Sampled reaches or days within the strata are selected randomly.

(A-24) The count for the last survey is zero.

(A-25) The stream is surveyed every day of the run.

(A-26) Arrivals and mortalities occur uniformly over a sampling interval.

(A-27) Redd residence time is the same for both males and females.

(A-28) The probability of capture is independent of previous capture history.

(A-29) Each fish has an equal probability of capture during a survey.

(A-30) Each fish has an equal probability of survival during a survey.

(A-31) Survival and captures of all fish are independent.

(A-32) Survival probabilities are independent of age.

(A-33) Mortalities occurring between surveys follow an exponential distribution.

(A-34) Survival probabilities (and hence Stream life) remain constant over the course of a run.

(A-35) The time between surveys is the same for all intervals.

(A-36) Catch rates (capture probabilities) are the same for all tagging events.

(A-37) All carcasses present during a survey have the same detection probability.

(A-38) Stream life times are distributed according to a two-parameter Weibull distribution.

(A-39) A tagged fish is dead if and only if it holds a pre-determined angle for a pre-determined amount of time.

3.4 Stream Life Estimation Techniques

3.4.1 Constant Stream Life Method

For the constant stream life method, a stream life estimate from one or more years in a particular stream is applied to other years, different streams, or both. Gangmark and Fulton (1952) used the average stream life of 7 days from sockeye in Karluk Lake, Alaska for “Bluebacks” (sockeye) in the Columbia River, Washington: $\hat{T} = 7$. Beidler and Nickelson (1980) took the average stream life of coho salmon in Oregon State to be 11 days: $\hat{T} = 11$. Ames (1984) averaged the estimated stream lives reported from four different studies done in 1952 and in the 1970’s and concluded that the stream life of the Pacific chum salmon was 10 days: $\hat{T} = 10$.

The major assumption of the Constant Stream Life Method is that the stream life remains constant from year to year (A-3) or from stream to stream (A-4) for each salmon species. Perrin and Irvine (1990) conducted an extensive survey of the literature and of fisheries managers to find what methods were being used to estimate salmon stream life and the values of the estimates for each of the five North American species of salmon. They found that stream life does not remain constant from year to year. For example, in 1984 the estimated stream life for chum in the Nekite River, B.C. was 19.0 days. The following year, using the same method, it was estimated to be 4.3 days. Perrin and Irvine (1990) also found that stream life did not remain constant from stream to stream. In 1983 estimates for stream life of early run pink salmon in Alaska ranged from 14.5 days to 40.5 days.

Due to the large variation in residence times between years and between streams, Perrin and Irvine (1990) recommended that stream life be estimated for each stream and season of interest. The direction and magnitude of any bias associated with using a constant stream life value is unknown. If the estimate of stream life is larger than the true stream life then the estimate of escapement will be negatively biased. If the estimate of stream life is smaller than the true value, then the estimate of escapement will be positively biased.

The variance associated with the use of a constant stream life value is zero, because a constant has no variability. Unfortunately,the apparent gain in precision is counterbalanced by bias when the uncertainty of the estimate is expressed as a mean squared error (MSE):

$$MSE(\hat{T}) = V(\hat{T}) + Bias(\hat{T})^2$$

Even though the variance is zero, the bias may be so large that the estimate is useless.

3.4.2 Time To Tag Recovery Method

Killick (1955) tagged live sockeye as they entered Adams River, B.C. and recovered tags from carcasses during daily foot surveys. Stream life was estimated to be the average number of days between tagging and recovery. Thomason and Jones (1984) did a similar study on pink and chum salmon in three rivers in Alaska. Stream life was estimated using the mode, mean, and median number of days from tagging to recovery. Perrin and Irvine (1990) called this the RT1 method, and found that it was the method most commonly used to estimate stream life for pink, chum, coho, and sockeye salmon.

The measured stream life for the jth tagged fish is:

$$T_j = t_j - t_{0j}, \quad j = 1, \ldots, m \tag{3.1}$$

where t_j is the day the jth fish is recovered dead, t_{0j} is the day the jth fish is originally tagged, and m is the number of tagged fish that are recovered.

The mode estimator is the T_j that occurs the most frequently. In a histogram of the survival times, the mode is the stream life value that corresponds to the top of the curve. The mode is not necessarily unique, as in the case of multimodal distributions. In the case of non-unique modes, Liao (1994) recommended taking the arithmetic mean of the modes.

The mean estimate is the average number of days it takes to recover each tag, which is the sum of the individual recovery times, T_j, divided by the total number of tags recovered, m (Equation from Fukushima and Smoker, 1997):

$$\hat{T}_{mean} = \frac{1}{m}\sum_{j=1}^{m} T_j \tag{3.2}$$

The median estimate is the day after tagging on which 50% of the tags which are recovered, $F(T_j) = 0.5$, have been found. If the individual recovery times, T_j, are relabeled in ascending order such that $T_1 < T_2 < \ldots < T_m$, then the median estimator is:

$$\hat{T}_{median} = T_{\frac{m+1}{2}}. \tag{3.3}$$

If there are an even number of recovered tags, then the arithmetic mean of the two central values is taken as the median:

$$\hat{T}_{median} = \frac{1}{2}(T_{\frac{m}{2}} + T_{\frac{m}{2}+1}). \tag{3.4}$$

All three stream life estimators are positively biased to an unknown degree due to the lag between mortality and recovery of tags (Perrin and Irvine, 1990). Both Killick (1955) and Thomason and Jones (1984) attempted to reduce this bias by excluding tags without bodies (lost tags) and tags from old carcasses in the stream life estimates, but the magnitude of the bias is still unknown. Carcasses that had tags but lost them must be excluded because entry time is unknown. Thomason and Jones (1984) also excluded mortalities due to predation from the stream life estimate. Excluding predation mortalities as "premature" mortalities has the potential to contribute further positive bias, because predation is a natural part of the process being studied.

Thomason and Jones (1984) found that the sampling distribution of 'time from tagging to recovery' was right skewed, meaning that fewer fish had stream lives shorter than the sample mode and more fish had stream lives longer than the sample mode. In this case, the sample mode is negatively biased (Lee, 1992) for the mean of tagging to recovery time by an unknown amount. The sample median is also negatively biased (Lee, 1992) for the mean of tagging to recovery time, also by an unknown amount.

If the fish tagged are a representative sample of the population (A-5) and tagging does not have an effect on stream life (A-6), the sample mean is unbiased for mean time from tagging to recovery, which includes both the stream life of the fish and the time from mortality to recovery. The variance of the sample mean is given as (Cochran, 1977, page 23):

$$V(\hat{T}) = \frac{S^2}{M}\left(1 - \frac{M}{N}\right) \qquad \text{where} \quad S^2 = \frac{1}{N-1}\sum_{j=1}^{N}(T_j - T)^2. \tag{3.5}$$

The variance is estimated by:

$$\hat{V}(\hat{T}) = \frac{s^2}{m}\left(1 - \frac{m}{\hat{N}}\right) \qquad \text{where} \quad s^2 = \frac{1}{m-1}\sum_{j=1}^{m}(T_j - \hat{T})^2, \tag{3.6}$$

which is slightly, but negligibly biased for the variance of $\hat{T}$ (Cochran, 1977, page 27). The second product term in Equation 3.5, $\left(1 - \frac{m}{\hat{N}}\right)$, is called the finite population correction (fpc), which must be estimated because the size of the population is not known. If the true proportion of the population tagged and recovered, m/N, is unknown but assumed to be very small, the fpc is close to unity and the variance can be approximated as:

$$\hat{V}(\hat{T}) = \frac{s^2}{m}. \tag{3.7}$$

Equations (3.5) and (3.6) depend on the assumption that the tagged fish are a representative sample of the population (A-5), which requires that fish be tagged throughout the season. If fish are tagged only during a portion of the season, the the estimate of variance may be biased in an unknown direction and to an unknown degree.

Iglewiez (1983) compared four different forms of scale estimators (i.e. estimators of the spread or variability of the data) including the standard deviation, mean absolute deviation, median absolute deviation, and fourth spread. He showed that the standard deviation (square root of sampling variance) and mean absolute deviation were not robust to outliers for symmetrical distributions. Of the remaining two, the median absolute deviation (MAD) estimator was the most efficient estimator of the spread for data that deviated strongly from a Gaussian distribution (Iglewiez, 1983, Table 12-5). Stream life data tend to deviate from a Gaussian distribution, so the MAD estimator may seem more appropriate than the sample variance. The MAD estimator is given by:

$$MAD = \text{median}(|T_j - \hat{T}_{median}|). \tag{3.8}$$

However, the MAD estimator is based on the sample median, rather than the sample mean. Also, the MAD is more comparable to the standard deviation than to the variance. Rosenberger and Gasko (1983) offered a formula for the actual variance of the median, but it requires knowing the distribution of the order statistics and is only applicable to data from symmetrical distributions. The reader is referred to Rosenberger and Gasko (1983) in Hoaglin et al. (1983) for further reading on robust location and scale parameter estimators.

3.4.3 Peak-to-peak Method

Perrin and Irvine (1990) called the Peak-to-Peak Method the RT2 method, and Liao (1994) called it the Mode-to-Mode method. Periodic counts are made of both live and dead fish, with the carcasses being removed or marked to prevent double counting (A-7). The average survey life is estimated as the number of days between the day of the highest live count and the day of the highest carcass count:

$$\hat{T} = t_{\max(d_i)} - t_{\max(l_i)} \tag{3.9}$$

In the case where the peak (of either live or dead counts) is not unique, the time of the peak must be estimated. Liao (1994) recommended averaging the times of the peaks in multi-modal distributions. If there were two peaks of live counts, $\max_1(l_i)$ and $\max_2(l_i)$ for example, the estimator would be:

$$\hat{T} = t_{\max(d_i)} - \left(\frac{t_{\max_1(l_i)} - t_{\max_2(l_i)}}{2} \right) \tag{3.10}$$

The Peak-to-Peak Estimator assumes that all fish have the same stream life (A-8), are detected as carcasses immediately upon death (A-9) and that the abundance of salmon can be used as a proxy for their arrival times (A-10). In which case, the time from the peak live count to the peak carcass count is approximately the time from the peak arrival of salmon to the peak time of death, which is the stream life of those fish. In practice, the peak live count is a result of the build up of fish that have been arriving over the course of the run, and cannot be used as a reliable proxy for arrival times. The fish present during the peak live count are at various stages in their stream lives, such that the difference in time between the peak live count and the peak dead count is shorter than the true average stream life. The fish alive at the time of the peak could have arrived at any time between one stream life prior to the peak and the peak. So on average, fish alive at the peak will have arrived half a stream life prior to the peak and have half a stream life left to live. Thus, if there is no lag between death and detection (A-9), then the Peak-to-Peak Estimator will intuitively be about half the true average stream life. Monte Carlo simulations have verified that the Peak-to-Peak Estimator underestimates the true stream life by about half, even under ideal detection circumstances (Liao, 1994, Table 5.1).

The Peak-to-Peak Estimator attempts to use the abundance curve as a proxy for the arrival curve. If it can be assumed that all fish have the same stream life (A-8), then the peak in fish arrival can be detected as the peak in detected carcasses one stream life time later. However, the largest pulse in abundance does not necessarily correspond to the largest pulse in arrival. Therefore the time between the peak live abundance and the peak carcass detection is irrelevant with respect to average stream life. Consequently, there is also no relevant variance estimate.

Killick (1955) found that the stream lives of individual sockeye varied from 12 to 19 days, with the greatest variations occurring at the time of peak spawning activity. Thus, the assumption of identical stream lives (A-8) seems unreliable, lending further evidence to not use the Peak-to-Peak Estimator.

Peak-to-Peak data can be used there are records from the number of carcasses for the entire run (A-11). If carcasses were detected with a known probability (A-12), the total number of carcasses detected, adjusted by observer efficiency, can be used to estimate escapement. The reader is referred to Chapter 8 for further details.

3.4.4 Cumulative Equivalence Points Method (RT3)

For the Cumulative Equivalence Points Method, live and dead fish are counted during periodic surveys. The cumulative counts of live and dead fish are plotted on the same timeline to form cumulative arrival and mortality curves. The time between equivalent points on the two curves is the estimate of residence time. Perrin and Irvine (1990) called this the RT3 method, and found that the most common percentiles were 10%, 50%, and 90%. The 50th percentile is the most common because it is the most robust, so Liao (1994) called this the "median-to-median" estimator of stream life.

The cumulative equivalence point estimator is:

$$\hat{T} = F_D^{-1}(q) - F_L^{-1}(q) \tag{3.11}$$

where $F_D^{-1}(q)$ and $F_L^{-1}(q)$ denote the times at which the cumulative dead and live count curves are at q, where q is a designated percentile, or point on the curves to be compared. For the median-to-median estimator, q=0.50, so the time at which half the total number of carcasses detected is the time, t, that satisfies:

$$F_D(\mathsf{t}) = \frac{\sum_{i=1}^{\mathsf{t}} d_i}{\sum_{i=1}^{m} d_i} = 0.50 \tag{3.12}$$

Similarly, the time at which half the total number of live fish are detected is the time that satisfies:

$$F_L(\mathsf{t}) = \frac{\sum_{i=1}^{\mathsf{t}} l_i}{\sum_{i=1}^{m} l_i} = 0.50 \tag{3.13}$$

The times at which half (or the desired quantile) of either live fish or carcasses may not be unique. The estimated time at which half of the carcasses are recovered may occur between two surveys and the cumulative curve at $q = 0.50$ would be horizontal. In that case, the midpoint between the two surveys is taken as the estimate (Liao, 1994, page 111).

Liao showed that even under ideal conditions of detection, the median-to-median estimator underestimated true stream life (Liao, 1994, Table 5.1). We propose that the negative bias is due to the fact that the cumulative live curve is actually a cumulative fish-days curve, since the live fish may be counted multiple times until they die. Hence, the median-to-median estimator uses the difference between when half of the total fish-days have occurred and when half the carcasses have been found to estimate average stream life. The median-to-median estimator assumes that the majority of the fish present at the time of the median abundance have just arrived so that the time it takes them to die and be found dead is, on average, one stream life. It is more likely that the fish present during the median are themselves half-way through their lives so that the time from median abundance to detection as carcasses is less than the average stream life.

The Median-to-Median Estimator equates the same point on two different types of curves. If the carcasses were not removed from the stream, but were left in the stream and had the possibility of being counted again, then the live and dead fish counts could be compared as the same kinds of counts. However, this comparison would require that fish mortality rate (stream life) be the same as carcass loss rate (retention time) (A-13). In other words, the distribution of live fish abundance would have to be the same shape as the distribution of carcass abundance. There is no biological reason for death rate and decay/loss rate to be the same.

No variance estimate for the Equivalence Points Estimator has ever been derived. Any attempt to estimate the variance would only consider sampling variance, not the bias. Because the method is biased, a better description of the error in sampling is the mean squared error (see Section 3.4.1).

3.4.5 Expected-to-Expected (Liao 1994) Method

One minor difficulty with the Peak-to-Peak and Median-to-Median Estimators is that the times of the peaks or medians may not be unique. Liao (1994) proposed an alternative stream life estimator derived from periodic live and dead counts, which he call the 'expected-to-expected' estimator. The Expected-to-Expected Estimator is the area between the standardized arrival and death curves, and always provides a unique estimate.

The cumulative carcass curve, $F_D(t)$, is generated from periodic surveys during which each carcass is removed after detection to avoid multiple counting (A-7). A cumulative curve can be plotted by adding each subsequent survey total to the previous count. Usually the probability of detection is assumed to be unity by default, $P_D = 1$, because no estimate of carcass detection probability is made (A-14). If there is an estimate of P_D a much simpler estimate of escapement can be made using the sum of the corrected carcass counts (see Chapter 8) (A-12).

The cumulative live curve can be constructed from passage counts made at weirs or with sonar (A-15). However, if there are reliable passage counts then a much more accurate and precise estimate of escapement is immediately available and there is no need for an estimate of stream life. See Chapter 5 for further details.

Liao (1994) constructed a cumulative arrival curve from the periodic counts of live and dead fish. Assuming that all fish remain in the study area once they have arrived (A-2), then the number of individuals arriving between survey occasions *i*-1 and *i*, is equal to the difference between the number of live individuals from occasion *i* to *i*-1, plus the number of mortalities occurring between the surveys:

$$B_i = D_i + L_i - L_{i-1} \tag{3.14}$$

To compare the arrival and death curves, they need to be standardized by the total escapement, which requires an estimate of escapement, $\hat{N}$ (A-16). The standardized arrival curve Liao (1994) is :

$$\hat{F}_B(i) = \sum_{k=1}^{i} \frac{\hat{B}_k}{\hat{N}}, \tag{3.15}$$

and the standardized mortality curve Liao (1994) is :

$$\hat{F}_D(i) = \sum_{k=1}^{i} \frac{\hat{D}_k}{\hat{N}}. \tag{3.16}$$

Note that we have dropped the t from the subscript so that i indicates the time of the ith survey. The area between these curves is $\hat{T}$ and is given by:

$$\hat{T} = \frac{1}{2} \sum_{i=1}^{M} (t_i - t_{i-1}) \left[\hat{F}_A(i) + F_A(\widehat{i-1}) + \hat{F}_D(i) + F_D(\widehat{i-1}) \right]. \tag{3.17}$$

The expected value of the Liao (1994) estimator is:

$$E[\hat{T}] = E\left[\frac{\hat{F}}{\hat{N}}\right] = T\left[1 + \frac{V(\hat{N})}{E^2(\hat{N})}\right] = T\left[1 + CV^2(\hat{N})\right] \tag{3.18}$$

so $\hat{T}$ is positively biased by a factor of $\left[1 + CV^2(\hat{N})\right]$. If the variance of the escapement estimate is estimable (A-17), Equation (3.18) can be adjusted so that the unbiased estimate of stream life is:

$$\hat{T*} = \hat{T}\left[1 + CV^2(\hat{N})\right]^{-1} \tag{3.19}$$

Liao (1994) approximated the variance of $\hat{T}$ using the delta method as:

$$V(\hat{T}) = T^2\left[CV^2(\hat{F}) + CV^2(\hat{N})\right] \tag{3.20}$$

where $\hat{F}$ is the estimator of fish days using the trapezoidal method. The variance of $\hat{T}$ in Equation 3.20 is estimated by:

$$\hat{V}(\hat{T}) = \hat{T}^2\left[\frac{\hat{V}(\hat{F})}{\hat{F}^2} + \frac{\hat{V}(\hat{N})}{\hat{N}^2}\right] \tag{3.21}$$

While the Expected-to-Expected Method can provide a unique and unbiased estimate of stream life, $\hat{T}^*$, it requires an estimate of total escapement (A-16). If an estimate of total escapement is available, then there would be no need for an AUC estimate of escapement, and no need to estimate average stream life.

3.4.6 Inverted Method

The AUC formula for escapement in Equation 2.1 can be solved for average stream life instead of escapement:

$$\text{Average Stream Life} = \frac{\text{Fish-days}}{\text{Escapement}} \tag{3.22}$$

Svoboda and Harrington-Tweit (1983) divided total fish-days (estimated using the Ames method) by an estimate of escapement (carcass mark and recovery technique) to estimate the average stream life of chum salmon in Yelm Creek, WA (cited in Perrin and Irvine, 1990. Bocking et al. (1988) estimated fish-days for French Creek on Vancouver Island, B.C. using a mortality model and divided fish-days by an independent tag-recapture estimate of escapement. Since fish-days was estimated two different ways, Perrin and Irvine (1990) considered these two approaches to be separate techniques (RT4 and RT6 respectively), but the theory behind both methods is the same. Any unbiased estimate of escapement (A-16) can be used in conjunction with any unbiased estimate of fish-days (A-18) to acquire an estimate of stream life, as long as both estimates pertain to the same population (A-19).

The Inverted Estimator is:

$$\hat{T} = \frac{\hat{F}}{\hat{N}}, \tag{3.23}$$

which is unbiased to the first and second terms of a Taylor Series Expansion and positively biased to higher orders. The major drawback of the Inverted Estimator is that it requires an estimate of escapement. If $\hat{N}$ was known, then there would be no need to estimate stream life and no use for the Inverted Estimator.

The variance of the average stream life estimate is approximated using the delta method and is estimated by:

$$\hat{V}(\hat{T}) = \hat{V}(\hat{F})\left(\frac{1}{\hat{N}}\right)^2 + \hat{V}(\hat{N})\left(\frac{\hat{F}}{\hat{N}^2}\right)^2 + 2Cov(\hat{F},\hat{N})\left(\frac{1}{\hat{N}}\right)\left(\frac{\hat{F}}{\hat{N}^2}\right), \tag{3.24}$$

so estimates of the variability of escapement (A-17) and fish-days (A-20) are required. If the estimates of fish-days, $\hat{F}$, and total escapement, $\hat{N}$, are made independently (A-21), the variance estimator simplifies to:

$$\begin{aligned} \hat{V}(\hat{T}) &= \hat{V}(\hat{F})\left(\tfrac{1}{\hat{N}}\right)^2 + \hat{V}(\hat{N})\left(\tfrac{\hat{F}}{\hat{N}^2}\right)^2 \\ &= \left(\tfrac{\hat{F}}{\hat{N}}\right)^2\left[\tfrac{\hat{V}(\hat{F})}{\hat{F}^2} + \tfrac{\hat{V}(\hat{N})}{\hat{N}^2}\right] \\ &= \hat{T}^2[\hat{CV}^2(\hat{F}) + \hat{CV}^2(\hat{N})]. \end{aligned} \tag{3.25}$$

3.4.7 Tag Depletion Curve Method (Bocking et al., 1988)

This method parallels that of the Inverted Estimator Method, but focuses on a tagged subset of the total fish. Fish are tagged and released into the study area and recovered or resighted periodically until no more live fish are detected. The periodic counts are plotted to make a 'tag depletion curve' and the area under this curve is total tag-days

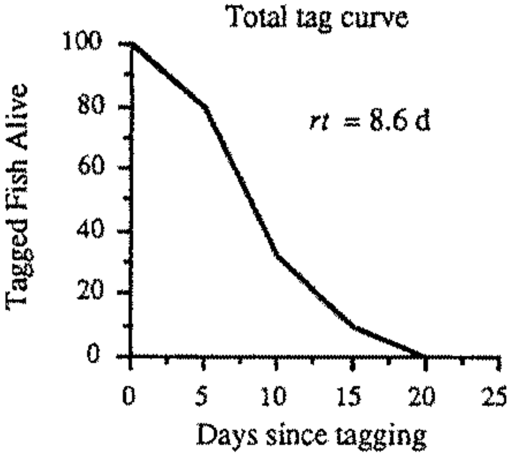

Figure 3.1: Hypothetical tag depletion curve generated from tagging 100 fish. The area under the curve divided by 100 fish gives the estimate of stream life, denoted *rt*. Taken from Figure 1 of English et al. (1992)

(c.f. total fish-days) as shown in Figure 3.1. The number of tag-days is divided by the known number of tags for an estimate of tagged-life leading Bocking et al. (1988) to call it the "tag depletion method." Perrin and Irvine (1990) referred to this the RT5 method.

Bocking et al. (1988) stratified their study streams into five or six spatial strata (A-22), and sampled two or three units (reaches) in each strata every three to seven days. The counts at each site were adjusted by observer efficiency (A-12), averaged, and expanded for each stratum. The expanded counts from each strata were summed to estimate the total number of tagged fish in the entire stream on the ith survey occasion, $\hat{M}_i$ (English et al., 1992):

$$\hat{M}_i = \sum_{h=1}^{H} \frac{N_h}{n_h} \sum_{j=1}^{n_h} \frac{m_{ihj}}{P_{ihj}}. \tag{3.26}$$

These counts are unbiased for the total number of tags present if the sampled reaches in each stratum are sampled randomly (A-23).

The adjusted counts, $\hat{M}_i$, were plotted against time to create the 'tag-depletion curve.' The area under this curve was calculated, using the trapezoidal method, and then divided by the number of tags used. Thus, the tag-depletion estimator of stream residence time is given by:

$$\hat{T} = \frac{\frac{1}{2}\sum_{i=2}^{k}(t_i - t_{i-1})(\hat{M}_i + \hat{M}_{i-1})}{M}, \tag{3.27}$$

or equivalently

$$\hat{T} = \frac{t_2 - t_1}{2} + \frac{\frac{1}{2}\sum_{i=2}^{k-1}(t_{i+1} - t_{i-1})\hat{M}_i}{M} \tag{3.28}$$

where $M_1 = M$ is the total number of tags released, k is the number of recovery surveys, and $M_k = 0$. Surveys must continue until no live tagged fish are found, which ensures that the last count of tagged fish is zero (A-24). The Tag Depletion Estimate of stream residence time is biased unless resighting surveys are conducted every day of the run (A-25). (The proof parallels that of the Trapezoidal Estimator in Section 2.7.2.) This estimate of stream life is only valid for the entire population of fish if tagging has no effect on the survival of the fish (A-6), if the tagged fish are a random sample from the entire run (A-5), and tag loss is not related to survival. That is, the inference is only valid if every fish has the same probability of being tagged and retaining the tag. Since only those fish arriving at the time of the start of the study are tagged, it must be assumed that stream life remains constant over the run (A-8), although it has been observed in several studies that there is a correlation between arrival date and stream life (Thomason and Jones, 1984; Perrin and Irvine, 1990; Hilborn et al., 1999). If the tag-depletion method is applied early in the run, the stream life estimate is likely to be positively biased. If the tag-depletion method is applied late in the run, the stream life estimate is likely to be negatively biased. Ideally, duration of the run should be stratified with fish sampled during each stratum and the strata estimates weighted by estimates of the numbers of fish arriving during each time stratum.

The variance of the estimator depends on the sampling scheme used to select reaches and days sampled, and is identical to that of the trapezoidal fish-days estimator in Equations (3.29)-(3.32) where L_i is replaced by M_i, and with an extra M^{-2} multiplied throughout. For simple random sampling of dates and reach sections, the estimator of the Tag Depletion Curve Estimator is:

$$\hat{V}(\hat{F}) = \frac{k}{4KM^2}\sum_{i=1}^{K}\left[(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}\frac{N_{hi}}{n_{hi}}\sum_{j=1}^{N_{hi}}\left\{\frac{\hat{V}(\hat{\hat{P}}_i)}{\hat{P}_i^{\,4}}[\hat{M}_i\hat{P}_i(1-\hat{P}_i)+(\hat{M}_i\hat{P}_i)^2]\right\}\right] \tag{3.29}$$

$$+\frac{k}{4KM^2}\sum_{i=1}^{K}\left[(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}\frac{N_{hi}}{n_{hi}}\sum_{j=1}^{N_{hi}}\left\{\frac{\hat{M}_i(1-\hat{P}_i)}{\hat{P}_i}\right\}\right] \tag{3.30}$$

$$+\frac{k}{4KM^2}\sum_{i=2}^{K-1}\left[(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}N_h(N_h-n_h)\frac{s_h^2}{n_h}\right] \tag{3.31}$$

$$+\frac{1}{4M^2}\frac{K^2S_Y^2}{r}\frac{(K-k)}{K} \tag{3.32}$$

where $S_Y^2 = \frac{\sum_{i=1}^{K}(Y_i-\bar{Y})^2}{K-1}$.

*3.4.8 Redd Residence Time (McNeil, 1964*b*)*

The Redd Residence Time (RRT) method uses an estimate of the amount of time a female spends at her redd as a proxy for stream life. McNeil (1964*b*) visually identified female pink salmon and counted the number of days each spent at her redd. Crone and Bond (1974) identified specific coho in Sashin Creek, CA using tags. Shardlow (2004) used a camera mounted above the spawning grounds to record the arrival and abandonment (departure) times of female sockeye in the Okanogan River, CA.

Since McNeil (1964*b*) was only interested in the the number of females on the spawning ground, his methods were appropriate. When used for total stream life, RRT is negatively biased because salmon tend to hold in pools and in other areas of fresh water before and after spawning. Crone and Bond (1974) assumed that the arrival and departure times of female salmon were uniformly distributed between samples (A-26). Crone and Bond (1974) surveyed salmon every day, so effectively they assumed that females arrived at the spawning sites up to half a day before they were first detected ($U(t_{i-1}, t_i)$), and abandoned the sites up to half a day after they were last detected ($U(t_i, t_{i+1})$). Thus Crone and Bond (1974) added two half days (i.e. one day) to each tagged fish's redd time. This correction accounts for error associated with periodic sampling if surveys are conducted daily (A-25), but does not account for holding times. Shardlow (2004) showed that RRT only accounted for 89% of the total fish days for sockeye salmon in the Okanogan River, Canada.

The redd residence time method also assumes that the redd residence time of females is the same as the stream residence time of males (A-27). Stream residence time and redd residence time are two different quantities, and it is unlikely that they are equivalent, especially across sexes. The redd residence time of males is different from the redd residence time of females. Female salmon linger over their redds and guard them until death (Briggs, 1953), whereas males attend ripe females until spawning takes place, then swim off to other still-ripe females.

Because redd residence time is not measuring stream life, RRT should not be used as a proxy for stream life. Any variance estimate would be based on the method used to estimate redd residence time (i.e. $V(RRT)$) and would not incorporate the bias term (i.e. $Bias^2(RRT)$) in the mean squared error.

3.4.9 Cormack (1964) Estimates of Survival Probabilities

The Cormack (1964) Method (Section 3.4.10), the Lady and Skalski (1998) Methods (Sections 3.4.11 and 3.4.13), and the Lady (1996) Methods (Sections 3.4.12 and 3.4.14) all use estimates of survival probabilities to determine death rates and from them, mortality times. This section describes the estimation of period-specific survival probabilities.

In the Cormack (1964) Method, between-period survival probabilities can be estimated from a multiple mark-recapture study using uniquely identifying tags. At the end of the study, there is a record of the capture history of each fish. Capture histories are denoted as a sequence of 1's and 0's corresponding to whether or not the individual was captured or not captured during each survey occasion. Since there are a finite number of mutually exclusive capture histories, the capture histories and the numbers of fish can be modeled according to a multinomial distribution:

$$L \propto \prod_h \theta_h^{u_h} \tag{3.33}$$

where θ_h represents the probability of capture history h and u_h is the number of individuals with capture history h.

Consider θ_{1011}, the probability that an individual was tagged during the first survey, was not seen on the second, then seen during both the third and fourth surveys. The probability of survival from the ith survey to the ith+1 survey is ϕ_i. Conversely, the probability of dying during that same interval is $(1 - \phi_i)$. The probability of detection on the ith survey is P_i, and the probability of escaping detection is (1-P_i). The survival and detection probabilities prior to the final survey occasion are inseparable and are denoted $\lambda = \phi_{k-1} P_k$, because it is impossible to discern if the unseen individuals died, or escaped detection (Cormack, 1964). Thus the probability of capture history '1011' is:

$$\theta_{1011} = \phi_1(1 - P_2)\phi_2 P_3 \phi_3 P_4 = \phi_1(1 - P_2)\phi_2 P_3 \lambda.$$

It is also useful to define χ_i, the probability that an individual seen during survey i, will not be seen thereafter. The probability of capture history 1010 is:

$$\theta_{1010} = \phi_1(1 - P_2)\phi_2 P_3 \chi_3,$$

where $\chi_3 = 1 - \lambda$. The χ_i are related recursively, such that:

$$\chi_i = (1 - \phi_i) + \phi_i \chi_{i+1}(1 - P_{i+1}).$$

Note the typo in Cormack (1964) Equation (1), where ϕ_1 should be ϕ_i.

The basic statistics from the Cormack Method are the numbers of individuals of each capture history, u_h, but they can be simplified into the following minimum sufficient statistics (Cormack, 1964):

m_i	Number of previously marked individuals resighted during the ith survey
c_i	Number of marked individuals seen for the last time during the ith survey
v_i	Number of marked individuals known to be alive after the ith survey

The likelihood of the Cormack (1964) model is:

$$L(\widetilde{P}, \widetilde{\phi} | \widetilde{m}, \widetilde{c}, \widetilde{v}) \propto \prod_{h=1}^{k-1} \phi_i^{v_h} P_{i+1}^{m_{i+1}} (1 - P_{i+1})^{v_i - m_{i+1}} \chi_h^{c_h} \tag{3.34}$$

where $v_{k-1} = m_k$. The likelihood in (3.34) assumes that the probability of capture is independent of previous capture history (A-28), that tagging has no effect on survival (A-6), and that no tags are lost.

Equation (3.34) can be solved to produce the following parameter estimates:

$$\hat{\phi}_i = \frac{\tau_i v_i s_{i+1} - \tau_i m_{i+1} c_{i+1}}{s_i v_i \tau_{i+1}}, \tag{3.35}$$

$$\hat{P}_i = \frac{m_i \tau_{i-1}}{s_{i-1} v_{i-1} \widehat{\phi_{i-1}}}, \tag{3.36}$$

and

$$\hat{\chi}_i = \frac{c_i}{s_i}, \tag{3.37}$$

where s_i is the number of fish captured during the ith survey. Note that we have used slightly different notation than Cormack (1964). We have used τ where Cormack used t and m where Cormack used a.

This method of estimating survival probabilities requires the assumption that all fish are identically and independently distributed with respect to capture history. That is, every fish has the same probability of capture or recapture (A-29) and survival for a given survey interval (A-30), and that the survival and capture of each fish is independent of the survival and capture of all others (A-31). Thus, tagging must have no effect on the survival or recapture of the fish (A-6) and the survival probabilities are independent of stream arrival (A-32). This method also assumes that no tags are lost or overlooked. If tags are lost or overlooked, survival and detection probabilities will be underestimated. Stream life, as estimated by the Cormack (1964) Method (Section 3.4.10), will be overestimated.

The variance of the estimate in Equation 3.35 is derived via the delta method as:

$$V(\hat{\phi}_i) = \frac{\chi_{i+1}(1-P_{i+1})^2\phi_i^2}{(1-\chi_{i+1})E[s_{i+1}]} + \frac{\chi_{i+1}^2 P_{i+1}^2 (1-P_{i+1})\phi_i^3}{(1-\chi_i)(1-\chi_{i+1})E[m_{i+1}]} + \frac{\chi_i\phi_i^2}{(1-\chi_i)E[s_i]}, \tag{3.38}$$

and is estimated by:

$$\hat{V}(\hat{\phi}_i) = \frac{\hat{\chi_{i+1}}(1-\hat{P_{i+1}})^2\hat{\phi}_i^2}{(1-\hat{\chi_{i+1}})s_{i+1}} + \frac{\hat{\chi_{i+1}}^2\hat{P_{i+1}}^2(1-\hat{P_{i+1}})\hat{\phi_i}^3}{(1-\hat{\chi}_i)(1-\hat{\chi_{i+1}})m_{i+1}} + \frac{\hat{\chi}_i\hat{\phi_i}^2}{(1-\hat{\chi}_i)s_i}, \tag{3.39}$$

The covariance of survival probability estimates from successive surveys is:

$$Cov(\hat{\phi}_i, \hat{\phi}_{i+1}) = -\frac{\phi_i\phi_{i+1}\chi_{i+1}(1-P_{i+1})}{(1-\chi_{i+1})E[s_i]}, \tag{3.40}$$

and is estimated by:

$$\hat{Cov}(\hat{\phi}_i, \hat{\phi_{i+1}}) = -\frac{\hat{\phi}_i\hat{\phi_{i+1}}\hat{\chi}_{i+1}(1-\hat{P_{i+1}})}{(1-\hat{\chi}_{i+1})s_i}, \tag{3.41}$$

For survey occasions separated by more than one occasion, the covariance is zero:

$$Cov(\phi_i, \phi_j) = 0 \quad \text{for} \quad |i-j| > 1. \tag{3.42}$$

3.4.10 Cormack (1964) Method

Cormack (1964) used annual survival probabilities to determine the life expectancy of fulmar petrels (*Fulmarus glacialis*) on an island in Orkney, but the techniques can be adjusted for any population of individuals that can be marked with uniquely identifying tags, and can be used to estimate salmon stream life.

Cormack (1964) assumed that mortality occurring within the interval between two surveys followed an exponential distribution (A-33), such that the probability distribution function of time of death on the interval from t_i to t_{i+1}, x_{ij}, conditional on being alive at the start of the interval is:

$$f(x_{ij}) = r_i e^{-r_i x_{ij}} \qquad t_i \leq x_{ij} < t_{i+1} \tag{3.43}$$

where r_i is the instantaneous death rate during the interval from t_i to t_{i+1}. Note, the expected value of the time to death for the jth individual, x_{ij}, is:

$$E(x_{ij}) = \frac{1}{r_i}.$$

The probability of survival during an interval is one minus the probability of death, which is the integral of Equation (3.43) over the time span of the interval:

$$\begin{aligned}
\phi_i &= 1 - \textstyle\int_{t_i}^{t_{i+1}} r_i e^{-r_i(x_i - t_i)} dx \\
&= 1 - \left[-e^{-r_i(x_i - t_i)}\right]_{t_i}^{ti+1} \\
&= 1 - \left[-e^{-r_i(t_{i+1} - t_i)} - -e^{-r_i(t_i - t_i)}\right] \qquad (3.44) \\
&= 1 - 1 + e^{-r_i(t_{i+1} - t_i)} \\
\phi_i &= e^{-r_i(t_{i+1} - t_i)}
\end{aligned}$$

Solving Equation 3.44 for death rate, r_i, yields:

$$\hat{r_i} = \frac{-\log(\phi_i)}{(t_{i+1} - t_i)} \tag{3.45}$$

As noted above, the expected time until death is $\frac{1}{r_i}$, assuming that death times follow an exponential distribution, so the Cormack (1964) estimator of life expectancy over the interval i to i+1 is:

$$E(x_i) = \frac{1}{r_i} = \frac{-(t_{i+1} - t_i)}{\log(\phi_i)} \tag{3.46}$$

Cormack (1964) performed his surveys annually, so all of the time intervals between surveys were unity: $t_{i+1} - t_i = 1$ for all $i = 1, 2, \ldots, k$. He also assumed that life expectancy was independent of age, so that an individual entering the study are after surveys started had the same life expectancy as an individual who had been in the study area since the first survey (A-32). Thus, each estimate of conditional survival probability, ϕ_i, could be thought of as an observation from a meta-population of ϕ's. An unbiased estimate of the underlying survival probability, Φ, would be the arithmetic mean of the observed ϕ_i's:

$$\bar{\phi} = \frac{1}{k-2}\sum_{i=1}^{k-2}\phi_i \tag{3.47}$$

Since the lifetime estimator uses the natural logarithm of survival, Cormack (1964) used the geometric mean of the survival probabilities:

$$\Phi = \left(\prod_{i=1}^{k-2}\phi_i\right)^{1/(k-2)}, \tag{3.48}$$

because the mean of the natural log of the survival is the log of the geometric mean of survival:

$$\begin{aligned} \log(\Phi) &= \log\left((\textstyle\prod_{i=1}^{k-2}\phi_i)^{1/(k-2)}\right) \\ &= \tfrac{1}{k-2}\log(\textstyle\prod_{i=1}^{k-2}\phi_i) \\ &= \tfrac{1}{k-2}\textstyle\sum_{i=1}^{k-2}\log(\phi_i) \end{aligned} \tag{3.49}$$

Thus, the estimator for life expectancy given by Cormack (1964) was:

$$\hat{T} = \frac{-1}{\log(\Phi)} \tag{3.50}$$

Equation (3.50) will be a biased estimate of life expectancy for studies that have unequal intervals between surveys. It also requires the assumption that the survival probabilities for all periods are the same (A-34). However, it has been found that stream life decreases over the course of the run. Therefore pooling the survival estimates into one value is inappropriate and could lead to bias with an unknown direction and magnitude. The estimator given in Equation (3.46) is more applicable to salmon escapements, where the time between surveys may vary from one to seven days or more, and survival for each period is estimated separately. However, Equation (3.46) only provides the life expectancy for a fish from the time of the ith survey. Many of the fish will have arrived prior to the ith survey, so Equation (3.46) also underestimates stream life.

Using the delta method, and assuming equal times between surveys (A-35), Cormack (1964) determined the variance of the life expectancy estimator to be:

$$\begin{aligned} V(\hat{T}) &= \tfrac{1}{(\log\Phi)^4(k-2)^2} \times \textstyle\sum_{i=1}^{k-2} \left\{ \tfrac{\chi_{i+1}P_{i+1}^2}{E(\tau_{i+1})} + \tfrac{\chi_{i+1}^2 P_{i+1}^2(1-P_{i+1})\phi_i}{(1-\chi_i)(1-\chi_{i+1})E(m_{i+1})} \right\} \\ &+ \tfrac{\chi_i}{E(\tau_1)} + \tfrac{\chi_{n-1}(1-2P_{k-1})}{E(\tau_{k-1})} \end{aligned} \tag{3.51}$$

which is estimated by substituting the parameters with their estimators and the expected values with their respective statistics.

3.4.11 Lady and Skalski (1998) Live Non-parametric Method

The Lady and Skalski (1998) method of estimating stream life uses estimates of period-specific survival probabilities to generate estimates of salmon life expectancy. The survival probabilities are estimated from the same tagging-resighting procedure and likelihood as in the Cormack (1964) model with only one initial tagging event at the start of the study. Consequently, the inference on life expectancy is restricted to those individuals who entered the study area at the time of initial tagging. The entire procedure can be repeated over the course of a run to obtain entry-time-specific estimates of stream life, which can be weighted by the proportion of the run arriving at each of the survey times to estimate the average stream life for the entire river.

The likelihood of survival and detection given by Lady and Skalski (1998) is:

$$L(\widetilde{\phi}, \widetilde{P}, \lambda | \widetilde{a}, \widetilde{c}) \propto \left(\prod_{i=2}^{k-2} \phi_i^{v_i} \right) \left(\prod_{i=2}^{k-1} P_i^{m_i} (1-P_i)^{v_{i-1}-m_i} \right) \left(\prod_{i=1}^{k-1} \chi_i^{c_i} \right) \lambda^{v_{k-1}} \tag{3.52}$$

where $v_i = M - \sum_{j=1}^{i} c_j$ is the number of marked individuals known to be alive at survey occasion $i+1$, and χ_i is the probability that an individual alive at the time of the ith survey will not be detected again.

Lady and Skalski (1998) assumed that the distribution of deaths between surveys was uniform (A-26), so that the expected time of death during an interval (given that death did occur during the interval) was the midpoint of the interval. The average stream life is the expected stream life of an individual, which is the sum over all periods of 'the probability of death' in a period, times 'the expected time of death' in the period (the midpoint). The probability of death on an interval includes the probabilities of surviving up until the start of an interval, because all fish were tagged at only the first survey time. The resulting Lady and Skalski (1998) estimator of stream life is:

$$\hat{T} = \frac{1}{2}\sum_{i=1}^{k-2}\left[(t_i + t_{i+1})(1-\hat{\phi}_i)\prod_{j=1}^{i-1}\hat{\phi}_j\right] + \frac{1}{2}(t_{k-1}+t_k)\prod_{j=1}^{k-2}\hat{\phi}_j \tag{3.53}$$

where t_i and t_{i+1} are survey times relative to the first survey so that $t_1 = 0$. If the study is carried out until all the fish are dead (A-24), the final survival probability is zero, $\phi_k = 0$, and Equation (3.53) simplifies to:

$$\hat{T} = \frac{1}{2}\sum_{i=1}^{k-1}\left[(t_i + t_{i+1})(1-\hat{\phi}_i)\prod_{j=1}^{i-1}\hat{\phi}_j\right]. \tag{3.54}$$

Inference from the estimators in Equations (3.53) and (3.54) apply only to individuals who entered the stream at the time of the initial tagging. To estimate the average stream life of all individuals over the entire study, tagging must occur multiple times. The expected stream life for individuals entering at each tagging event can be estimated as above and the overall average residence time calculated as a weighted mean of each estimate, $\hat{T}_j$. The weights, w_j, are the proportions of the overall population that entered at each new tagging event. If, for example, additional fish were tagged as they entered the stream during each of Z time strata, the estimate would be:

$$\hat{T} = \sum_{j=1}^{Z}\hat{T}_j w_j \tag{3.55}$$

where $\hat{T}_j$ is the stream life estimate for the fish entering the study during the jth of Z time strata. The weights, w_j, can be obtained by tracking the proportion of the run tagged on a tagging day. If the salmon are captured in a weir at the entrance to the study area (A-15), the weights would be the proportion of fish entering the stream during the tagging period:

$$w_j = \frac{N_j}{N}$$

However, having a weir with which to count proportions of fish would negate the need for a stream life estimator. Instead, the weights can be constructed by assuming a constant catch rate P (A-36) and an index of escapement abundance. The number caught and marked is the product of the number arriving during that period, N_j, and the capture probability:

$$M_j = N_j P \tag{3.56}$$

The total number marked over the study is:

$$\sum_{j=1}^{Z} M_j = \sum_{j=1}^{Z} N_j P \tag{3.57}$$

so the weights are the number marked from each initial release period divided by the total:

$$w_j = \frac{M_j}{\sum_{j=1}^{Z} M_j} = \frac{N_j P}{\sum_{j=1}^{Z} N_j P} = \frac{N_j}{N} \tag{3.58}$$

provided that the capture probability, P, remained the same over all tagging events (A-36).

Lady and Skalski (1998) found it convenient to write the stream life estimator, $\hat{T}$, in matrix notation. Including the weights, the overall estimate of stream life is:

$$\hat{T} = \sum_{j=1}^{Z} \hat{T}_j w_j = \sum_{j=1}^{Z} \frac{w_j}{2} \tilde{\tau}' \tilde{\delta} \tag{3.59}$$

where

$$\begin{aligned}
\tilde{\tau}' &= [t_1 + t_2 \quad t_2 + t_3 \quad ... \quad t_{k-1} + t_k] \\
\tilde{\delta}' &= [\delta_1 \quad \delta_2 \quad ... \quad \delta_{k-2} \quad \delta k - 1] \\
\delta_1 &= 1 - \hat{\phi}_1 \\
\delta_2 &= (1 - \hat{\phi}_2)\hat{\phi}_1 \\
\delta_{k-2} &= (1 - \hat{\phi}_{k-2}) \textstyle\prod_{i=1}^{k-3} \hat{\phi}_i \\
\delta_{k-1} &= \textstyle\prod_{i=1}^{k-2} \hat{\phi}_i.
\end{aligned}$$

The variance of the overall stream life estimate is:

$$V(\hat{T}) = V\left(\sum_{i=1}^{k-1}\hat{T}_j w_j\right) = \sum_{j=1}^{k-1}\frac{(w_j)^2}{4}\tilde{\tau}'V(\tilde{\delta})\tilde{\tau} \tag{3.60}$$

The elements of $V(\tilde{\delta})$ are estimated via the delta method as:

$$\hat{Cov}(\delta_i, \delta_j) = \sum_{x=1}^{k-1}\sum_{y=1}^{k-1}\left(\frac{\partial\delta_i}{\partial\hat{\phi}_x}\right)\left(\frac{\partial\delta_j}{\partial\hat{\phi}_y}\right)\hat{Cov}(\hat{\phi}_x, \hat{\phi}_y) \tag{3.61}$$

for i, j = 1,2,...,k-1.

The partial derivatives are:

$$\begin{array}{rcl}
\frac{\partial\delta_i}{\partial\hat{\phi}_j} & = & 0 \qquad \text{for } j > i \\
\frac{\partial\delta_i}{\partial\hat{\phi}_i} & = & -\prod_{j=1}^{i-1}\hat{\phi}_j \qquad \text{for } i = 1, ..., k-2 \\
\frac{\partial\delta_i}{\partial\hat{\phi}_j} & = & (1-\hat{\phi}_i)\left(\prod_{x=1}^{j-1}\hat{\phi}_x\right)\left(\prod_{y=j+1}^{i-1}\hat{\phi}_y\right) \qquad \text{for } i = 1, ..., k-2 \quad \text{and} \quad j < i \\
\frac{\partial\delta_{k-1}}{\partial\hat{\phi}_i} & = & \left(\prod_{j=1}^{i-1}\hat{\phi}_j\right)\left(\prod_{j=i+1}^{k-2}\hat{\phi}_j\right) \qquad \text{for } i = 1, ..., k-2
\end{array} \tag{3.62}$$

where $\prod_{i=a}^{b} f = 1$ for $a > b$ and any function f.

The variances of the survival probabilities are very similar to those of Cormack (1964) in Equations 3.38 and (3.39):

$$\begin{aligned}
\hat{V}(\hat{\phi}_i|\phi_i) &= \frac{\hat{\chi}_{i+1}(1-\hat{P}_{i+1})^2\hat{\phi}_i^2}{(1-\hat{\chi}_{i+1})\hat{E}(m_{i+1})} \\
&+ \frac{\hat{\chi}_{i+1}^2\hat{P}_{i+1}^2(1-\hat{P}_{i+1})^2\hat{\phi}_i^3}{(1-\hat{\chi}_i)(1-\hat{\chi}_{i+1})\hat{E}(m_{i+1})} \\
&+ \frac{\hat{\chi}_i\hat{\phi}_i^2}{(1-\hat{\chi}_i)\hat{E}(m_i)}
\end{aligned} \tag{3.63}$$

for i = 2, ... ,k-2, where

$$\hat{E}(m_i) = M\left(\prod_{j=1}^{i-1}\hat{\phi}_j\right)\hat{P}_i$$

is the estimate of the expected number detected at the ith survey occasion, and

$$\hat{P}_2 = \frac{m_2}{M\hat{\phi}_1} \tag{3.64}$$

$$
\begin{aligned}
\hat{P}_i &= \frac{m_i(m_{i-1} - c_{i-1})}{m_{i-1}v_{i-1}\hat{\phi}_{i-1}} \qquad \text{for } i = 2, ..., k-1. & (3.65) \\
\hat{\chi}_1 &= \frac{c_1}{M} & (3.66) \\
\hat{\chi}_i &= \frac{c_i}{m_i} \qquad \text{for } i = 1, ..., k-1. & (3.67)
\end{aligned}
$$

The covariances of the survival probabilities are:

$$
\hat{Cov}(\hat{\phi}_i, \hat{\phi}_{i+1}) = \frac{\hat{\phi}_i \hat{\phi}_{i+1} \hat{\chi}_{i+1}(1 - \hat{P}_{i+1})}{(1 - \hat{\chi}_{i+1})\hat{E}(m_{i+1})} \qquad \text{for } i = 1, ..., k-3 \tag{3.68}
$$

and

$$
\hat{Cov}(\hat{\phi}_i, \hat{\phi}_{i+1}) = 0 \qquad \text{for } |i - j| > 1. \tag{3.69}
$$

Again, note the similarity of the covariances with those derived by Cormack (1964) shown in Equations (3.40) and (3.41).

For an example of the Live Non-parametric Method, please see (Lady, 1996, Chapter 4).

3.4.12 Lady (1996) Live and Dead Non-parametric Method

When conducting periodic surveys of study sites, records can be kept of dead individuals as well as live ones. Lady (1996 page 28) developed a likelihood model similar to the Live Non-parametric Model that included carcass detections (Lady and Skalski, 1998). As in the Live Non-parametric Model, the inference for each tagging event only applies to fish entering the system during the tagging event, so the technique must be applied repeatedly over the course of the run in order to generate a valid estimate of stream life.

The likelihood is the same as that in the Cormack-Jolly-Seber method in Section 3.4.9 given by Equation 3.33, only the capture histories now include the possibility of individuals being found dead. This is denoted as a “2” in the capture history. The tag is recovered from the carcass and the carcass is removed from the stream or severed in half to preclude it from being recounted (A-7). Thus scores for all survey occasions following a “2” in the capture history are “0.” Note also that all capture histories start with “1,” because a fish must be captured the be marked. For example, 1120 is the capture history of an individual tagged at the first sample seen alive at the second, and found dead at the third. The probability of having capture history “1120” is:

$$\theta_{1120} = \phi_1 P_2 (1 - \phi_2) D_3,$$

where D_i is the probability that a carcass (dead fish) present at the time of the ith survey will be detected at the time of that survey. The probability of being marked in the first survey, not being seen in the second and being found dead in the third is:

$$\theta_{1020} = (1 - \phi_1)(1 - D_2) D_3 + \phi_1 (1 - P_2)(1 - \phi_2) D_3.$$

In this model, not all fish must be observed dead. It is possible that a fish may be detected alive on the last survey. For example, the probability of capture history “1001” where a fish is tagged, not seen for two surveys and then detected alive on the fourth survey is:

$$\theta_{1001} = \phi_1 (1 - P_2) \phi_2 (1 - P_3) \phi_3 P_4.$$

It is helpful to define χ_i^{**} as the probability that an individual alive at the time of the ith survey will not be detected after the ith survey, either alive or dead. Thus:

$$\chi_i^{**} = (1 - \phi_i) \prod_{j=i+1}^{k} (1 - D_j) + \phi_i (1 - P_{i+1}) \chi_{i+1}^{**} \tag{3.70}$$

and $\chi_k^{**} = 1$.

It is also helpful to define $\phi_{i,j}^*$ as the probability that an individual alive at the time of the ith survey will be detected dead on the jth survey and not detected on any survey in between:

$$\phi_{i,j}^* = (1 - \phi_i) \prod_{j=i+1}^{k} (1 - D_m) + \phi_i (1 - P_{i+1}) \phi_{i+1,j}^* \tag{3.71}$$

for $i = 1, ...,$ k-2 and $j > i + 1$. For sequential surveys, $\phi_{i,i+1}^* =$ 1 - ϕ_i for $i = 1$,...,k-1.

Lady (1996, page 37) gave the likelihood in terms of the sufficient statistics as:

$$\begin{aligned} L(\widetilde{\phi}, \widetilde{P}, \widetilde{D} | \widetilde{m}, \widetilde{c}, \widetilde{f}) \;\propto\; & \left(\prod_{i=2}^{k-1} P_i^{m_i}(1-P_i)^{v_i-m_i}\right) \\ \times\; & \left(\prod_{i=1}^{k-1} \prod_{j=i+1}^{k} (\phi_{i,j}^{*})^{f_{i,j}}\right) \left(\prod_{i=1}^{k-1} (\chi_i^{**})^{c_i} \phi_i^{v_i}\right) \\ \times\; & \left(\prod_{i=2}^{k} D_i^{f_{\bullet j}}\right) P_k^{v_{k-1}} \end{aligned} \tag{3.72}$$

where $f_{i,j}$ is the number of individuals seen alive at the time of the ith survey, found dead at the time of the jth survey (and not detected in between), $f_{\bullet j}$ is the total number of individuals detected dead on the jth survey, and $v_i = M - \sum_{j=1}^{i} c_j - \sum_{j=1}^{i} f_{j\bullet}$ is the number known to be alive at the time of survey $i+1$. Note that we have used m and M where Lady (1996) used a and R, respectively.

The Live and Dead Non-parametric Method requires the same set of assumptions as the Live Non-parametric Method with the additional assumption that all carcasses present during the ith survey have the same probability of detection (A-37). In reality, older carcasses have been decaying longer and have a greater risk of being removed from the study site by scavengers or flooding. These risks are reduced when surveys are conducted on a daily basis.

The parameters, their variances and covariances are estimable for studies with four or more sampling occasions, but only via iterative numerical procedures (Lady, 1996, page 37). Once time-specific estimates of stream residence time, T_j, are obtained, the overall stream life and variance can be calculated as in Section 3.4.11.

For an example of the Live and Dead Non-parametric Method, please see (Lady, 1996, Chapter 4).

3.4.13 Lady and Skalski (1998) Live Parametric Method

The Lady and Skalski (1998) method assumes that there is a process underlying the death times of individuals, and that the period-specific estimates of survival probabilities are observations from the underlying distribution. Lady and Skalski (1998) assumed that mortality follows a 2-parameter Weibull(α,β) distribution (A-38), where

α is the scale parameter and β is the shape parameter. They chose the Weibull distribution, in part, for its increasing hazard function (when $\beta > 1$), which seemed appropriate for modeling individuals who will die by the end of the study (A-7) (Lady and Skalski, 1998). The period-specific survival probabilities of the Live Non-parametric Model (Section 3.4.11) are re-parameterized in terms of the assumed distribution, and the expected stream residence time is the expected value of the assumed distribution.

If death times are distributed according to the Weibull(α,β) distribution (A-38), the probability of an individual surviving from tagging to some time t is

$$\phi(t) = P(\mathsf{t} > t) = e^{-(\alpha t)^{\beta}}.$$

The period-specific survival probabilities can be parameterized in terms of the Weibull distribution as:

$$\phi_i = P(\mathsf{t} > t_{i+1} | \mathsf{t} > t_i) = \frac{P(\mathsf{t} > t_{i+1})}{P(\mathsf{t} > t_i)} = \frac{e^{-(\alpha t_{i+1})^{\beta}}}{e^{-(\alpha t_i)^{\beta}}} = e^{-\alpha^{\beta}(t_{i+1}^{\beta} - t_i^{\beta})} \tag{3.73}$$

Substituting Equation (3.73) into the likelihood from Equation (3.52) for a study with k sampling occasions is:

$$L(\alpha, \beta, \widetilde{P} | \widetilde{a}, \widetilde{c}) \propto \left(\prod_{i=2}^{k-1} P_i^{m_i} (1 - P_i)^{v_i - m_i} e^{-c_i(\alpha t_i)^{\beta}} \right) \left(\prod_{i=1}^{k-1} \chi_i^{c_i} \right) P_R^{v_{k-1}} e^{(-v_{k-1}(\alpha t_k)^{\beta})} \tag{3.74}$$

where $v_i = M - \sum_{j=1}^{i} c_j - \sum_{j=1}^{i} f_{j\bullet}$ is the number known to be alive at survey occasion $i + 1$, $f_{j\bullet}$ is the number detected alive for the last time at survey occasion i and detected dead later, and χ_i is the probability of not being detected again, either alive or dead, after survey occasion i.

The expected stream residence time is the mean of the Weibull(α,β) distribution, which is (Meeker and Escobar, 1998, page 86):

$$T = \frac{1}{\alpha} \Gamma \left(1 + \frac{1}{\beta} \right),$$

and is estimated by:

$$\hat{T} = \frac{1}{\hat{\alpha}} \Gamma \left(1 + \frac{1}{\hat{\beta}} \right), \tag{3.75}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the maximum likelihood estimates derived from Equation (3.74).

The variance of the estimator given in Equation (3.4.13) can be approximated using the delta method (Lady, 1996):

$$V(\hat{T}) = V(\alpha)\left(\frac{\delta T}{\delta \alpha}\right)^2 + V(\beta)\left(\frac{\delta T}{\delta \beta}\right)^2 + 2Cov(\alpha, \beta)\left(\frac{\delta T}{\delta \alpha}\right)\left(\frac{\delta T}{\delta \beta}\right) \tag{3.76}$$

where $\frac{\delta T}{\delta \alpha} = \frac{-1}{\alpha^2}\Gamma\left(1 + \frac{1}{\beta}\right)$ and $\frac{\delta T}{\delta \beta} = \frac{-1}{\alpha\beta^2}\int_0^\infty t^{1/\beta} log(t) e^{-t} dt$.

Equation (3.76) is approximated by substituting the parameters with their MLE's:

$$\begin{aligned} \hat{V}(\hat{T}) &= \hat{V}(\hat{\alpha})\left(\frac{-1}{\hat{\alpha^2}}\Gamma\left[1 + \frac{1}{\hat{\beta}}\right]\right)^2 \\ &+ \hat{V}(\hat{\beta})\left(\frac{-1}{\hat{\alpha}\hat{\beta}^2}\int_0^\infty t^{1/\hat{\beta}}\log(t) e^{-t} dt\right)^2 \\ &+ 2\hat{Cov}(\hat{\alpha}, \hat{\beta})\left(\frac{-1}{\hat{\alpha^2}}\Gamma\left[1 + \frac{1}{\hat{\beta}}\right]\right)\left(\frac{-1}{\hat{\alpha}\hat{\beta}^2}\int_0^\infty t^{1/\hat{\beta}}\log(t) e^{-t} dt\right) \end{aligned} \tag{3.77}$$

where $\hat{V}(\hat{\alpha})$, $\hat{V}(\hat{\beta})$, and $\hat{Cov}(\hat{\alpha}, \hat{\beta})$ are derived from the information matrix of the likelihood in Equation (3.74).

For an example of the Live Parametric Method, please see (Lady, 1996, Chapter 4).

3.4.14 Lady (1996) Live and Dead Parametric Method

The Live and Dead Parametric technique is a combination of the Lady (1996) Live and Dead Non-parametric method (Section 3.4.12) and the Lady and Skalski (1998) Live Parametric method (Section 3.4.13). It relies on mark-recapture estimates of survival probabilities, based on counts of both live and dead individuals, which have been re-parameterized according to the parameters of the distribution which has been chosen to model the mortality process of individuals. Lady (1996) chose to re-parameterize according to a Weibull(α,β) distribution because of the Weibull's decreasing hazard function (A-38) and closed-form mean.

The likelihood for the live and dead parametric model is given by (Lady, 1996, page 41):

$$
\begin{aligned}
L(\alpha, \beta, \widetilde{P}, \widetilde{D} | \widetilde{m}, \widetilde{c}, \widetilde{f}) &\propto \left(\prod_{i=2}^{k-1} e^{(c_i + f_{i\bullet})(\alpha t_i)^{\beta}} P_i^{m_i} (1 - P_i)^{v_i - m_i} \right) \\
&\times \left(\prod_{i=1}^{k-1} \prod_{j=i+1}^{k} (\phi^*_{(W)(i,j)})^{f_{i,j}} \right) \left(\prod_{i=1}^{k-1} (\chi^*_{(W)i})^{c_i} \right) \left(\prod_{i=2}^{K} D_i^{f_{\bullet i}} \right) \\
&\times e^{-v_{k-1}(\alpha t_k)^{\beta}} \times P_r^{v_{k-1}}
\end{aligned}
\tag{3.78}
$$

where $\chi^*_{(W)i}$ is the probability that a tagged fish alive at sampling occasion i will not be detected again, parameterized using the Weibull distribution, and $\phi^*_{(W)(i,j)}$ is the probability that an individual alive at sampling occasion i will be detected dead at sampling occasion j and not in between, parameterized using the Weibull distribution.

As with the Lady (1996) Live and Dead Non-parametric Method (see Section 3.4.12), the MLE's and variances can be estimated but for studies with three or more sampling occasions, using numerical iterative procedures (Lady, 1996, page 42).

For an example of the Live and Dead Parametric Method, please see (Lady, 1996, Chapter 4).

3.4.15 Timer Tags, Non-parametric (Shardlow et al., 2007)

Shardlow et al. (2007) developed an electronic tag with two tilt switches, a timer and a battery. The tags can be applied to salmon entering the study area and recovered from carcasses. Rather than recording the time from tagging to recovery, the timer in the tag records the time from tagging to death, i.e. when the tagged fish assumes a preset body angle for longer than a preset time (A-39). For example, Shardlow et al. (2007) set the timers to record death if the tag detected angles between 45 and 315 for longer than 3 minutes.

Both parametric and non-parametric estimators can be derived from the timer tag information. For the non-parametric estimator, individuals tagged must be a representative sample from the total escapement. If tagging has no influence on stream life (A-7) and every fish has an equal probability of being tagged (A-5) then the tagged fish constitute a random sample from the run and a valid estimate of average stream life for the population is:

$$\hat{T} = \frac{1}{m} \sum_{j=1}^{m} T_j \tag{3.79}$$

which is unbiased for average stream life. The variance is (Cochran 1977 pg. 23):

$$V(\hat{T}) = \frac{\sum_{j=1}^{N} (T_j - \bar{T})^2}{m(N-1)} \left(1 - \frac{m}{N}\right) \tag{3.80}$$

and is estimated by:

$$\hat{V}(\hat{T}) = \frac{\sum_{j=1}^{m} (T_j - \hat{T})^2}{m(m-1)} \left(1 - \frac{m}{N}\right) \tag{3.81}$$

Often, the sampling fraction, $\left(1 - \frac{m}{N}\right)$, is unknown and a completely random sample of the individuals is logistically unfeasible. In that case, the run is divided into Z time strata and tags are applied during each stratum. The estimates from each stratum are weighted by the proportion of fish entering the stream during that stratum. The estimate of stream life for individuals entering the study area in stratum i is:

$$\bar{T}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} T_{ij} \tag{3.82}$$

where T_{ij} is the stream life time of fish j tagged during the ith tagging stratum.

The variance on the time-specific estimates is the sampling variance:

$$V(\bar{T}_i) = \frac{\sum_{j=1}^{N_i} (T_{ij} - T_i)^2}{m_i(N_i - 1)}, \tag{3.83}$$

which is estimated by:

$$\hat{V}(\bar{T}_i) = \frac{\sum_{j=1}^{m} (T_{ij} - \bar{T}_i)^2}{m_i(m_i - 1)}. \tag{3.84}$$

The time specific estimates are weighted by the proportion of the escapement arriving about the same time as the tagged individuals:

$$\hat{T} = \sum_{i=1}^{k-1} \bar{T}_i w_i \tag{3.85}$$

where the weights, w_i are determined as described in the Lady and Skalski (1998) Live Non-Parametric Methods of Section 3.4.11.

Assuming that, given entry time, stream life times of individual fish are independent of one another (A-31), the variance of the weighted stream life estimate is:

$$V(\hat{T}) = \sum_{j=1}^{k-1} (w_i)^2 V(\bar{T}_i) \tag{3.86}$$

and is estimated by:

$$\hat{V}(\hat{T}) = \sum_{j=1}^{k-1} (w_i)^2 \hat{V}(\bar{T}_i) \tag{3.87}$$

where $\hat{V}(\bar{T}_i)$ is given by Equation 3.84.

Non-parametric Timer Tag Examples

The data for these examples comes from Table 38 of Killick (1955) and are reproduced in Table 3.1. For these examples, we assume that values in the T column were obtained from timer tags and are accurate measures of the time from initial stream entry to death.

The mean estimators are calculated using Equation (3.2) and are shown in the "Mean" row of Table 3.1. For the earliest group of female sockeye (A),

$$\begin{aligned}\bar{T}(A) &= \tfrac{1}{156}[6 + 7 + 8(6) + 9(6) + 10(15) + 11(9) \\ &\quad +12(20) + 13(23) + 14(21) + 15(12) + 16(13) \\ &\quad +17(8) + 18(8) + 19(8) + 20(2) + 21 + 22] \\ &= \tfrac{1}{156}[2112] \\ &= 13.59 \quad \text{days,}\end{aligned} \tag{3.88}$$

with a variance estimate of:

$$\hat{V}(\bar{T}(A)) = \frac{s_A^2}{156} = \frac{1561.744}{(156)(155)} = 0.0646 \tag{3.89}$$

For the middle release group (B),

$$\bar{T}(B) = 11.81 \quad \text{days.} \tag{3.90}$$

Table 3.1: Frequency distribution of the number of days from tagging to recovery for early (A), middle (B), and late (C) release groups of female sockeye in Forfar Creek, 1952. T is stream life, i.e. the number of days from initial tagging to death. Data are taken from Table 38 of Killick (1955).

T	Group A	Group B	Group C	A+B+C
3			2	2
4			1	1
5			0	0
6	1		1	2
7	1	2	0	3
8	6	3	5	14
9	6	6	3	15
10	15	4	5	24
11	9	6	2	17
12	20	12	0	32
13	23	7	1	31
14	21	4		25
15	12	5		17
16	13	1		14
17	8	2		10
18	8			8
19	8			8
20	3			3
21	1			1
22	1			1
Tags recovered	156	52	20	228
Mean	13.59	11.81	8.40	12.73
s^2	10.08	6.00	7.00	11.17

with a variance estimate of:

$$\hat{V}(\bar{T}(B)) = \frac{s_B^2}{(52)} = 0.1777 \tag{3.91}$$

For the last group (C),

$$\bar{T}(C) = 8.40 \quad \text{days.} \tag{3.92}$$

with a variance estimate of:

$$\hat{V}(\bar{T}(C)) = \frac{s_C^2}{(20)} = 1.7670 \tag{3.93}$$

For this example, we assume that half the population of sockeye females arrived during the early group, one quarter arrived during the middle group and the remaining quarter arrived late. Weights for groups A, B, and C would be $w_A = 1/2$, $w_B = 1/4$, and $w_C = 1/4$. We further assume that the weights are known without error. If these assumptions are valid, then the overall mean estimate of stream life would be:

$$\hat{T} = 13.59(w_A) + 11.81(w_B) + 8.40(w_C) = 11.85, \tag{3.94}$$

with an overall estimate of variance of:

$$\begin{aligned} \hat{V}(\hat{T}) &= \left(\tfrac{1}{2}\right)^2 \hat{V}(\hat{T}_{mean}(A)) + \left(\tfrac{1}{4}\right)^2 \hat{V}(\hat{T}_{mean}(B)) + \left(\tfrac{1}{4}\right)^2 \hat{V}(\hat{T}_{mean}(C)) \\ &= \left(\tfrac{0.0646}{4}\right) + \left(\tfrac{0.1777}{16}\right) + \left(\tfrac{1.7670}{16}\right) \\ &= 0.1377 \end{aligned} \tag{3.95}$$

The standard error of the seasonal estimate of stream life is:

$$\hat{SE}(\hat{T}) = \sqrt{0.1377} = 0.3711, \tag{3.96}$$

and the coefficient of variation is estimated:

$$\hat{CV}(\hat{T}) = \frac{\sqrt{0.1377}}{11.85} = 0.03131. \tag{3.97}$$

If the weights were also estimated, say with a standard error on each weight estimate of 0.05, so that $V(w_i) = 0.0025$ for $i = A, B, C$, the variance estimate for stream life would be:

$$\begin{aligned}\hat{V}(\hat{T}) &= (w_A)^2\hat{V}(\hat{T}_A) + (\hat{T}_A)\hat{V}(w_A) - \hat{V}(\hat{T}_A)\hat{V}(w_A) \\ &\quad +(w_B)^2\hat{V}(\hat{T}_B) + (\hat{T}_B)\hat{V}(w_B) - \hat{V}(\hat{T}_B)\hat{V}(w_B) \\ &\quad +(w_C)^2\hat{V}(\hat{T}_C) + (\hat{T}_C)\hat{V}(w_C) - \hat{V}(\hat{T}_C)\hat{V}(w_C) \qquad (3.98) \\ &= 0.4777 + 0.3592 + 0.2824 \\ &= 1.119\end{aligned}$$

The standard error would be $\hat{SE}(\hat{T}) = 1.058$ and the coefficient of variation would be $\hat{CV}(\hat{T}) = 0.0893$.

If it can be assumed that all female sockeye had the same probability of being tagged, regardless of which group (A, B or C) they arrived with, then the data can be pooled, and a common mean can be estimated. This assumption is equivalent to assuming fish are tagged proportional to arrival abundance (see column "ABC" in Table 3.1).

3.4.16 Timer Tags, Parametric (Shardlow et al., 2007)

This method also uses the timer tags developed by Shardlow et al. (2007). The tags are applied to salmon entering the study area, record the time of the fish's death (assumed when the fish remains beyond a threshold angle for a pre-determined time), and are recovered from carcasses. For the parametric timer-tag estimator, the death times of individuals are assumed to be independent observations (A-31) from a specified distribution. The most common distributions used for continuous death times (or failure times) are the exponential (e.g. Cormack, 1964) and the Weibull (e.g. Lady and Skalski, 1998), though the lognormal has also been used to model salmon stream life (e.g. Labelle, 1994). The data are fit to the presumed distribution and the average stream life is taken to be the mean of that distribution. The parameters and the empirical variances of the parameters are estimated by maximum likelihood estimation (MLE) and the variance of the overall stream life estimate is typically estimated using the delta method.

If the assumed distribution of death times is a 2-parameter Weibull(η,β) (A-38), then η and β are estimated by fitting a Weibull(α,β) to the data. The expected stream residence time is estimated as the mean of the distribution as (Meeker and Escobar 1998 page 85):

$$\hat{T} = \hat{\eta}\Gamma\left(1 + \frac{1}{\hat{\beta}}\right), \tag{3.99}$$

where $\eta = 1/\alpha$.

The variance of the estimated stream life can be approximated using the delta method as in Equation (3.76):

$$V(\hat{T}) = V(\alpha)\left(\frac{\delta T}{\delta \alpha}\right)^2 + V(\beta)\left(\frac{\delta T}{\delta \beta}\right)^2 + 2Cov(\alpha,\beta)\left(\frac{\delta T}{\delta \alpha}\right)\left(\frac{\delta T}{\delta \beta}\right) \tag{3.100}$$

where

$$\frac{\delta T}{\delta \alpha} = \Gamma\left(1 + \frac{1}{\beta}\right)$$

and

$$\frac{\delta T}{\delta \beta} = \frac{-\alpha}{\beta^2}\int_0^\infty t^{1/\beta} log(t) e^{-t} dt.$$

Equation (3.100) is approximated by substituting the parameters with their fitted estimates:

$$\begin{aligned} \hat{V}(\hat{T}) &= \hat{V}(\hat{\alpha})\left(\Gamma\left[1 + \tfrac{1}{\hat{\beta}}\right]\right)^2 \\ &+ \hat{V}(\hat{\beta})\left(\tfrac{-\hat{\eta}}{\hat{\beta}^2}\int_0^\infty t^{1/\hat{\beta}} \log(t) e^{-t} dt\right)^2 \\ &+ 2\hat{Cov}(\hat{\alpha},\hat{\beta})\left(\Gamma\left[1 + \tfrac{1}{\hat{\beta}}\right]\right)\left(\tfrac{-\hat{\eta}}{\hat{\beta}^2}\int_0^\infty t^{1/\hat{\beta}} \log(t) e^{-t} dt\right) \end{aligned} \tag{3.101}$$

Alternatively, if the designated distribution is an Exponential(rate = ν) (A-33), then the estimate of stream life is $\frac{1}{\nu}$ and the variance of the estimate is:

$$V\left(\frac{1}{\hat{\nu}}\right) = \frac{V(\hat{\nu})}{\nu^4}. \tag{3.102}$$

The estimate of the variance would be:

$$\hat{V}\left(\frac{1}{\hat{\nu}}\right) = \frac{V(\hat{\nu})}{\hat{\nu}^4}. \tag{3.103}$$

The parametric estimate of stream life is preferable to the non-parametric estimate when the assumption of an underlying distribution is supported by the data. The parametric estimate will be more precise, i.e. have a smaller variance, than the non-parametric estimate. If the underlying model does not describe the data well, i.e. if a model fails a test of fit, then the parametric estimate may be biased to an unknown degree. The non-parametric estimator is more robust, i.e. unbiased regardless of the nature of the underlying distribution.

Parametric Timer Tags Example

The data for this example are taken from Column "ABC" in Table 38 of Killick (1955) and are reproduced in Table 3.1. For the purpose of this example, we assume that the values in the T column were obtained from timer tags and are accurate measures of the time from initial stream entry to death.

A two-parameter Weibull distribution was fit to the data using function "weibreg" in R. The results for both the shape (β) and scale (η) estimates and variance estimates are reported in terms of the natural logarithm of the parameters (Table 3.2). The reported estimates can be exponentiated to estimate the mean stream life because the estimates are maximum likelihood estimates. Therefore, the estimated mean stream life is:

$$\hat{T} = (13.996)\Gamma\left(1 + \frac{1}{4.181}\right) = 12.72 \quad \text{days.} \tag{3.104}$$

The variance estimate requires some adjustment due to the reported variances and covariance applied to the logarithm of the parameters. The estimate can be rewritten as:

$$\hat{T} = e^{\widehat{\log\eta}} \quad \Gamma\left(1 + \frac{1}{e^{\widehat{\log\beta}}}\right), \tag{3.105}$$

Table 3.2: Maximum likelihood estimates for the two-parameter Weibull fit to example data. The R output provides the natural logarithm of the parameters of interest.

Parameter	Estimate	log(Estimate)	V(log(Estimate))	Cov(log(Estimate))
Scale, η	13.996	2.639	0.0002794	0.0002689
Shape, β	4.181	1.430	0.002539	0.0002689

The variance for this re-parameterized estimate of stream life is:

$$V(\hat{T}) = V(\log\eta)\left(\frac{\delta T}{\delta\log\eta}\right)^2 + V(\log\beta)\left(\frac{\delta T}{\delta\log\beta}\right)^2 + 2Cov(\log\eta,\log\beta)\left(\frac{\delta T}{\delta\log\eta}\right)\left(\frac{\delta T}{\delta\log\beta}\right), \tag{3.106}$$

where

$$\frac{\delta T}{\delta\log\eta} = \left(e^{\log\eta}\right)\Gamma\left(1+\frac{1}{e^{\log\beta}}\right) = \eta\Gamma\left(1+\frac{1}{\beta}\right)$$

and

$$\frac{\delta T}{\delta\beta} = \frac{-\eta}{\beta}\int_0^\infty t^{1/\beta}log(t)e^{-t}dt.$$

Thus, the variance estimate for the parametric average stream life estimate using the two-parameter Weibull(η, β) distribution is:

$$\begin{aligned}\hat{V}(\hat{T}) &= (0.000279)(12.718)^2 + (0.00254)(1.647)^2 + 2(0.000269)(12.718)(1.647)\\ &= 0.0633 \quad \text{days.}\end{aligned} \tag{3.107}$$

The standard error is $\hat{SE}(\hat{T}) = 0.252$ and the coefficient of variation is $\hat{CV}(\hat{T}) = 0.0198$.

To linearize the stream life time data, i.e. the survival curve, the survival curve and stream lives must be transformed. The logarithm of the negative logarithm of the empirical survival curve provides the y-axis points and the logarithm of the stream lives, $\log(T_j)$, provides the x-axis points. Such a plot is shown in Figure 3.2. The overlaid line uses $\beta = 4.18$ as the slope and $-\beta l\eta = -11.03$ as the intercept.

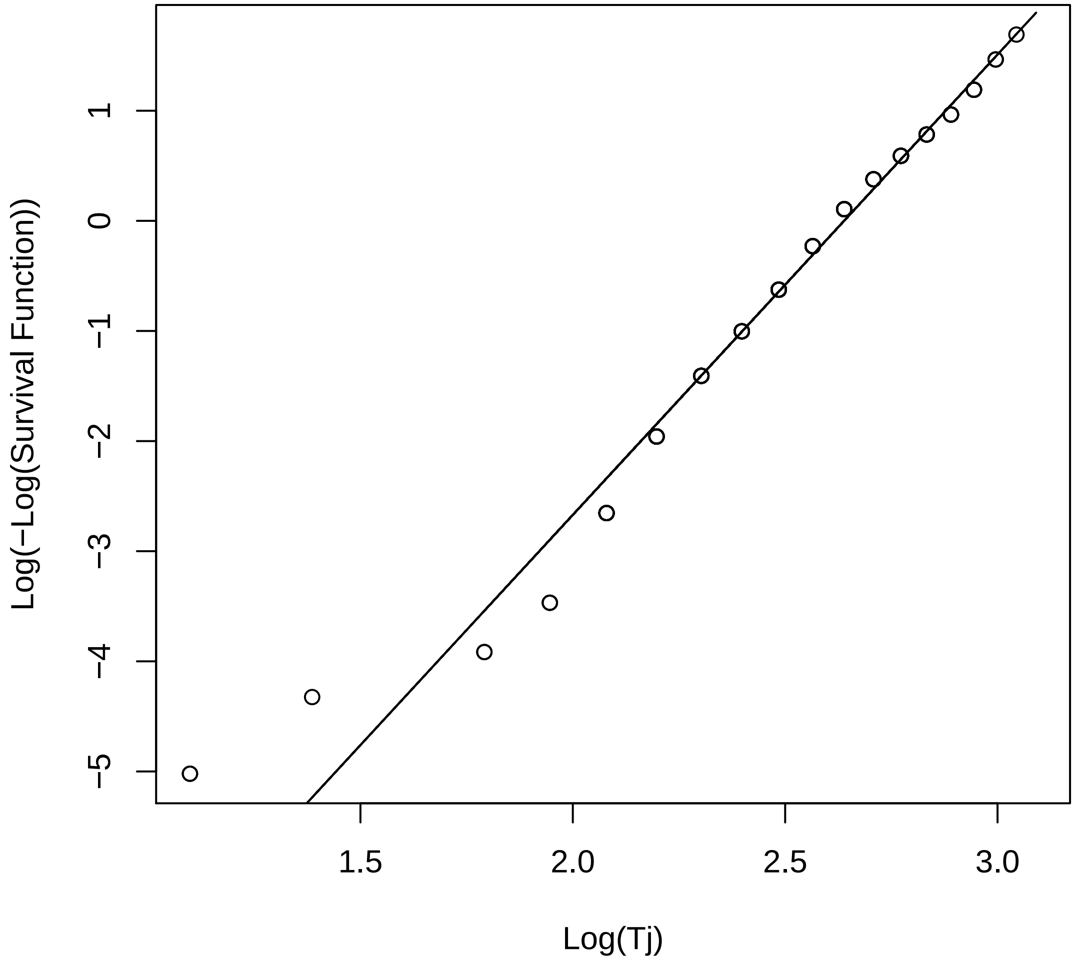

Figure 3.2: Stream life data linearized using to the two-parameter Weibull distribution (open circles) with the maximum likelihood best fit line superimposed. The linearized intercept is $-\beta l\eta = -11.03$ and the linearized slope is $\beta = 4.18$.

3.5 Discussion of Residence Time Estimators

An unbiased estimate of stream life is necessary for an unbiased AUC estimate of escapement. There are 15 different ways to estimate stream life, though only 5 are capable of producing unbiased estimates of stream life.

The Beidler and Nickelson (1980) method of estimating average stream residence time is a hopeful, but misguided approach. A universal and inflexible estimate of stream life is inappropriate in the light of the huge variation between streams, years, and run timing found from empirical studies, which invalidate the main assumption of the technique. Therefore, the constant stream life estimator of stream life should not be used (Table 3.3).

Of the three variations on the time-to-tag-recovery method, the "mean" estimator is the most appropriate because it provides an estimate of the variance that can be used in the AUC formula. While the estimate is positively biased for stream life due to the lag between mortality and recovery, the bias can be reduced by using only recently deceased carcasses. Killick (1955) also stressed the importance of surveying for carcasses every day in order to minimize the lag time between mortality and recovery. The time-to-tag-recovery method is straight-forward in execution and computation, but remains positively biased, and should not be used (Table 3.3).

Neither the Peak-to-Peak nor the Median-to-Median methods should be used to estimate salmon stream life (Table 3.3). Both methods use abundance as a proxy for arrival and so are not actually measuring stream life. Both methods have been shown theoretically and empirically to produce severely biased estimates of stream life (Killick, 1955; Liao, 1994). For example, Liao (1994) showed that the bias of the Peak-to-Peak estimator may be on the order of half the average stream life.

Neither the Expected-to-Expected Estimator nor the Inverted stream life method are useful for estimating a stream life for salmon escapement. Both methods require an estimate of escapement to generate their estimate of stream life.

Table 3.3: Evaluation of stream life estimation methods for use in the area-under-the-curve method.

Name	Estimator	Variance	Assumptions	Empirical Experience	Recommendation
Constant	Poor	N/A	Unreasonable	Inapplicable	Do Not Use
Time-to-tag recovery	Poor	Good	Unreasonable	Robust	Do Not Use
Peak-to-peak	Poor	None	Unreasonable	Inapplicable	Do Not Use
Median-to-median	Poor	None	Unreasonable	Inapplicable	Do Not Use
Expected-to-expected	Fair	Good	Conflictual	Inapplicable	Do Not Use
Inverted	Fair	Good	Conflictual	Inapplicable	Do Not Use
Tag-depletion	Fair	Good	Reasonable	Fair	Good Alternative
Redd residence	Poor	None	Unreasonable	Inapplicable	Do Not Use
Cormack	Fair	Good	Conflictual	Not much testing	Do Not Use
Live Non-parametric	Fair	Good	Reasonable	Not much testing	Good Alternative
Live and Dead Non-parametric	Good	Good	Reasonable	Not much testing	Good Alternative
Live Parametric	Fair	Good	Reasonable	Not much testing	Good Alternative
Live and Dead Parametric	Good	Good	Reasonable	Not much testing	Good Alternative
Timer Tags, Non-parametric	Good	Good	Reasonable	Not much testing	Best Choice
Timer Tags, Parametric	Good	Good	Reasonable	Not much testing	Best Choice

Bocking et al.'s (1988) tag-depletion method only provides an unbiased estimate of residence time if the surveys are conducted daily. Under relatively strict conditions, the tag-depletion method of estimating stream life can provide an unbiased estimate. If the season is short enough for changes in stream life to be considered negligible, or if the method can be repeated over the run and the multiple estimates weighted in an average, then the tag depletion method may be considered a method to estimate stream life.

The redd residence time of female salmon should not be used as a proxy for stream life. Like the Peak-to-Peak and Median-to-Median estimators, Redd Residence Time is not actually measuring the amount of time a salmon spends in its natal stream. Applying female redd residence time to females underestimates stream life due to holding behaviors of salmon before and after spawning. Female redd residence time (FRRT) and male stream life (MSL) are not comparable, and therefore FRRT should not be used as a proxy for MSL (Table 3.3).

The Cormack (1964) method of estimating life expectancy is not appropriate for salmon escapement. It does not allow for differences in survival probability between early arriving and late arriving fish, and assumes that the life expectancies of all fish remain constant during the run. A fish that arrived in the third week of the run is assumed to have the same survival probability as a fish that was present on the first day. However, all fish are assumed to have the same expected stream life. If two fish with the same expected stream life arrive at different times, the early arriving fish has a lower survival probability than the late arriving fish. Consequently, the assumptions required to support the Cormack (1964) method are at odds with each other, and inapplicable to salmon escapement (Table 3.3).

The Lady (1996) and Lady and Skalski (1998) methods use some of the same assumptions as the Cormack (1964) method, but apply to fish only at the initial tagging event. Consequently, all tagged fish are expected to have the same probability of survival and stream life and the assumptions are reasonably met. Where possible, including carcass detections makes the estimates more precise than estimates derived from Live Only counts. Lady (1996) recommended the parametric estimator, because it was more accurate than the non-parametric estimator, especially in cases where many of the tagged fish were still alive at the end of the study. However, if the data do not fit the assumed survival function (i.e. Weibull) then the non-parametric estimator is more accurate (i.e. less biased) and should be used (Lady, 1996, page 146).

Timer-tags provide an almost exact measurement of time from tagging (i.e. arrival) to death and are among the best options for estimating stream life, assuming the assumption of tilt angle holds. The measurement error is on the order of minutes (Shardlow et al., 2007), which is negligible when stream life is reported in numbers of days. Though timer tags have only been tested on one wild population, controlled tests indicated that deaths were recorded accurately (Shardlow et al., 2007). Therefore, the validity of the timer-tag method will depend on how representative the sample of tagged salmon is of the escapement. Provided that every fish has an equal probability of being tagged, the application gives a simple, unbiased estimate of average stream life. Timer-tags can also provide an unbiased estimate of stream life if the run is stratified by time and estimates from each strata are weighted by the proportion of fish entering during each stratum. When the survival curve based on individual stream lives follow a distribution such as the Weibull or Exponential, the parametric estimator will provide a more precise estimate of stream life. When the survival curve does not fit a specified distribution, the non-parametric estimator will be more accurate.

Chapter 4

MARK-RECAPTURE TECHNIQUES

4.1 Introduction to Mark-Recapture Techniques

Most mark-recapture methods incorporate some form of the Petersen estimator (Labelle, 1994). The Petersen estimator is a comparison of ratios between marked and unmarked individuals from the population of interest. Individuals from the population are captured and marked in such a way as to make them identifiable as having been seen before. Adult salmon are usually marked with a Petersen disk tag (e.g. Rajwani and Schwarz, 1997) or a spaghetti tag (e.g. Smith et al., 2005), sometimes in conjunction with an operculum punch (e.g. Bocking, 1991).

Gangmark and Fulton (1952) referred to the number of salmon in a river at any given time as the "standing crop" because new fish are added to the population by immigration and subtracted from the population by emigration (straying) or mortality (deaths). Such populations are referred to as "open." A "closed" population is one subject to neither additions nor subtractions of individuals.

The mark-recapture techniques used to estimate abundance differ between open and closed populations. Despite spawning salmon populations being "open," closed population models have been used to estimate escapement. Closed population methods can be valid as long as the population is closed "on one end," that is, closed to either additions or subtractions.

4.2 Mark-Recapture Notation

Due to the large number of symbols used in mark-recapture methods, the notation used in this chapter is divided into two parts: one for closed populations and one for open populations. Some parameters or statistics are denoted with more than one symbol. An attempt has been made to reflect each original author's notation, however some symbols have been changed to avoid having multiple definitions assigned to the same symbol. Most estimates are denoted with carets, for example, escapement, N, is estimated by $\hat{N}$. Vectors and matrices are denoted as bold characters. Sums across strata are denoted with dots, for example, the total number of fish tagged in stratum i and recovered any time is $\sum_{j=1}^{t} c_{ij} = c_{i\bullet}$. Expected values, variances, and covariances are denoted by $E()$, $V()$, and $Cov()$, respectively.

The following is the notation used in the closed population methods:

a_i= Number of fish tagged in tagging stratum i (same as m_i).

b_j= Number of fish recovered in recovery stratum j (c.g. n_i).

c_{ij}= Number of fish tagged in stratum i and recovered stratum j (same as m_{ij}).

$\mathbf{D}_x$= Denotes a diagonal matrix with elements from the vector $\mathbf{x}$.

$L(\mathbf{x})$= Denotes the likelihood of the parameters in $\mathbf{x}$.

λ= Tag miss rate, i.e. probability a tag will be overlooked given that a tag is present.

M_i= Known number of fish tagged in stratum i (same as a_i).

m_{ij}= Number of fish tagged in stratum i that are recovered in recovery stratum j (same as c_{ij}).

N= Escapement.

n_i= Number of fish captured at survey i (c.g. b_j).

$P(i|j)$= Probability that a fish entered during tagging stratum i given that it was recovered in stratum j.

$P(j|i)$= Probability that a fish is recovered in stratum j given that it entered during tagging stratum i.

$P(i)$= Unconditional probability of a fish entering the study during survey i.

p_i= The probability that a fish entering during stratum i will be detected later in any stratum.

p_j= The probability of a fish surviving and if it is in stratum j, being recovered there.

ϕ^*= The probability of survival.

r= Subsampling rate, i.e. the proportion of carcasses re-examined for tags.

s= The number of release (marking) strata.

t= The number of recovery strata.

θ_{ij}= The probability that a fish arriving during stratum i and being caught in stratum j given that it survived.

The notation for open population mark-recapture methods is based largely on that of Schwarz and Arnason (1996) and is similar to that of Seber (1982). The following notation is used in open population methods:

B_i net number of fish that enter after i and survive to sample $i+1$.

B_i^* gross number of fish that enter between i and $i+1$.

b_ω= The number of fish with capture history ω which do not ultimately become losses on capture.

β_i fraction of total births that enter during interval between i and $i+1$.

χ_i= probability of not being recaptured after i conditional on being captured at i.

d_ω= The number of fish with capture history ω that ultimately become losses on capture.

$f(x)$= Generically, a function of x.

k= The number of mark-recapture surveys.

l_i= The number of fish not returned to stream after survey occasion i, losses on capture.

λ_i= probability a fish is recaptured after i conditional on being (re)captured at i.

ln= The natural logarithm, i.e. logarithm with base e.

M_i= The number of marked fish in the stream at time i.

m_i= The number of marked fish in sample taken at time i.

N_i= Abundance of fish in study area at time i.

n_i= The number of fish in sample taken at time i.

p_i= The probability of capture at survey occasion i.

Ψ_i= probability a fish enters stream, lives, and is not seen before survey i.

ϕ_i= probability of surviving from i to $i+1$ conditional on being alive at i.

R_i= The number of fish returned to stream tagged after survey occasion i.

r_i= the number of R_i that are subsequently recaptured alive.

τ_i= probability a fish is recaptured at i conditional on being recaptured on or after i.

U_i= Total number of unmarked fish in study area at time i.

u_i= The number of unmarked fish in sample taken at time i.

ν= the probability that a fish captured at i will be released.

ω= Denotes capture history.

z_i= The number of fish recaptured before and after survey i, but not at survey i.

4.3 Mark-Recapture Assumptions

The following assumptions are required by one or more of the mark-recapture methods. Not all assumptions are required by each method, and no single assumption is used in every method. The validity of each assumption is discussed as it arises in the development of the estimator.

(A-1). Every fish has the same (and independent) probability of being tagged.

(A-2). Every fish has the same (and independent) probability of being detected, regardless of tag status.

(A-3). Fish do not leave the spawning ground once they have arrived.

(A-4). Carcasses are not removed from the stream except by surveyors.

(A-5). Recaptures continue until all fish have arrived in the spawning area.

(A-6). All fish are correctly identified with regard to tag status, i.e. no tags are lost or overlooked.

(A-7). The probability of detection is independent of tag status.

(A-8). Carcasses are sampled without replacement, i.e. they are 'dead pitched."

(A-9). The number of tags recovered is at least 10.

(A-10). No tags are missed when carcasses are checked for tags the second time.

(A-11). The probability of detecting a tag (given that a tag is present) is independent for every carcass examination.

(A-12). All fish within a stratum have the same probability of tagging.

(A-13). All carcasses within a recovery stratum have the same probability of detection.

(A-14). Tagging is proportional to the population size.

(A-15). Tagging has no impact on the survival of tagged fish.

(A-16). Tagging has no impact on the detectability of carcasses.

(A-17). The number of tagging and recovery strata are equal.

(A-18). The probability of a fish escaping is independent of tagging strata.

(A-19). All surveys are conducted independently.

(A-20). All surveys are conducted instantaneously.

(A-21). The probability of being returned to the stream after being captured is the same for all fish.

(A-22). The probability of being returned to the stream after capture is independent of tag status.

(A-23). The probability of being returned to the stream after capture is independent of capture history.

(A-24). Every carcass is has the same (and independent) probability of retention.

(A-25). Each unmarked salmon has the same probability of being captured.

(A-26). Tagging has no impact on the probability of capture.

(A-27). All fish have the same probability of survival between surveys.

(A-28). Emigration is permanent, i.e. fish enter the stream only once.

(A-29). There is no immigration (study must commence at the start of the run).

(A-30). The number of fish arriving prior to the start of the study would be small enough to be reasonably ignored.

(A-31). The study is carried out until the end of the run.

(A-32). The method used to estimate the initial abundance is independent of the method used to estimate subsequent abundances.

(A-33). The number of hatchery fish removed from the study is known without error.

(A-34). The number of wild fish removed from the study is known without error.

(A-35). Mortalities occur uniformly over the interval.

4.4 Techniques for Closed Populations

Despite the fact that the number of salmon spawning at any given time of the run fluctuates due to new arrivals and deaths, closed population techniques have been used to estimate salmon escapement. “Petersen methods” (Seber, 1982, page 59) can still provide a consistent estimate of abundance provided that *only* immigration (birth) *or* death (emigration), but not both, occur over the course of the study (Kendall, 1999). In the case of no emigration (i.e. no fish leave or die) the estimator is valid for the population size at the time of the second sample. In the case of no immigration (i.e. no new arrivals) the Petersen Estimator estimates abundance at the start of the study. During salmon runs, clearly both immigration and emigration do occur. Consequently, Petersen methods must be carefully applied when both immigration and death occur in the same study.

Typically in closed population mark-recapture studies, salmon are tagged while alive, either as they enter or when they are already on the spawning ground, and “recaptured” (detected) as carcasses (Simpson, 1984; Bocking, 1991; Rajwani and Schwarz, 1997; Maselko et al., 2003). If it is reasonable to assume no carcass loss, then an estimate of abundance applies to the time of the recaptures, which should be after all the fish have arrived, and hence provide an estimate of escapement.

4.4.1 Pooled Petersen Method (Schubert, 2000)

The Pooled Petersen Estimator is essentially the Petersen Estimator spread over the course of the run. Salmon are tagged or marked as they enter the study area and then released. Carcasses are collected periodically and examined for tags and both the number examined and the number of tags recovered are recorded.

If all fish are independent and identically distributed with respect to the probability of tagging (A-1) or the probability of detection (A-2), then it is possible to construct a likelihood for the total escapement. The likelihood is appropriate only if fish do not stray from the study area (A-3) and carcasses are not removed from the stream except by surveyors (A-4). The last two assumptions ensure that the population is closed to emigration. If the population is closed to emigration, then the Pooled Petersen estimate applies to the abundance at the end of the study. If the recaptures occur after all fish have arrived (A-5), then the estimate applies to the total escapement.

Usually the carcasses are "dead pitched," meaning that they are removed from the stream and tossed above the high-water mark (e.g. Johnston et al., 1986). If carcasses are not dead pitched, i.e. if the carcasses are sampled with replacement the probability of recapturing m tags can be described using a binomial distribution:

$$Pr(\text{recaptures} = m) = \begin{pmatrix} n_2 \\ m \end{pmatrix} \left(\frac{M}{N}\right)^m \left(1 - \frac{M}{N}\right)^{N-m} \tag{4.1}$$

where M is the total number of fish tagged, n_2 is the number of carcasses checked for tags, m is the number of tags found, and N is escapement.

The estimate of escapement for sampling with replacement is:

$$\hat{N} = \frac{Mn_2}{m}, \tag{4.2}$$

which is biased, especially at small sample sizes. Bailey (1951, 1952) suggested a modification that provides a nearly unbiased estimate of abundance for sampling with replacement:

$$\hat{N}_B^* = \frac{M(n_2 + 1)}{(m + 1)}. \tag{4.3}$$

The Bailey modification has been used by Crone and Bond (1974) and Maselko et al. (2003). A nearly unbiased estimate of the variance of Bailey's modification (Bailey, 1952) is:

$$\hat{V}(\hat{N}_B^*) = \frac{M^2(n_2 + 1)(n_2 - m)}{(m + 1)^2(m + 2)}. \tag{4.4}$$

Equation (4.3) is reasonably unbiased as long as the tag status of all fish is correctly identified (A-6). If tags are lost or overlooked, m will be undercounted and the escapement estimate will be positively biased. Similarly, the probability of detection must be independent of tag status (A-7). If tagged fish are recovered with a higher probability than untagged fish, escapement will be underestimated. There is evidence that tagging may actually reduce the probability of detection. Schubert (2000) found that tagged males did not swim as far into the stream as untagged males and that a disproportionate number of tagged males were trapped in his carcass weir. He concluded that tagged males were more likely to be washed out of the study stream and were therefore less likely to be detected. If tagged carcasses have lower detection rates than untagged carcasses then the escapement estimate will be positively biased.

In order for the Pooled Petersen Estimator given in Equation (4.2) to be a valid estimate of total escapement, the population must be closed to emigration. That means that carcasses do not leave the study area unless they have been recovered by surveyors (A-4). However, carcasses and moribund fish are frequently washed downstream and out of the study area (Simpson, 1984; Johnston et al., 1986; Schubert, 2000). These losses can be reduced by a weir at the downstream end of the study area (Simpson, 1984), but weirs are often rendered inoperable during periods of high run-off, precisely when the rate of carcass loss is greatest (Johnston et al., 1986; Shardlow et al., 1986). Additional carcasses are lost to scavengers, decay, and deep pools (e.g. Johnston et al., 1986). Even live fish may be removed by predators such as bears or bald eagles (Dunwiddie and Kuntz, 2001; Gende et al., 2001).

Another key assumption of the closed population models is that either 1) every fish has the same probability of being tagged while alive (A-1), or 2) every fish has the same probability of being detected when dead (as a carcass) (A-2). For every entering fish to have an equal probability of being tagged, the same proportion of the run must be tagged every day. This requires knowledge of the total escapement and the daily portion of the total escapement *a priori*. If the size of a run was known in advance, there would be no need to conduct a mark-recapture study. Some studies have tagged fish in proportion to the previous day's escapement (e.g. Rajwani and Schwarz, 1997), or using visual estimates of the numbers of spawners downstream of the study area (e.g. Schubert, 2000). These methods are subjective, but the pooled Petersen estimator seems to be robust to minor violations in the assumption of a constant tagging probability. In fact, Junge (1963) showed that tagging and recovery need not be random as long as the probability of recovery is independent of the probability of being tagged.

If consistent tagging cannot be assumed, then one must rely on the assumption that every carcass has the same probability of being detected (A-2), or at least that recovery is not dependent on tag status (A-7). Maselko et al. (2003) asserted that every carcass had an equal probability of detection if recovery surveys were "at fixed intervals over the course of sampling." If mortalities are random with respect to surveys, then even fixed surveys can recover random samples of carcasses. The recovery surveys do not need to be at fixed intervals. Every carcass has the same probability of detection only if sex does not influence the detectability of dead salmon. Shardlow et al. (1986, page 12) found that female salmon had higher carcass recoveries than males, and cited several studies with similar findings. They attributed the difference in carcass recoveries to behavioral differences between males and females. Females guarded their redds until death, remaining on the spawning grounds as carcasses, whereas males held in deep pools or were washed out to sea "in a moribund state."

Schubert (2000) also found that male sockeye were more likely to be flushed out of the study area and were therefore less likely to be detected during carcass surveys than females. The males and females can be estimated separately (e.g. Schubert, 2000), removing the problem of differential carcass retention rates due to sex.

If carcasses are sampled without replacement (A-8) the probability of recapturing m tags can be described with a hypergeometric distribution:

$$Pr(\text{recoveries} = m) = \frac{\begin{pmatrix} n_2 \\ m \end{pmatrix} \begin{pmatrix} N & - & M \\ n_2 & - & m \end{pmatrix}}{\begin{pmatrix} N \\ m \end{pmatrix}} \tag{4.5}$$

where M is the total number of fish tagged, n_2 is the number of carcasses checked for tags, m is the number of tags found, and N is escapement. The estimate of escapement is the same as that in Equation (4.2).

The estimator in Equation (4.2) is biased for either sampling strategy (with or without replacement) with the bias increasing as the number of recaptures decreases (Chapman, 1951). Chapman (1951) provided an estimator that is unbiased when the sum of the two sample sizes is greater than the population, $M + n_2 \geq N$:

$$\hat{N}_C^* = \frac{(M+1)(n_2+1)}{(m+1)} - 1. \tag{4.6}$$

When the sum of the two samples is less than the population, $M + n_2 < N$, Equation (4.6) is reasonably unbiased as long as the number of recaptures (detections) is greater than 7 (Robson and Regier, 1964). Chapman (1951, page 148) recommended that the number of recaptures be at least 10 (A-9) (Seber, 1982, page 60). The Chapman modification has been used to estimate salmon escapement by Frith and Nelson (1994) and Jacobs (2002).

The variance of the Chapman-modified pooled Petersen estimator is estimated by (Seber, 1982, page 138):

$$\hat{V}(\hat{N}_C^*) = \frac{(M+1)(n_2+1)(M-m)(n_2-m)}{(m+1)^2(m+2)}. \tag{4.7}$$

Even with the bias corrections, the Pooled Petersen estimates of escapement (Equations 4.3 and 4.6) are unbiased only when the assumptions are met, which Simpson (1984) contended had never been accomplished. Simpson (1984) compiled data from ten different streams and multiple years between 1938 and 1980 for which comparisons between the Pooled Petersen estimator and fence counts had been made. He found that the Pooled Petersen estimate routinely overestimated escapement compared to wier count data. The findings of Schubert (2000) and Maselko et al. (2003) agree with those of Simpson (1984) that Pooled Petersen estimates of escapement are positively biased. Schubert (2000) conjectured that the Pooled Petersen estimate was positively biased for escapement due to unequal probabilities of detection of marked carcasses.

Adjustment for overlooked tags (Rajwani and Schwarz, 1997)

A key assumption in the Pooled Petersen Method is that the tag status of all fish is correctly identified (A-6). If a tag is lost between release and detection, or if a tag is simply overlooked during recovery surveys, the recorded number of tagged fish recovered will under-represent the number of truly tagged fish. Consequently, the estimates of escapement given in either Equation (4.3) or Equation (4.6) will be positively biased.

Rajwani and Schwarz (1997) derived a maximum likelihood method for estimating abundance in the presence of overlooked tags. Sockeye salmon were tagged as they entered the Fraser River, Canada. Carcasses were examined for tags and those found with tags were cut in half and thrown from the study area. Carcasses that were found without tags were placed back in the river. Later, n_3 of the allegedly untagged carcasses were re-examined for tags and m_3 tags were found. Thus, the n_2 carcasses initially examined for tags could be classified into one of four categories: 1) tag was found the first time the carcass was detected, 2) tag was missed and then found, 3) the fish had not been tagged, or 4) the tag was missed on both examinations. Rajwani and Schwarz (1997) used a multinomial distribution to describe the numbers of fish in each of the four categories, which requires that the probability of detecting a tag, given that a tag is present, must be independent for every examination (A-11). in other words, the detectability of a tag must not depend on whether or not the carcass has been labeled "unmarked."

Rajwani and Schwarz (1997) found that the maximum likelihood estimate of the total number of tags present in a sample was:

$$\hat{m} = m_2 + \frac{m_3}{n_3}(n_2 - m_2), \tag{4.8}$$

which replaces m in the Pooled Petersen estimate of escapement. Rajwani and Schwarz (1997) recommend the Bailey (1951) correction to estimate escapement:

$$\hat{N} = \frac{n_1(n_2 + 1)}{(\hat{m} + 1)}, \tag{4.9}$$

which is still positively biased when "the miss rate is large, the subsampling rate is small, and the expected number of tags is small" (Rajwani and Schwarz, 1997) but is otherwise reasonably unbiased. (Rajwani and Schwarz, 1997) gave no indication as to what constituted "large" or "small."

Estimating the number of missed tags increases the variance of escapement, which Rajwani and Schwarz (1997) calculated using a variance inflation factor (VIF):

$$VIF = 1 + \frac{N\lambda(1 - r)}{(N - n_1(1 - \lambda))r}, \tag{4.10}$$

where $\lambda = \frac{(n_2 - m_2)m_3}{\hat{m}n_3}$ is the miss rate (i.e. probability a tag will be overlooked) and $r = n_3/(n_2 - m_2)$ is the subsampling rate (probability of being re-examined after being found without a tag). If no tags are missed ($\lambda = 0$) or all "untagged" carcasses are re-examined ($r = 1$) then the variance of the escapement estimate is the same as that of Equation (4.4). In other cases, the variance estimate of $\hat{N}$ is Equation 4.4 multiplied by the VIF.

Rajwani and Schwarz (1997) placed carcasses that were classified as untagged back in the river, while tagged carcasses were removed. Since recovery surveys were conducted over the entire course of the run, the carcasses classified as "untagged" were available for re-detection during the entire course of the study. Thus, untagged carcasses were likely to be counted multiple times, making the number of untagged carcasses seem larger than it was, and causing positive bias to the escapement estimate. However, had untagged carcasses been thrown above the high water mark and cut in half after re-examination, the method would have been valid. Rajwani and Schwarz (1997) recommended thorough examination of every carcass the first time to ensure that no tags are missed and no adjustment would be necessary (A-6).

4.4.2 Schaefer (1951) Method

The Pooled Petersen estimator (Section 4.4.1) requires that all escaped individuals have the same probability of being marked *or* the same probability of being recovered. Effort devoted to tagging or recovering salmon may change over the course of the run due to resource or environmental limitations. Tagging or detection probabilities cannot be considered the same for all members of the escapement and "samples drawn from any particular part of the season do not represent all parts of the population equally" (Schaefer, 1951). Schaefer (1951) proposed an estimator for escapement based on several time strata such that fish in each arrival (tagging) strata can be assumed to have the same probability of being tagged (A-12) or that each carcass present within a recovery strata can be assumed to have the same probability of being detected (A-13). Live individuals are tagged upon entry with time-specific tags. The tags are recovered from carcasses during subsequent time strata and the numbers of carcasses from each tagging stratum and the number of untagged carcasses are recorded. Abundance estimates are made for each of the recovery strata, and escapement is the sum of the strata estimates.

Let $P(j|i)$ denote the probability that an individual tagged during the ith time stratum dies and is recovered during the jth recovery stratum. Let $P(i|j)$ denote the probability that a carcass detected during the jth stratum entered the survey area during the ith tagging stratum (whether it was tagged or not). By Bayes' Theorem:

$$P(i|j) = \frac{P(j|i)P(i)}{\sum_i P(j|i)P(i)} \tag{4.11}$$

Assuming that tagging was proportional to the population size (A-14), the probability of having entered the stream during the ith stratum, P(i), can be estimated as the fraction of the marked individuals entering during the ith stratum divided by the total number of marked individuals:

$$\hat{P}(i) = \frac{m_{i\bullet}}{m_{\bullet\bullet}}, \tag{4.12}$$

as long as tags were not lost or overlooked (A-6). Lost and overlooked tags will inflate a stratified estimate just as they would a Pooled Petersen estimate.

Tagging must be proportional to arrival for Schaefer (1951) estimates of the probability of being in a tagging strata to be valid. However, proportional tagging negates the need to stratify a population.

The probability of dying and being recovered in the jth stratum given being marked in the ith stratum, P($j|i$), can be estimated as the fraction of the number marked in i and recovered in j divided by the total marked in i:

$$\hat{P}(j|i) = \frac{m_{ij}}{m_{i\bullet}} \tag{4.13}$$

From Equations 4.12 and 4.13 we can estimate the probability that a carcass found during time stratum j had been tagged during time stratum i as:

$$\hat{P}(i|j) = \frac{\frac{m_{i\bullet}}{m_{\bullet\bullet}}\frac{m_{ij}}{m_{i\bullet}}}{\sum_i \frac{m_{i\bullet}}{m_{\bullet\bullet}}\frac{m_{ij}}{m_{i\bullet}}} = \frac{\frac{m_{ij}}{m_{\bullet\bullet}}}{\frac{m_{\bullet j}}{m_{\bullet\bullet}}} = \frac{m_{ij}}{m_{\bullet j}} \tag{4.14}$$

An estimate of the number of individuals detected in j from i is the probability of being from i given being found in j, times the number found in j:

$$\hat{n}_{ij} = n_j \hat{P}(i|j) = n_j \frac{m_{ij}}{m_{\bullet j}} \tag{4.15}$$

Now, an estimate of the actual number of individuals in j from i, N_{ij}, is the estimate of the number detected, $\hat{n}_{ij}$, divided by the detection probability of individuals from that stratum, p_i. The probability of an individual from stratum i being detected later in any stratum is:

$$p_i = \frac{\sum_j n_{ij}}{N_i}$$

and is estimated by:

$$\hat{p}_i = \frac{m_{i\bullet}}{M_{i\bullet}}. \tag{4.16}$$

To apply the recovery probability of tagged fish to the untagged fish, tagging must have no impact on the survival (A-15) or detectability (A-16) of a tagged fish. As mentioned in Section 4.4.1, Schubert (2000) found evidence that tagged salmon were less likely to be recovered than untagged salmon, and tagging often shortens the survival of salmon. Shardlow et al. (1986) noted that handling unripe chinook caused premature death. Crone and Bond (1974) found that tagged sockeye had shorter redd lives than untagged sockeye. Johnston et al. (1986) accounted for tagging mortality in pink, chum, and coho by holding newly tagged fish for several hours to exclude from analysis those that died due to handling. Since tagged fish are likely to have a lower survival probability than untagged fish for all periods, the estimate of survival will be negatively biased. Consequently estimates of new arrivals will be positively biased, and escapement will be too large. Even in a simulation with these assumptions met, the stratified estimator over estimated escapement (Law, 1994).

The estimate of the total number of fish tagged in i and available for recovery in j is:

$$\hat{N}_{ij} = \frac{\hat{n}_{ij}}{\hat{p}_i}.$$

The total abundance is the sum of the abundances from each of the tagging strata to each of the recovery strata:

$$\hat{N} = \sum_i \sum_j \hat{N}_{ij} = \sum_i \sum_j n_j \frac{m_{ij}}{m_{\bullet j}} \frac{M_{i\bullet}}{m_{i\bullet}} \tag{4.17}$$

Chapman and Junge (1956) showed, using a numerical example, that the Schaefer (1951) estimator was inconsistent. They provided an alternative derived from the estimating equations:

$$\sum_{i=1}^{k} m_{ij}\frac{\hat{N}_{i\bullet}}{M_i} = n_{\bullet j} \tag{4.18}$$

where $j = 1, 2, ..., k$. Ricker (1975) asserted that the Chapman and Junge modification was more complicated than it was worth since it required the same assumptions as the Petersen Estimator.

Schaefer (1951) did not provide a variance estimator. No closed formula for the variance exists without making assumptions about the distribution of the estimator or its components. Chapman and Junge (1956) provided a closed formula for the variance of their estimator based on the assumption that the conditional distributions of tagged and recovered individuals were multihypergeometric. Johnston et al. (1986) estimated the variance using a Monte Carlo simulation assuming a Poisson distribution on the data. The resulting 95% confidence interval did not encompass the known (wier) escapement value (Johnston et al., 1986)

4.4.3 Darroch (1961) Estimator

Darroch (1961) was the first to develop a full likelihood approach to stratified tagging and recapturing. He analyzed the data of Schaeffer (1951), where sockeye salmon were tagged as they entered the stream and were recovered as carcasses, to illustrate his method. The total number recaptured was assumed to follow a binomial distribution. The number of salmon from each tagging stratum were assumed to be multinomially distributed among recovery strata. The estimate of abundance for each strata was derived from the joint likelihood of the multinomial and binomial distributions. Total escapement was estimated as the sum of strata totals.

The derivation of the Darroch (1961) Estimator is long and arduous. We refer the reader to the original work, and present the likelihood and results given by Darroch (1961).

Let θ^*_{ij} be the probability that a fish entering the study area during time stratum i

becomes a carcass available for capture at time stratum j, and let p_j^* be the probability that a carcass present at j is detected. Let $\sum_j \theta_{ij}^* = \phi_i^*$ be the survival probability of fish entering during tagging strata i. The number of carcasses tagged in stratum i can distribute themselves among the recovery strata according to the multinomial likelihood (Darroch's Equation (1*)):

$$P(\mathbf{c_{ij}}|\mathbf{a_i}) = \frac{\prod_i a_i!}{\prod_i (a_i - c_{i.})! \prod_{ij} c_{ij}!} \prod_i (1 - \sum_i \theta_{ij}^* p_j^*)^{a_i - c_i} \prod_{ij} (\theta_{ij}^* p_j^*)^{c_{ij}}. \tag{4.19}$$

The total number of individuals recaptured in a given stratum is modeled binomially as (Darroch's Equation (2*)):

$$P(\mathbf{b_j}) = \prod_j \begin{pmatrix} n_j^* \\ b_j \end{pmatrix} (p_j^*)^{b_j} (1 - p_j^*)^{n_j^* - b_j}. \tag{4.20}$$

The joint likelihood is the product of (4.19) and (4.20). The movement and detection probabilities, θ_{ij}^* and p_j^* are non-identifiable in that the likelihood remains the same for any constant k relating to the movement and detection probabilities like so: $k\theta_{ij}^*$ and $k^{-1}p_j^*$. Darroch therefore assumed that all survival probabilities were equal and used the common ϕ^* to generate $\theta_{ij} = \theta_{ij}^*/\phi^*$ as the probability that any fish from tagging strata i is present in recovery strata j given that it is alive at the time, and $p_j = \phi^* p_j^*$ as the probability of a fish surviving and being caught given that it is in recovery strata j. If it can be assumed that there is neither straying (A-3), nor carcass loss (A-4) over the course of the study, then $\phi^* = 1$ and estimates resulting from the maximization of the joint likelihood do not need to be adjusted.

The assumption of equal survival (A-27) is applied differently depending on the design of the study. In a typical stratified closed population mark-recapture study, survival applies to the probability that a live fish will stay in the spawning area until death and remain as a carcass. In a study like that of Underwood et al. (2007) survival applies only to live fish. Underwood et al. (2007) stratified the run in time and space by tagging chum salmon downstream and later recapturing them in fish wheels upstream. The population was considered closed because all salmon passing the tagging sites had to migrate beyond the recapture sites to find suitable spawning grounds. Within this context it is reasonable to expect the same probability of survival because all fish in a tagging cohort are subject to approximately the same predation levels and other environmental conditions. Travel times from the first location of tagging to the recovery location are allowed to vary, which results in mixing in recovery strata.

When recovering carcasses rather than live fish, "survival probability" refers to the probability that a fish will be not be removed from the study while alive or as a carcass. Fish from a tagging cohort are expected to die at different times, and carcasses decay and washout at different rates due to temperature and river flow changes. Consequently, when applied to recovery of carcasses, the assumption of equal "survival" is unreliable and Darroch Method should not be used.

The joint likelihood contains $st + t - s$ parameters but only st statistics. Consequently, the parameters are not identifiable when the number of recovery strata is greater than the number of tagging strata, $s < t$. In this case, Darroch (1961) recommended pooling recovery strata until the numbers of tagging and recovery strata are equal. When the tagging strata outnumber the recovery strata, $s > t$ then all parameters are estimable, but only through the use of numeric iteration.

When the number of tagging strata is the same as the number of recovery strata (A-17), the Darroch (1961) estimate of escapement, in vector notation is:

$$\hat{N} = \mathbf{b}'\mathbf{C}^{-1}\mathbf{a} \tag{4.21}$$

where $\mathbf{b}$ is a vector with elements b_j, the number of untagged fish captured during recovery stratum j, $\mathbf{a}$ is a vector with elements a_i, the number of fish tagged in release stratum i, and $\mathbf{C}$ is a matrix with elements c_{ij}, the number of tagged fish released in stratum i and recovered in stratum j. The estimate of escapement given in Equation (4.21) relies on the inverse of the survival probability matrix, $\mathbf{C}^{-1}$. Equation (4.21) is consistent but slightly biased for escapement. The bias decreases with increased sample size and Darroch (1961) considered it to be negligible.

Equation (4.21) requires that all tags are read (i.e. that the c_{ij} are known without error). Recaptured fish can escape before the tags can be read. Underwood et al. (2007) expanded the number of unknown tags, u_j by the portion of the known c_{ij} for each combination of strata. This adjustment assumes that the probability of a fish escaping is independent of tagging strata (A-18). Since fish escaping is usually regarded as a random event, Assumption (A-18) seems reasonable. The estimated number of stratum i tags recovered in j is:

$$\hat{c}_{ij} = c_{ij} + u_j \frac{c_{ij}}{\sum_{i=1}^{s} c_{ij}} \tag{4.22}$$

The variance of the escapement estimate depends on the variance of the inverse of detection probability. When the number of tagging and recovery strata are equal (A-17), the variance-covariance matrix of the inverse detection probabilities is estimated by:

$$\mathbf{\Sigma} = \mathbf{D}_\rho \mathbf{\Theta}^{-1} \mathbf{D}_\mu \mathbf{D}_a^{-1} \mathbf{\Theta}'^{-1} \mathbf{D}_\rho, \tag{4.23}$$

where $\mathbf{D}_\rho$ is the diagonal matrix with elements $\mathbf{C}^{-1}\mathbf{a}$, $\mathbf{\Theta} = \mathbf{D}_\mathbf{a}^{-1}\mathbf{C}\mathbf{D}_\rho$ is the full rank matrix with elements θ_{ij}, $\mathbf{D}_\mu$ is the diagonal matrix with elements $\mu_i = \sum_{j=1}^{t} \frac{\theta_{ij}}{p_j} - 1$, and $\mathbf{D}_a$ is the diagonal matrix with elements a_i as the number tagged in each tagging strata.

The variance of the escapement estimate is estimated by:

$$\hat{V}(\hat{N}) = \mathbf{b}'\mathbf{\Sigma}\mathbf{b}. \tag{4.24}$$

If the number of tags from stratum i recovered in stratum j, c_{ij}, must be estimated, Equation (4.24) will underestimate the true variance.

Equation (4.24) will generally be larger than the Pooled Petersen Estimator. Therefore, Darroch (1961) recommended stratifying mark-recapture studies only to ensure that the required assumptions are valid.

Darroch Method Example

The data for this example were taken from Schaefer (1951). Darroch (1961) collapsed Schaefer's eight release strata and nine recovery strata into four release and three recovery strata. We collapsed Schaefer's data even further into three release and three recovery strata (Table 4.1).

Table 4.1: Data for the Darroch Method example, taken from Schaefer (1951). The c_{ij} are the numbers of fish released in stratum i and recaptured in stratum j. The a_i are the number of fish tagged in stratum i. The b_j are the number of fish captured in stratum j.

	c_{i1}	c_{i2}	c_{i3}	a_i
c_{1j}	59	29	2	484
c_{2j}	45	275	43	1468
c_{3j}	0	39	28	399
b_j	847	5981	3124	

Using Equation (4.21), the Darroch estimate of escapement is:

$$\hat{N} = [847 \quad 5981 \quad 3124] \begin{bmatrix} 59 & 29 & 2 \\ 45 & 275 & 43 \\ 0 & 39 & 28 \end{bmatrix}^{-1} \begin{bmatrix} 484 \\ 1468 \\ 399 \end{bmatrix} = 54,318 \quad \text{fish.} \tag{4.25}$$

The matrices required for the variance estimate are:

$$\mathbf{D}_\rho = Diag\left(\begin{bmatrix} 59 & 29 & 2 \\ 45 & 275 & 43 \\ 0 & 39 & 28 \end{bmatrix}^{-1} \begin{bmatrix} 484 \\ 1468 \\ 399 \end{bmatrix} \right) = \begin{bmatrix} 6.56 & 0 & 0 \\ 0 & 2.60 & 0 \\ 0 & 0 & 10.62 \end{bmatrix}, \tag{4.26}$$

$$\begin{aligned}\mathbf{\Theta} &= \left(\begin{bmatrix} 484 & 0 & 0 \\ 0 & 1468 & 0 \\ 0 & 0 & 399 \end{bmatrix}^{-1} \begin{bmatrix} 59 & 29 & 2 \\ 45 & 275 & 43 \\ 0 & 39 & 28 \end{bmatrix}^{-1} \begin{bmatrix} 6.56 & 0 & 0 \\ 0 & 2.60 & 0 \\ 0 & 0 & 10.62 \end{bmatrix}\right) \\ &= \begin{bmatrix} 0.800 & 0.156 & 0.044 \\ 0.201 & 0.488 & 0.311 \\ 0.000 & 0.254 & 0.746 \end{bmatrix}\end{aligned} \tag{4.27}$$

$$\begin{aligned}\mathbf{D}_\mu &= Diag\left(\begin{bmatrix} 0.800 & 0.156 & 0.044 \\ 0.201 & 0.488 & 0.311 \\ 0.000 & 0.254 & 0.746 \end{bmatrix} \begin{bmatrix} 6.564 \\ 2.603 \\ 10.625 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) \\ &= \begin{bmatrix} 5.124 & 0 & 0 \\ 0 & 4.896 & 0 \\ 0 & 0 & 7.584 \end{bmatrix}\end{aligned}, \tag{4.28}$$

and

$$\mathbf{D}_a = \begin{bmatrix} 484 & 0 & 0 \\ 0 & 1468 & 0 \\ 0 & 0 & 399 \end{bmatrix}. \tag{4.29}$$

The matrices in Equations (4.26) - (4.29) combine according to Equation (4.23) to form the estimated variance-covariance matrix of the inverse detection probabilities, $\hat{\Sigma}$:

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} 0.917 & -0.315 & 0.672 \\ -0.315 & 0.401 & -1.379 \\ 0.672 & -1.379 & 6.923 \end{bmatrix}. \tag{4.30}$$

The variance estimate for the Darroch estimate of escapement is:

$$\begin{aligned}\hat{V}(\hat{N}) &= [847 \quad 5981 \quad 3124] \begin{bmatrix} 0.917 & -0.315 & 0.672 \\ -0.315 & 0.401 & -1.379 \\ 0.672 & -1.379 & 6.923 \end{bmatrix} \begin{bmatrix} 847 \\ 5981 \\ 3124 \end{bmatrix} \\ &= 31,405,099\end{aligned} \tag{4.31}$$

The standard error on escapement is $\hat{SE}(\hat{N}) = 5604$ fish and the coefficient of variation is $\hat{CV}(\hat{N}) = 0.1032$.

4.5 Techniques for Open Populations

A major concern for the application of closed-population techniques to salmon escapement is that the population is open to new arrivals and mortalities. Open-population techniques seem more appropriate.

4.5.1 Jolly-Seber (1965) Method

The Jolly-Seber Method involves marking individuals with uniquely identifying tags so that the capture history of each individual may be recorded. In each subsequent sample, previously tagged individuals are recorded and untagged individuals are given tags. Some proportion of these individuals are re-released (which allows for losses on capture) and the procedure is repeated for a total of at least five survey occasions.

The method developed separately by Jolly (1965) and Seber (1965) does not give escapement estimates directly. Instead it estimates abundance at survey times, N_i, and survival between survey times, ϕ_i. These estimates are combined to estimate the number of individuals entering the population ("births" or "arrivals") between surveys and surviving until the next survey. Prior to the start of a run, and after the death of the last fish, there are no adult salmon in the survey area. Therefore, the sum of the arrivals is the escapement.

Seber (1982, page 196) gave the most thorough explanation of the Jolly-Seber Method, after the original papers by Jolly (1965) and Seber (1965), although neither Jolly nor Seber included a derivation of the maximum likelihood estimators in their papers, due to computational limitations. Instead they provided estimates based on the likelihood estimation. Pollock et al. (1990) provided an intuitive development of the Jolly-Seber estimators and this is considered to be the standard reference for

Jolly-Seber experiments. As the method is so well known and well documented, we provide the main points and the details pertinent to the application of the Jolly-Seber Method to estimating salmon escapement.

Seber (1982) considered the number of unmarked individuals at the time of each survey, U_i, to be a fixed parameter. If each unmarked salmon has the same probability of being captured (A-25) and the probability of capture is independent of the capture of the other salmon (A-2), then the number of unmarked salmon in each sample can be considered a Binomial sample of u_i from U_i, with probability equal to that of capture, p_i. Assuming each survey is conducted independently (A-19), the likelihood of the number of unmarked individuals can be written (Seber's equation 5.4):

$$L(\tilde{u_i}|\widetilde{U_i}) = \prod_{i=1}^{k} \left\{ \left(\begin{array}{c} U_i \\ u_i \end{array} \right) p_i^{u_i} (1-p_i)^{U_i - u_i} \right\} \tag{4.32}$$

The first time an individual is captured, it is unmarked. Therefore, conditioning on the number of unmarked fish in sample i, the numbers of salmon with capture histories starting at sample i can be considered realizations from a multinomial distribution. If the probability of being returned to the stream after being captured is the same for all fish (A-21) and independent of capture history (A-22), then the number of individuals returned to the population can be modeled as a realization from a binomial distribution. The conditional likelihood of being recaptured and re-released is (Seber's equation 5.3):

$$\begin{aligned} L(b_\omega, d_\omega|\tilde{u_i}) &= \frac{\prod_{i=1}^{s}(u_i)!}{\prod_\omega (b_\omega)!(d_\omega)!} \prod_{i=1}^{s-1} \left[\chi_i^{R_i - r_i} (\phi_i q_{1+1})^{z_{i+1}} (\phi_i p_{i+1})^{m_{i+1}} \right] \\ &\times \prod_{i=1}^{s} \left[\nu_i^{R_i} (1-\nu_i)^{n_i - R_i} \right] \end{aligned} \tag{4.33}$$

The complete likelihood is the product of Equations 4.32 and 4.33. The abundance at the time of the ith survey, N_i, can be reparameterized as the number of unmarked individuals present, U_i, plus the number of marked individuals, M_i, and is therefore represented in the likelihood, even if not explicitly.

Estimates of abundances and survival probabilities given by Jolly (1965) and Seber (1965) are biased, and, in fact, are not maximum likelihood estimates. Generally nearly-bias-corrected estimates, distinguished from the biased estimates with a “*”, are used (Seber, 1982, page 204):

$$\begin{array}{rcll} \hat{M}_i^* & = & \frac{R_i+1}{r_i+1} z_i + m_i & i = 2, \ldots, k-1 \\ \hat{\phi}_i^* & = & \frac{\hat{M}_{i+1}^*}{\hat{M}_i - m_i + R_i} & i = 2, \ldots, k-2 \\ \hat{N}_i^* & = & \frac{n_i+1}{m_i+1} \hat{M}_i^* & i = 2, \ldots, k-1 \end{array} \tag{4.34}$$

where $\hat{M}_i = \frac{R_i z_i}{r_i} + m_i$. The survival probabilities, ϕ_i^* include losses to the population due to emigration. The likelihood does not contain separate terms for emigration or temporary immigration. Consequently, the Jolly-Seber Method requires the assumption that emigration is permanent (A-28). This assumption is reasonable as long as the study area is defined as the entire spawning area or above a weir or fence. Even without a fence, the assumption is satisfied as long as the start of the study area marks a significant confluence (e.g. lake or span of unsuitable spawning habitat). Often this is the case for chinook, chum, coho, and pink salmon. Sockeye spawn in lakes as well as streams, so steps must be made to ensure that sockeye do not migrate into and out of the stream being studied if the lake-spawning population is being enumerated separately from the stream-spawning population (c.f. Schubert, 2000). More often than not the assumption of permanent immigration can be considered reasonably satisfied.

The likelihood requires that all fish are identically and independently distributed with respect to survival (A-24) and capture probabilities ((A-25) and (A-2)). Implicit in Assumptions (A-25) and (A-2) is the additional assumption that tagging has no influence on survival (A-15) or recapture probability (A-26). If tagged fish are caught more readily than untagged fish (e.g. by getting tags caught in nets), then estimates of abundance will be negatively biased. If tagged fish are more difficult to capture, then estimates will be positively biased. Similarly, abundance estimates will be positively biased if tags are lost or overlooked (A-6).

Another pivotal assumption in the Jolly-Seber Model is that every fish present after a given survey has the same probability of surviving to the next survey (A-27), regardless of tag status or capture history. This assumption is not valid for two reasons: tagging often shortens survival of salmon (Section 4.4.2); fish arriving (and spawning) earlier in the run tend to die earlier than fish that arrive later (Killick, 1955). Manly (1970) and Schwarz et al. (1993) showed through simulation that when the probability of mortality increased with age, Jolly-Seber estimates of abundance and survival were negatively biased, resulting in positively biased estimates of new arrivals and abundance. Schwarz et al. (1993) further showed that if survival is purely random (i.e. not dependent on arrival timing) then the Jolly-Seber estimate of abundance is negatively biased.

The number of new arrivals, is not included in the likelihood and must be estimated as a function of the time-specific abundances and survival probabilities:

$$\hat{B}_i^* = \hat{N}_{i+1}^* - \hat{\phi}_i^*(\hat{N}_i^* - n_i + R_i), \tag{4.35}$$

where n_i is the number of fish captured during the ith survey and R_i is the number of fish returned to the stream after the ith survey. The Jolly-Seber method requires at least three capture periods to provide a single estimate of abundance, and at least two estimates of abundance are required to estimate new arrivals. Consequently, the Jolly-Seber Method requires at least 4 release periods to produce an estimate of escapement.

The initial abundance, $\hat{N}_1^*$ cannot be estimated because there are no tagged individuals to recapture. Since $\hat{N}_1^*$ cannot be estimated, it is clear from Equation 4.35 that $\hat{B}_1^*$ also cannot be estimated. If it is assumed that no mortality takes place prior to the start of the study (A-29), then $\hat{B}_1^* = \hat{N}_1^*$, since all the individuals present at the start of the study were recruited prior to t_1. Seber (1982, page 201) recommended estimating $\hat{N}_1^*$ from the trend in the series of estimates (i.e. extrapolate backward). Alternately, an independent estimate of the initial abundance (e.g. weir or foot survey in the early days of the run) could be used. Schwarz et al. (1993) asserted that in a well designed study, N_1 would be small enough to be reasonably ignored (A-30)

Equation 4.35 shows that the number of new recruits entering after the penultimate survey, B_{k-1} and B_k, cannot be estimated because the abundance at and after the last survey, N_k^* and N_{k+1}^*, is unknown. When applied to escapement, the study is conducted until the end of the run (A-31) so there will be no recruitment after the penultimate survey, B_{k-1}, and B_k will be zero.

The estimate of escapement is the sum of all the new recruits:

$$\hat{N}^* = \hat{B}_1^* + \sum_{i=2}^{k-2} B_i^* + \hat{B}_{k-1} + \hat{B}_k. \tag{4.36}$$

Again $\hat{B}_{k-1}$ and $\hat{B}_k$ are assumed to be zero, i.e. there is no recruitment after the penultimate survey.

Samples for the Jolly-Seber method must be taken instantaneously (A-20) because abundance estimates are assumed to be estimating a discrete and unique number of fish. If the true number of fish present during a survey is changing, the abundance estimate is invalid. Obviously it takes time to perform a survey, but the "instantaneous" assumption is satisfied as long as the samples are taken quickly with respect to arrivals and mortalities of the salmon being surveyed. This seems reasonable for small streams where surveys can be conducted in a few hours, but less reasonable for streams that take more than a day to survey.

Variances and covariances of the Jolly-Seber abundance and survival estimates are derived using the Delta Method. Abundance and survival variances are used to derive the variance estimates for the estimates of births, also using the Delta Method. The variance estimate of total escapement is also derived from the variances and covariances of the birth estimates using the Delta Method. The variances of new arrivals estimates, $\hat{B}_i$, described in Seber (1982, page 203) and by Pollock et al.

(1990, page 21) are:

$$\begin{aligned}\hat{V}(\hat{B}_i|B_i) &= \tfrac{B_i^2(M_{i+1}-m_{i+1})(M_{i+1}-m_{i+1}+R_{i+1})}{M_{i+1}^2}\left(\tfrac{1}{r_{i+1}}-\tfrac{1}{R_{i+1}}\right)\\ &+ \tfrac{M_i-m_i}{M_i-m_i+R_i}\left[\tfrac{\phi_i R_i(n_i-m_i)}{m_i}\right]^2\left(\tfrac{1}{r_i}-\tfrac{1}{R_i}\right)\\ &+ \tfrac{(N_i-n_i)(N_{i+1}-B_i)(n_i-m_i)(1-\phi_i)}{n_i(M_i-m_i+R_i)}\\ &+ N_{i+1}(N_{i+1}-n_{i+1})\left(\tfrac{n_{i+1}-m_{i+1}}{n_{i+1}m_{i+1}}\right)\\ &+ \phi_i^2 N_i(N_i-n_i)\left(\tfrac{n_i-m_i}{n_i m_i}\right)\end{aligned} \tag{4.37}$$

for $i = 2, 3, ..., k-2$. The covariances of new arrivals estimates as given by Seber (1982, page 203) and by Pollock et al. (1990, page 21) are:

$$\begin{aligned}\hat{Cov}(\hat{B}_i, \hat{B}_{i+1}) &= -\phi_{i+1}(N_{i+1}-n_{i+1})(1-\tfrac{m_{i+1}}{n_{i+1}})\\ &\quad\times\left[\tfrac{B_i R_{i+1}}{M_{i+1}}\left(\tfrac{1}{r_{i+1}}-\tfrac{1}{R_{i+1}}\right)+\tfrac{N_{i+1}}{m_{i+1}}\right]\end{aligned} \tag{4.38}$$

and

$$\hat{Cov}(\hat{B}_i, \hat{B}_j) = 0 \quad \text{for } |i-j| > 1. \tag{4.39}$$

The variance estimate for the total escapement is:

$$V(\hat{N}|\mathbf{B}) = V(\hat{N}_1) + \sum_{i=2}^{k-2} V(\hat{B}_i) + \sum_i \sum_j Cov(\hat{B}_i, \hat{B}_j), \quad i \neq j. \tag{4.40}$$

The variance and covariance estimates in Equations 4.37 and 4.38 are asymptotic and only valid for large sample sizes. $V(\hat{N}_1)$ will depend on the method used to estimate initial abundance, N_1, and is assumed to be independent of the subsequent recruitment estimates (A-32). If initial abundance, $\hat{N}_1$, is assumed to be zero (A-30) or is counted without error, then $V(\hat{N}_1) = 0$.

Jolly-Seber Example

The data for this example were taken from Table 1 of Schwarz and Arnason (1996) for the repeated mark and recapture of male salmon to the Chase River in British Columbia and are shown in Table 4.2. Parameter estimates were computed using

Table 4.2: Data, computed statistics, and variance and covariance estimates for the Jolly-Seber Method example. Data taken from Table 1 of Schwarz and Arnason (1996).

Survey (i)	1	2	3	4	5	6	7	8
n_i	85	35	97	84	67	51	39	18
m_i	0	12	14	25	39	28	6	14
R_i	85	34	72	78	56	37	29	—
r_i	28	19	31	34	14	5	7	—
z_i	—	16	21	27	22	8	7	—
$\hat{M}_i^*$	—	40.00	61.91	85.94	122.60	78.67	32.25	—
$\hat{\phi}_i^*$	—	0.11	0.06	0.06	0.06	0.10	—	—
$\hat{N}_i^*$	—	110.77	404.45	280.97	208.42	141.06	184.29	—
$\hat{B}_i$	—	392.30	259.97	192.74	128.87	171.79	—	—
$\hat{V}(\hat{B}_i)$	—	11844.13	2866.92	1867.07	2268.10	8239.26	—	—
$\hat{Cov}(\hat{B}_i, \hat{B}_{i+1})$	—	-3525.09	-347.07	-91.98	-210.91	—	—	—

Equations (4.34) and (4.35). Variances and covariances were calculated using Equations (4.37) and (4.38).

Assuming that the number of salmon in the stream before the study was negligible and that the study is carried out until there were no longer new recruits (for at least the last two surveys), the Jolly-Seber estimate of male escapement is the sum of the male recruits:

$$\hat{N} = \sum_{i=2}^{6} \hat{B}_i = 1146 \quad \text{male salmon.} \tag{4.41}$$

The escapement variance estimate is the total of the variances and covariances of each of the "birth" estimates. The covariances of non-consecutive birth estimates are zero. The covariances of consecutive birth estimates must be counted twice (because the covariance of B_i and B_{i+1} is the same as the covariance of B_{i+1} and B_i). The resulting Jolly-Seber variance estimate is:

$$\begin{aligned} \hat{V}(\hat{N}) &= 11844.13 + 2866.92 + 1867.07 + 2268.10 + 8239.26 \\ &+ 2(-3525.09 + -347.07 + -91.98 + -210.91) \\ &= 27085.47 + 2(-4175.051) \\ &= 18735.37 \end{aligned} \tag{4.42}$$

The estimate of the standard error is $\hat{SE}(\hat{N}) = 136.88$ and the coefficient of variation is estimated by $\hat{CV}(\hat{N}) = 0.1195$.

4.5.2 New Arrivals Adjustment (Schwarz et al., 1993)

The original Jolly-Seber Method uses estimates of the number of new individuals entering the population between surveys and surviving to the next survey. These estimates of new arrivals underestimate the true number of individuals entering the population because they do not account for individuals that enter and die before the next sampling period. For example, an individual that enters the study area on the day after a weekly survey and dies five days later will not be included in the original Jolly-Seber (1965) Model.

Schwarz et al. (1993) provided new arrival adjustments factors to account for arrivals that die prior to the next survey. In addition to the assumptions required for the Jolly-Seber Method, the new arrivals adjustment requires assumptions about the distribution of arrivals between surveys. Recruitment patterns determine how the probability of survival between periods should be applied. Crosbie and Manly (1985) assumed that the distribution of new arrivals was uniform across the entire study. Sykes and Botsford (1986) assumed that all recruitment occurred at the midpoint of the interval. Schwarz et al. (1993) developed a framework whereby the distribution of arrivals could be modeled separately for any period. The pattern of recruitment could be modeled to occur six ways (Table 4.3): 1) immediately after the previous survey, 2) declining over the interval, 3) at the midpoint of the interval, 4) uniformly over the interval, 5) increasing over the interval or 6) immediately before the sample. For example, arrivals occurring prior to the peak of the run might be modeled as increasing over time between surveys, but after the peak, arrivals might be modeled as decreasing over the interval. The applicability of the adjustment factors and the accuracy of the new recruits estimate depend on how well the assumed distribution of arrivals fits the data. Consequently the direction and magnitude of any biases are unknown without additional information regarding arrival patterns.

Estimates of arrivals, $\hat{B}_i'$, are the product of the Jolly-Seber estimates of the net number of new individuals (traditional J-S births), and adjustment factors, which are functions of the survival probabilities. MLE estimates, $\hat{B}_i$ and $\hat{\phi}_i$ given by Seber (1982), or the bias-corrected estimates $\hat{B}_i^*$ and $\hat{\phi}_i^*$ may be used. The bias-corrected estimates performed better in a simulation than the MLE's (Schwarz et al., 1993):

$$\hat{B}_i' = \hat{B}_i^* \times f(\hat{\phi}_i^*)$$

where $i = 2, 3, ..., k-1$ and $f(\hat{\phi}_i^*)$ is the adjustment factor appropriate for the assumed recruitment patterns. Adjustment factors are derived from the formula for the net number of arrivals on the interval from survey i to $i+1$. Since the estimate pertains

Table 4.3: Distribution of recruitment between surveys and corresponding adjustment factors for estimating total recruits from estimated net recruits. Taken from Table 3 of Schwarz et al. (1993)

Arrival pattern	$f(\phi_i)$
All after previous	$\frac{1}{\phi_i}$
Declining	$\frac{\ln(\phi_i)^2}{2+2\phi_i\ln(\phi_i)-2\phi_i}$
All at midpoint	$\frac{1}{\sqrt{\phi_i}}$
Uniform	$\frac{\ln(\phi_i)}{\phi_i-1}$
Increasing	$\frac{\ln(\phi_i)^2}{2-2\phi_i\ln(\phi_i)-2\phi_i}$
All before sample	1

only to the interval, the endpoints can be reset to zero and one. Assuming uniform mortality over the interval (A-35), the probability that a fish entering a stream at time t $(0 < t < 1)$ will survive to the next survey time at $t = 1$ is $\phi^{(1-t)}$. If a fish arrives at the start of the interval, $t = 0$, then the fish has ϕ probability of survival. If the fish enters at the end of the interval, $t = 1$, then it has $\phi^{(1-1)} = 1$ probability of surviving to the end of the interval, because it already has. If the new arrivals follow the probability distribution $f(t)$ then the net number of arrivals is the product of the total (gross) number of new recruits, the distribution function of arrivals, and the survival probability at time of entry:

$$B_i = \int_0^1 B_i^* f(t) \phi_i^{(1-t)} dt. \tag{4.43}$$

Solving Equation (4.43) for B_i^* produces the net new arrivals times the adjustment factor:

$$B_i^* = B_i f(\phi_i). \tag{4.44}$$

For example, if recruitment is assumed uniform, $f(t) = 1$, Equation (4.43) becomes:

$$B_i = \int_0^1 B_i^* \phi_i^{(1-t)} dt. \tag{4.45}$$

Let $u = 1 - t$ and $du = -1dt$ such that $dt = -1du$. Now we have:

$$B_i = B_i^* \int_0^1 -\phi_i^{(u)} du, \tag{4.46}$$

which is:

$$B_i = B_i^* \Big[\frac{-\phi_i^{(u)}}{\ln \phi_i}\Big]\Big|_0^1 = B_i^* \Big[\frac{-\phi_i^{(1-t)}}{\ln \phi_i}\Big]\Big|_0^1, \tag{4.47}$$

or equivalently:

$$B_i = B_i^* \Big[\frac{-\phi_i^0}{\ln \phi_i} - \frac{-\phi_i^1}{\ln \phi_i}\Big] = B_i^* \Big[\frac{\phi_i - 1}{\ln \phi_i}\Big]. \tag{4.48}$$

for $t = 0$ to $t = 1$. Therefore, the total number of new arrivals is:

$$B_i^* = B_i \Big[\frac{\ln \phi_i}{\phi_i - 1}\Big], \tag{4.49}$$

and the adjustment factor for uniform arrival of new arrivals is:

$$f(\phi_i) = \left[\frac{\ln \phi_i}{\phi_i - 1}\right]. \tag{4.50}$$

The traditional Jolly-Seber Method does not allow for an estimate of the number of new recruits just after the first survey, B_1. Schwarz et al. (1993) provided an additional adjustment to estimate B_1. They noted that N_2 is made up of the net number of recruits prior to the first survey, B_0, times the survival from the first to the second surveys, ϕ_1, plus the number of recruits following the first survey, B_1:

$$N_2 = B_0\phi_1 + B_1 \tag{4.51}$$

The estimate of N_2 is multiplied by the adjustment factor, $f(\hat{\phi}_1^*)$, to estimate the gross number of new individuals present prior to the second survey.

The adjusted estimate of escapement, $\hat{N}'$, is now the sum of the adjusted recruitment estimates:

$$\hat{N}' = (\hat{B}_0\hat{\phi}_1 + \hat{B}_1)f(\hat{\phi}_1^*) + \sum_{i=2}^{k-1} \hat{B}_i' + \hat{B}_k' \tag{4.52}$$

where $\hat{B}_k'$ is assumed to be zero (A-31)

Schwarz et al. (1993) pointed out that the estimate in Equation 4.52 is slightly negatively biased because $N_2 f(\hat{\phi}_1)$ estimates $B_0\phi_1 f(\hat{\phi}_1)$ which is less than B_0, and because it ignores $\hat{B}_s'$. This bias is expected to be small, especially if the study commences near the start of a run (A-30), and is conducted until fish are no longer entering the stream (A-31)

Schwarz et al. (1993) provided asymptotic variance and covariance estimates for net arrivals, abundances, and survival probabilities in their Appendix D. The estimates differed from those given by Pollock et al. (1990) because Schwarz et al. (1993) conditioned on total escapement rather than each individual arrival. The escapement variance estimated using the new arrivals adjustment factors can be calculated using the Delta Method, and will be larger than those of the traditional Jolly-Seber Method due to the uncertainty contributed by the adjustment factors.

4.5.3 Schwarz and Arnason (1996) Method

New arrivals are not included directly in the classic Jolly-Seber likelihood. Consequently, it is not possible to constrain the arrival estimates to be positive. Negative estimates of arrival are possible but not biologically meaningful. Often negative estimates are truncated, or rounded up to zero, resulting in a positively biased escapement estimator (Schwarz et al., 1993). Schwarz and Arnason (1996) developed a way to include new arrivals into the likelihood by considering the total escapement as being multinomially distributed between surveys. They credit Crosbie and Manly (1985) for the multinomial portion of the model.

Conditioning on the total escapement, N, the number of new arrivals available at tagging occasion i is distributed multinomially, with the probability of an individual from the population, N, arriving between the ith and the ith+1 tagging occasion is β_i:

$$\{B_1, ..., B_{k-1}\} \sim \text{Multinomial}(N; \beta_1, ..., \beta_{k-1}).$$

The number of individuals in the study area prior to the start of the study, B_0, is $B_0 = N - \sum_{i=1}^{k-1} B_i$, assuming that the study is conducted until there are no more new arrivals and B_k=0 (A-31).

The proportion of unmarked individuals available for capture at occasion $i + 1$, Ψ_{i+1}, is the proportion of the new arrivals available from the occasion before (Ψ_i) that did not get captured (1-P_i), but survived (ϕ_i), plus the proportion of new arrivals (B_i):

$$\Psi_{i+1} = \Psi_i(1 - P_i)\phi_i + \beta_i \qquad \text{where} \qquad \Psi_1 = \beta_0. \tag{4.53}$$

The number of unmarked individuals in each sample also can be considered to be distributed as a multinomial:

$$\{u_1, ..., u_k\} \sim \text{Multinomial}(N; \Psi_1 P_1, ..., \Psi_k P - k) \tag{4.54}$$

Schwarz and Arnason (1996) factored the likelihood of the unmarked individuals into two parts, which they called L'''_{1a} and L'''_{1b}, by conditioning on the total number of unmarked individuals captured over the course of the study:

$$
\begin{aligned}
L_{1a} \times L_{1b} &= \begin{pmatrix} N \\ u. \end{pmatrix} \left(\sum_{i=1}^{k} \Psi_i P_i \right)^{u.} \left(1 - \sum_{i=1}^{k} \Psi_i P_i \right)^{N-u.} \\
&\times \begin{pmatrix} u. \\ u_1, u_2, ..., u_k \end{pmatrix} \prod_{i=1}^{k} \left(\frac{\Psi_i P_i}{\sum_{i=1}^{k} \Psi_i P_i} \right)^{u.} \qquad (4.55)
\end{aligned}
$$

Equation (4.55) describes the distribution of the initial captures and is only the first portion of the full likelihood. The second portion, L_2, describes the losses (deaths) on capture and is typically modeled using the binomial distribution:

$$
L_2 = \prod_{i=1}^{k} \begin{pmatrix} n_i \\ l_i \end{pmatrix} (1 - \nu_i)^{l_i} (\nu_i)^{n_i - l_i}, \qquad (4.56)
$$

which requires that the probability of being returned to the stream is the same for all fish (A-21), and be independent of capture history (A-23)

The third portion of the likelihood, L_3, describes the recapture probability, and can be factored into conditionally independent binomials:

$$
L_3 = \prod_{i=1}^{k-1} \begin{pmatrix} R_i \\ r_i \end{pmatrix} (\lambda_i)^{r_i} (1 - \lambda_i)^{R_i - r_i} \prod_{i=2}^{k-1} \begin{pmatrix} m_i + z_i \\ m_i \end{pmatrix} (\tau_i)^{m_i} (1 - \tau_i)^{z_i} \qquad (4.57)
$$

The complete likelihood is the product of the components described in Equations (4.55)-(4.57):

$$
L = L_{1a}''' \times L_{1b}''' \times L_2 \times L_3. \qquad (4.58)
$$

Schwarz and Arnason (1996) described the steps necessary to maximize Equation (4.58) numerically. They also described software that would perform the maximization. Although the theoretical framework is different, the maximum likelihood estimates of survival ($\tilde{\phi}$), capture probability ($\tilde{P}$), net births ($\tilde{B}$), and periodic abundances ($\tilde{N}$) are the same as those given by the Jolly-Seber Method (Section 4.5.1). The major difference is that the arrivals are constrained to be positive and will be biologically meaningful.

Variances of the survival and capture probabilities remain unchanged from those given by Pollock et al. (1990). Variances of net births and abundances are larger due to the treatment of the new arrivals as random variables, and can be derived using the inverse of the Fisher Information matrix from the likelihood (Schwarz and Arnason, 1996).

4.5.4 Manly-Parr (1968) Method

One of the drawbacks of the Jolly-Seber method is that it requires the assumption that the probability of a fish dying is independent of the arrival time of the fish (A-27). If fish are assumed to have the same stream life, then a fish arriving early in the study would have a lower probability of survival than a fish arriving later in the study and Assumption (A-27) is violated. Manly and Parr (1968) derived estimators for periodic abundances, survival proportions, and numbers of births that did not rely on assumptions about survival probabilities. Instead, Manly and Parr (1968) assumed that the probability of detection was the same for all members of the population so the expected value of the number captured was the abundance, N_i, times the detection probability, p_i:

$$E(n_i|N_i) = N_i p_i. \tag{4.59}$$

Abundance at the time of the ith survey is:

$$\hat{N}_i = \frac{n_i}{\hat{p}_i}. \tag{4.60}$$

The detection probability for the ith survey can be estimated using the number of marked fish sampled, m_i, from the total number of marked fish present, M_i:

$$\hat{p}_i = \frac{m_i}{M_i}. \tag{4.61}$$

The number of marked fish in the population at i can be estimated as the sum of the marked fish captured, m_i, and those captured any time before i and after i, but not at i, z_i. The Manly-Parr estimate of survey abundance is:

$$\hat{N}_i = \frac{n_i(m_i + z_i)}{m_i}, \tag{4.62}$$

which Manly (1969) showed was the maximum likelihood estimate using a multinomial distribution. The four possible categories of fish known to be alive at the time of the ith sample are: captured with a tag; captured without a tag; not captured with a tag; and not captured without a tag (Table 4.4). Equation (4.62) is a Petersen type estimator and is biased (Manly, 1969). Seber (1982, page 234) recommended the Chapman modification:

$$\hat{N}_i^* = \frac{(n_i + 1)(m_i + z_i + 1)}{(m_i + 1)} - 1, \tag{4.63}$$

which is approximately unbiased for abundance at the time of the ith survey.

Table 4.4: The number of fish within each of the four categories possible for fish alive at the time of the ith sample. Recreated from (Manly, 1969, Table 1). See also Seber (1982, page 233).

	Captured in ith survey	Not captured in ith survey	Total
Having tag	m_i	z_i	$m_i + z_i$
Not having tag	$n_i - m_i$	$N_i - z_i - n_i$	$N_i - m_i - z_i$
Total	n_i	$N_i - n_i$	N_i

Like the Jolly-Seber method, the Manly-Parr method requires that the probability of recapture be independent of tag status (A-7) and that fish enter the spawning area only once (A-28). If there is handling-induced mortality, i.e. a violation of Assumption (A-7), then a marked fish will have an increased chance of recapture after the ith survey if it is not captured during the ith survey (Seber, 1982, page 235). If tagged fish leave the spawning area and are recaptured after they return, then the number of tagged fish alive during any survey while the fish are gone will be over-estimated, the detection probability will be under-estimated, and abundance estimates will be positively biased.

While the Manly-Parr abundance estimates are simpler than those derived using the Jolly-Seber method, the birth estimates are constructed in exactly the same way:

$$\hat{B}_i = \hat{N}_{i+1} - \hat{\phi}_i \hat{N}_i, \tag{4.64}$$

where $\hat{\phi}_i$ is the estimate of survival from the ith to the ith+1 survey. The survival estimates are (Seber, 1982, 236):

$$\hat{\phi}_i = \frac{m_{i,i+1}}{n_i p_{i+1}}, \tag{4.65}$$

where $m_{i,i+1}$ is the number of individual fish caught in the ith and ith+1 surveys.

Abundances for the first and last survey cannot be estimated, so arrivals prior to the first, last, and penultimate surveys also cannot be estimated. The first and last recruits adjustment used in the Jolly-Seber method must also be applied to Manly-Parr estimates.

Manly and Parr (1968) found asymptotic means and variances for abundances, but not for survival proportions or new arrivals (Seber, 1982, page 237). The variance for $\hat{N}_i$ (Seber, 1982) is:

$$\hat{V}(\hat{N}_i) = \frac{\hat{N}_i(z_i)(n_i - m_i)}{m_i^2}. \tag{4.66}$$

The standard error is:

$$\hat{SE}(\hat{N}_i) = \frac{\sqrt{\hat{N}_i(z_i)(n_i - m_i)}}{m_i}, \tag{4.67}$$

and the coefficient of variation is:

$$\hat{CV}(\hat{N}_i) = \frac{\sqrt{(z_i)(n_i - m_i)}}{m_i \sqrt{\hat{N}_i}}. \tag{4.68}$$

Manly (personal communication) recommended a bootstrap procedure to estimate variances of survivals and new arrivals, since no asymptotic variances have yet been derived.

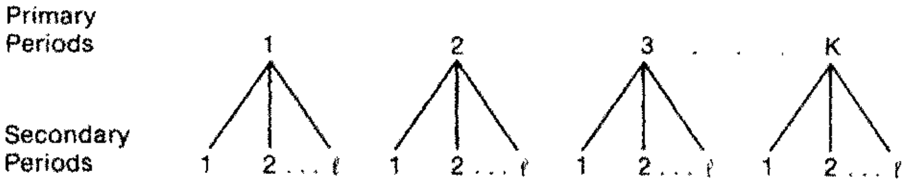

Figure 4.1: A schematic of the Robust Design Method of sampling. Taken from Figure 7.1 of Pollock et al. (1990).

4.6 Robust Design Method (Pollock et al., 1990)

Pollock (1982) developed a mark-recapture technique to incorporate individuals with unequal probabilities of capture. The method is a hybrid of both open and closed population techniques, and is best conveyed by Figure 4.1. The salmon run is divided into k primary periods which are then divided into l secondary sampling events. The population is assumed to be closed for the duration of the secondary sampling and the abundance can be estimated using closed population techniques. Between primary periods, the population is allowed to fluctuate and the between-period survival is estimated using the Jolly-Seber Method (See Section 4.5.1). The data from the secondary samples is pooled to make "one" sample, where an individual is "recaptured" for the primary sample if it is seen alive during any of the secondary samples within that primary period.

The formulae for estimates of new arrivals are the same as in Equation 4.35 for the Jolly-Seber Method (Section 4.5.1):

$$\hat{B}_i = \hat{N}_{i+1} - \hat{\phi}_i(\hat{N}_i - n_i + R_i).$$

The abundances, N_i, need to be estimated on a short enough time frame so that the population can be assumed closed to arrivals and mortalities.

Under the Robust Method, the abundance at the first primary sampling occasion can be estimated, and therefore B_1 is estimable as well. Assuming that surveys are conducted until there are no new arrivals ($B_s = 0$) (A-31), the estimate of escapement is:

$$\hat{N} = \sum_{i=1}^{s-1} B_i. \tag{4.69}$$

The same adjustments that can be applied to the Jolly-Seber Method can also be applied to the primary periods of the Robust Method.

Variances of the $\hat{N}_i$ are derived according to the method used for the secondary (closed population) sampling. Survival estimates, $\hat{\phi}_i$, are derived according to the Cormack-Jolly-Seber Method in Sections 3.4.9 and 4.5.1. Because the survival and abundance estimates are derived using different methods, Pollock (1982) assumed that their covariances were zero. Thus the variance of the birth estimates, $\hat{B}_i$, can be written:

$$V(\hat{B}_i) = V(\hat{N}_{i+1}) + V(\phi_i N_i) + V\left(\phi_i(n_i - R_i)\right) \tag{4.70}$$

Using the "Exact variance formula" for a product of two independent random variables (Seber 1982 pg. 9) the middle term is:

$$V(\phi_i N_i) = \phi_i^2 V(N_i) + N_i^2 V(\phi_i) + V(\phi_i)V(N_i). \tag{4.71}$$

Therefore the variance, as given by Pollock (1982) is:

$$V(\hat{B}_i) = V(\hat{N}_{i+1}) + \phi_i^2 V(N_i) + N_i^2 V(\phi_i) + V(\phi_i)V(N_i) + (n_i - R_i)^2 V(\phi_i) \tag{4.72}$$

The variance estimate is derived by replacing parameters with their estimated values and negating the product of the abundance and survival variances:

$$\hat{V}(\hat{B}_i) = \hat{V}(\hat{N}_{i+1}) + \hat{\phi}_i^2 \hat{V}(N_i) + \hat{N}_i^2 \hat{V}(\hat{\phi}_i) - \hat{V}(\hat{\phi}_i)\hat{V}(\hat{N}_i) + (n_i - R_i)^2 \hat{V}(\hat{\phi}_i) \tag{4.73}$$

Because covariances of the new arrivals estimates depend only on the estimates of abundances and not on survival probabilities, covariances of $\hat{B}_i$ are the same as in Equation 4.38 where $\hat{N}_i$ are estimated from secondary samples:

$$\begin{aligned} \hat{Cov}(\hat{B}_i, \hat{B}_{i+1}) &= -\phi_{i+1}(N_{i+1} - n_{i+1})(1 - \tfrac{m_{i+1}}{n_{i+1}}) \\ &\quad \times \left[\tfrac{B_i R_{i+1}}{M_{i+1}} \left(\tfrac{1}{r_{i+1}} - \tfrac{1}{R_{i+1}}\right) + \tfrac{N_{i+1}}{m_{i+1}}\right] \end{aligned} \tag{4.74}$$

and

$$\hat{Cov}(\hat{B}_i, \hat{B}_j) = 0 \quad \text{for } |i-j| > 1. \tag{4.75}$$

Finally, the variance of the robust design escapement estimator is:

$$V(\hat{N}) = \sum_{i=1}^{s-1} V(\hat{B}_i) + \sum_i \sum_j Cov(\hat{B}_i, \hat{B}_j), \quad i \neq j \tag{4.76}$$

Kendall et al. (1995) developed a maximum likelihood framework for the robust method, which allows for variation in catchability over time. In simulation, the Likelihood Robust Method estimates had smaller coefficients of variation than the Jolly-Seber and Robust Method estimates when the correct likelihood model was fitted to the simulated data. Kendall et al. (1995) did not address issues of likelihood model selection or the performance of the Likelihood Robust Method when the underlying model was not correct. Without empirical comparisons, it is impossible to say how well the Likelihood Robust Estimator performs in practice.

4.7 Change In Ratio Method

Hatchery fish are often marked prior to release as smolts by removing the adipose fin (i.e. fin clipping). The lack of an adipose fin on an adult salmon alerts harvesters to the potential presence of a coded wire tag (CWT) and distinguishes hatchery from wild fish. Hatchery fish home to the hatchery and are preferentially removed by hatchery workers, such that the relative abundance (ratio) of hatchery to wild fish changes upstream of the hatchery (intervention). Using the number of marked (hatchery) fish removed from the population and estimates of the relative abundances before (downstream) and after (upstream) the harvest, the total number of spawners can be estimated. The Change In Ratio (CIR) method can be thought of as a Pooled Petersen estimator where the population is open to emigration (loss of hatchery fish). The estimate applies to the population prior to the harvest. It should be noted, however, that CIR is not necessarily a mark-recapture method. In other applications the ratio examined is of males to females before and after male-only harvest (Skalski and Millspaugh, 2006).

Let M denote the true number of adipose fin-clipped hatchery (marked) fish returning to the study stream. Let U be the number of wild (unmarked) fish such that the total escapement is the sum of the two, $N = M + U$. Assume that each fish has the same probability of capture within a sampling period (A-2). If fish are sampled so that they are caught once below the hatchery and once above the hatchery, then they are sampled without replacement and the likelihood is:

$$L(N, M|n_i, m_i) = \frac{\binom{M}{m_1}\binom{N-M}{n_1-m_1}}{\binom{N}{n_1}} \bullet \frac{\binom{M+\Delta M}{m_2}\binom{N-M+\Delta U}{n_2-m_2}}{\binom{N+\Delta M+\Delta U}{n_2}}. \tag{4.77}$$

A sample of n_1 fish captured below the hatchery has m_1 marked fish, taken out of the M total marked fish. The expected value of the number of fin-clipped fish in the first sample is:

$$E(m_1) = n_1 \left(\frac{M}{N}\right). \tag{4.78}$$

A second, independent (A-32) sample of n_2 fish captured above the hatchery has m_2 hatchery fish, drawn out of the M total hatchery fish plus the change in the number of hatchery fish, ΔM. Since the hatchery fish are removed from the population, ΔM is negative, and assumed known without error (A-33). The $n_2 - m_2$ unmarked fish are drawn from the original number of unmarked fish U plus any changes, ΔU. If any wild fish are harvested, ΔU will be negative, and is assumed known without error (A-34). If no wild fish are removed, ΔU will be zero. The expected value of the number of marked fish in the second sample is:

$$E(m_1) = n_1 \left(\frac{M+\Delta M}{N+\Delta M+\Delta U}\right). \tag{4.79}$$

Let $P_1 = \left(\frac{M}{N}\right)$ be the probability that a fish caught in the first sample is a hatchery fish. From Equation (4.78) the number of marked fish is equal to the total escapement times the proportion of the escapement that is marked, $M = NP_1$. Denote the probability that a fish caught in the second sample is a hatchery fish P_2:

$$P_2 = \left(\frac{M + \Delta M}{N + \Delta M + \Delta U} \right), \tag{4.80}$$

which is the proportion of hatchery fish present during the second sample. Substituting NP_1 into Equation (4.80) yields:

$$P_2 = \frac{NP_1 + \Delta M}{N + \Delta M + \Delta U}. \tag{4.81}$$

Solving Equation 4.81 for N shows:

$$N = \frac{(\Delta M + \Delta U)P_2 - \Delta M}{(P_1 - P_2)}, \tag{4.82}$$

which is the same as that given in Equation (2) of Paulik and Robson (1969) where $R = -(\Delta M + \Delta U)$ and $R_x = -\Delta M$.

Downstream of the hatchery the probability of being marked is estimated by:

$$\hat{P} = \frac{m_1}{n_1}$$

and upstream of the hatchery the probability of being marked is estimated by:

$$\hat{P} = \frac{m_2}{n_2}.$$

Estimates of the proportion marked are substituted into Equation (4.82) to estimate escapement:

$$\hat{N} = \frac{(\Delta M + \Delta U)\hat{P}_2 - \Delta M}{(\hat{P}_1 - \hat{P}_2)}. \tag{4.83}$$

Equation (4.83) is the maximum likelihood estimate for total escapement, including both hatchery and wild salmon. It is unbiased to the first order Taylor Series expansion, but is positively biased to the third and higher order terms.

Paulik and Robson (1969) discussed bias contributed to the estimator when the assumption of a known number of subtractions or additions is violated ((A-33) and

(A-34)). If either hatchery or wild fish were harvested but not reported, or died from fishing wounds, then the estimate given in Equation (4.83) will underestimate the true escapement (Skalski and Millspaugh, 2006). The assumption of a closed population, excluding harvest of the hatchery fish, can be relaxed if the natural mortality of wild fish is the same as that of hatchery fish (Paulik and Robson, 1969). Unless there are significant differences in run timing or genetics between hatchery and wild fish, this seems to be a reasonable assumption. If all fish are assumed to have equal survival less than 100%, Equation (4.83) will estimate escapement prior to removal of hatchery fish (Skalski and Millspaugh, 2006). If detection probabilities of hatchery fish are different from those of wild fish, the estimate of escapement will be biased (Skalski and Millspaugh, 2006).

The variance of the estimator can be approximated to the first order using the Delta Method, or the "propagation-of-error" method (Paulik and Robson, 1969):

$$\hat{V}(\hat{N}) = \frac{1}{(\hat{P}_1 - \hat{P}_2)^2}\Big[\hat{V}(\hat{P}_1)\hat{N}^2 + \hat{V}(\hat{P}_2)(N + \Delta M + \Delta U)^2\Big], \tag{4.84}$$

where the estimated variance of $\hat{P}_1$ is:

$$\hat{V}(\hat{P}_1) = \frac{P_1(1-P_1)}{n_1}\left(\frac{N-n_1}{N-1}\right), \tag{4.85}$$

and the estimated variance of $\hat{P}_2$ is:

$$\hat{V}(\hat{P}_2) = \frac{P_2(1-P_2)}{n_2}\left(\frac{N+\Delta M+\Delta U-n_2}{N+\Delta M+\Delta U-1}\right). \tag{4.86}$$

Equations (4.85) and (4.86) require that salmon be sampled without replacement. Equation (4.84) requires the assumption that estimates of P_1 and P_2 are independent (A-32), which is most often the case (Paulik and Robson, 1969).

Skalski and Millspaugh (2006) described a program called USER (User Specified Estimation Routine) which computes the likelihood-based CIR estimates of abundance and variance. Variances are computed using profile-likelihoods.

The description of the Change-In-Ratio method applies only if every hatchery fish is marked prior to release. If only a known proportion of hatchery releases are marked, $\hat{P}_1$ and $\hat{P}_2$ must be adjusted, and the variance estimate of Equation (4.84) will underestimate the true uncertainty of $\hat{N}$. The true uncertainty will increase as the marked fraction of hatchery releases decreases. For a thorough discussion on the Change-In-Ratio method where the proportion of marked hatchery releases is less than 100%, see Hankin (1982).

Change-In-Ratio Example

The data for this example are taken from Alexandersdottir (2005). The ratio of marked chinook salmon returning to the Nisqually River, Washington, was estimated as the proportion marked in a sample from the commercial fishery harvest downstream of the hatchery (Table 4.5). Both marked and unmarked fish were harvested in the hatchery. The ratio of marked chinook above the hatchery was estimated as the marked ratio in a test fishery conducted by the Nisqually sampling staff (Alexandersdottir, 2005).

Table 4.5: Data used in the Change-In-Ratio Method example, from chinook salmon in the Nisqually River, Washington in 2003. Data taken from Alexandersdottir (2005).

Fish checked for mark at $t = 1$	n_1	14,510
Marked fish found at $t = 1$	m_1	11,804
Marked fish taken at hatchery	ΔM	-8,617
Unmarked fish taken at hatchery	ΔU	-1,203
Fish checked for mark at $t = 2$	n_2	73
Marked fish found at $t = 2$	m_2	13
Total run size (harvest + escapement)	N_1	26,707

Proportions of marked salmon before and after the hatchery harvest are:

$$\hat{P}_1 = \frac{11,804}{14,510} = 0.81351 \quad \text{and} \quad \hat{P}_2 = \frac{13}{73} = 0.17808, \tag{4.87}$$

respectively. The CIR estimate of escapement is derived via Equation (4.83) as:

$$\hat{N} = \frac{(-8,617 + -1,203)(0.8135) + 8,617}{(0.8135 - 0.1781)} = 10,809 \text{ chinook.} \tag{4.88}$$

Because the proportion of marked chinook was estimated from harvested fish, the variance of $\hat{P}_1$ must be estimated using the number of salmon in the total run that could be commercially harvested, N_1. The variance estimate of $\hat{P}_1$ is:

$$\hat{V}(\hat{P}_1) = \frac{0.8135(1 - 0.8135)}{14,510}\left(\frac{26,707 - 14,510}{26,707 - 1}\right) = 0.0000047753. \tag{4.89}$$

The variance estimate for $\hat{P}_2$ is:

$$\hat{V}(\hat{P}_2) = \frac{0.1781(1 - 0.1781)}{73}\left(\frac{10,809 - 8,617 - 1,203 - 73}{10,809 - 8,617 - 1,203 - 1}\right) = 0.0018589. \tag{4.90}$$

The variance of the escapement estimate is:

$$\begin{aligned} \hat{V}(\hat{N}) &= \tfrac{1}{(0.8135-0.1781)^2}\Big[(0.000004775)(10,809)^2 + (0.001859)(989)^2\Big] \\ &= 5,883.8 \end{aligned} \tag{4.91}$$

The standard error is $\hat{SE}(\hat{N}) = 76.706$ and the coefficient of variation is $\hat{CV}(\hat{N}) = 0.0070966$. Note that the sample used to estimate the proportion of marked fish at the start of the study, n_1, is very large. Consequently, the variance of $\hat{P}_1$ is very small. If the same $\hat{P}_1$ had been generated from the escapement only (after the commercial harvest had taken place) from 1,451 captured chinook, the variance on $\hat{P}_1$ would have been 0.000090530, and the overall variance estimate would have been 30,697 chinook.

4.8 Discussion and Comparison of Mark-Recapture Techniques

When applied to salmon escapement, most Mark-Recapture methods fail due to violations of assumptions. There are cases, however, when assumptions can be met, and an unbiased estimate of escapement is possible.

The Pooled Petersen estimator has been shown to be positively biased both statistically and empirically. Corrections proposed by Chapman (1951) (cited in Seber, 1982) and Bailey (1951) provide nearly-unbiased estimates. The Pooled Petersen estimator is often recommended for ease of use and robustness to assumption violations (e.g. Schubert, 2000). Of all the mark-recapture methods considered, the Pooled Petersen Estimator has the smallest known variance because it is estimating fewer parameters. However, the Pooled Petersen Estimator is biased due to violations of the assumptions and is not recommended (Table 4.6).

The Schaefer Estimator (Section 4.4.2) incorporates different tagging and recovery probabilities but requires assumptions that are unreasonable to ensure. The Schaefer Estimator requires nearly all the assumptions of the Darroch Method (Section 4.4.3) , but does not provide a variance estimate (Table 4.6). The Darroch Method does provide a variance estimate so the Darroch method supersedes the Schaeffer Method.

The Darroch Method is the best mark-recapture method available to estimate salmon escapement, but not as it is traditionally used (Table 4.6). Traditionally, salmon are tagged while alive and recovered as carcasses, with the assumption that bodies of all the salmon remain in the study area until the study is complete (e.g. Maselko et al., 2003). However, carcasses and moribund fish are frequently washed downstream and out of the study area (Simpson, 1984; Johnston et al., 1986; Schubert, 2000). Entire recovery strata can be lost and the total escapement estimate is not possible. These losses can be mitigated by a carcass weir at the downstream end of the study area (Simpson, 1984). Carcasses trapped against the weir can be included in recovery counts. However, weirs may be inoperable during high run-off, precisely when the rate of carcass loss is greatest (Johnston et al., 1986; Shardlow et al., 1986). Additional carcasses are lost to scavengers, decay, and deep pools. Live fish are also subject to predation (Johnston et al., 1986). These removals may not be included in recovery counts.

Underwood et al. (2007) appropriately applied the Darroch Estimator by con-

Table 4.6: Evaluation of mark-recapture methods for estimating salmon escapement

Name	Estimator	Variance	Assumptions	Empirical	Recommendation
Pooled Petersen	Fair	Good	Unreasonable	Biased or Robust	Cautionary
Schaeffer	Fair	None	Unreasonable	Unnecessary	Do Not Use
Darroch	Fair	Good	Reasonable	Requires closed migration route	Best Choice
Jolly-Seber	Poor	Good	Unreasonable	Imprecise	Do Not Use
Schwarz & Arnason	Good	Good	Unreasonable	Not much tested	Do Not Use
Manly-Parr	Poor	None	Reasonable	Not much tested	Do Not Use
Robust Design	Good	Good	Unreasonable	Not much tested	Do Not Use
CIR	Good	Good	Reasonable	Requires intervention and partial removal	Best Choice

structing tagging and recapturing strata that defined a geographically closed population of migrating spawners. By tagging salmon downstream and recovering them upstream, Underwood et al. (2007) ensured that assumptions supporting the Darroch Estimator were reasonably met (Table 4.6). The Darroch Estimate of escapement is a maximum likelihood estimate and the variance is a derivative thereof. Underwood et al. (2007) appears to be the only study where in the assumptions of the Darroch Estimator are reasonably satisfied.

The Jolly-Seber Method (Section 4.5.1)of estimating salmon escapement is convoluted and includes biases due to assumption violations, even with the adjustments of Schwarz et al. (1993). Manly (1970) showed that when the mortality was dependent on arrival time the Jolly-Seber estimate of escapement was positively biased. Additionally, the Jolly-Seber estimate is a combination of multiple parameters which results in very large variance estimates (Manly et al., 2003). The Jolly-Seber Method is not recommended for estimating escapement using live fish (Table 4.6).

The Schwarz-Arnason Method (Section 4.5.3) yields larger variance estimates than the Jolly-Seber Method and requires even more assumptions, although the additional assumptions ensure that estimates are biologically meaningful. The benefit of having biologically meaningful birth estimates may outweigh the model's sensitivities to assumptions. However, as the trade-off is uncertain and method remains untested in practice, we do not recommend it (Table 4.6).

The Manly-Parr Method (Section 4.5.4) requires fewer assumptions about survival than the Jolly-Seber method, making the estimator seem more reasonable. However, the Manly-Parr method can not provide estimates for the number of salmon that arrive and die between surveys, so the Manly-Parr estimator of escapement is negatively biased. The Manly-Parr method does not include a variance estimate on total escapement. A bootstrapping procedure may provide variance estimates, but there is little empirical testing of the Manly-Parr method with or without variance estimates. Due to the bias and the lack of empirical testing, we do not recommend the Manly-Parr

Method (Table 4.6).

The Robust Design Method (Section 4.6) improves upon the Jolly-Seber method by allowing for an estimate of new recruits at the first survey. The Robust method relies on all the assumptions of both the open and closed population models. If primary surveys are conducted on a short interval so that the population can be reasonably assumed to be closed, then the abundance estimates will be valid. However, the Jolly Seber method requires that the probability of survival is the same for all fish between surveys. Realistically, fish that arrived early are more likely to die over a survey interval than late-arriving fish. Thus, Robust birth estimates will likely be biased, similar to the Jolly-Seber birth estimates, resulting in a biased total escapement estimate, and should therefore not be used (Table 4.6).

The Change-In-Ratio Method (Section 4.7) is not commonly used, as it requires a hatchery or some other intervention point below wild salmon spawning grounds. It does provide a reasonably unbiased estimator and an estimate of variance without unreasonable assumptions. An estimate of the total number of hatchery returns can also be estimated and the wild population estimated as the difference between the total escapement and the hatchery returns. If the circumstances allow, the CIR Method is a good choice to estimate salmon escapement, though CIR estimates typically have large variances (Skalski, personal communication).

Chapter 5

PASSAGE COUNT METHODS

5.1 Introduction to Passage Count Methods

In Passage Count Methods, fish are enumerated as they migrate past a reference point. The reference point may be a tower from which an observer notes the passing salmon (e.g. Reynolds et al., 2007), or an acoustic transducer which detects, tracks, and records fish electronically. Passage counts are commonly made at some obstruction in the river where migrating fish are funneled past a viewing station (e.g. Hatch et al., 1998) or hold beneath a weir until being counted and released upstream (e.g. Ripley, 1949; Bocking et al., 1988).

5.2 Passage Count Notation

The following notation will be used throughout the Passage Count Methods chapter. Estimates will be denoted with hats. For example, escapement, N, is estimated by $\hat{N}$. Population variance, S^2, (the true variation between units in the population) is estimated by s^2. $E(x)$ and $V(x)$ are functions denoting the expected value and variance of xy. Note that some symbols used in passage counts are also used in mark-recapture methods, but with completely meanings.

AF	Adjustment factor relating passage counts to actual number of unique fish passing
C_{hi}	True number fish passing during hour h of day i
c_{hi}	Number of passing fish counted during hour h of day i
D_{ihj}	Number of downstream migrants detected during the jth minute of the hth hour on the ith run day
F_j	Combined harvest from commercial, tribal and sport fisheries for the season above counting point j
FB_j	Number of unique radio-tagged fish that fell back one or more times at dam j
H	Hours of count data recorded per day
K	Number of days in a given run
k	Number of days the passage point is monitored, $k = 1, ..., K$
LP_j	Number of unique radio-tagged fish passing dam j via a fish ladder
NLP_j	Number of unique radio-tagged fish passing dam j via navigation locks
N	Number of adult spawners in a given run; escapement
R_j	Number of unique radio-tagged fish that reascended dam j after falling back
S^2	True sampling variance, estimated by s^2
TLP_j	Total number of times radio-tagged fish pass over dam j via the fish ladder
U_{ihj}	Number of upstream migrants detected during the jth minute of the hth hour of the ith run day
X	Proportion of downstream migrants that are actively down migrating

5.3 Passage Count Assumptions

Below are assumptions use by one or more of the Passage Count Methods. Not all assumptions are required by any one method, and no single assumption is universal to all methods. Applicable assumptions are referred to in the descriptions of individual methods.

(A-1). There are no errors in species identification.

(A-2). Visual, video, or acoustic counts are made without error.

(A-3). Counts are made during all hours of fish migration.

(A-4). Counts are made every day of the run.

(A-5). The hours monitored are a random sample of a 24-hour day.

(A-6). Commercial, native, and sport fisheries account for all mortalities between dams, i.e. there is no natural mortality of fish prior to their arrival on the spawning grounds.

(A-7). No fish die in dam turbines.

(A-8). The total harvest between dams is estimable.

(A-9). The variance of the harvest estimate is estimable.

(A-10). Days monitored are random with respect to fish arrival distributions.

(A-11). The days monitored are a random sample of all days in a run.

(A-12). All salmon pass through the viewing area, i.e. no fish travel past the dam through navigation locks.

(A-13). No salmon fall back over the dam.

(A-14). No salmon reascend the dam via the viewing area.

(A-15). Proportions of tagged fish taking all migration routes are identical to those of untagged fish.

(A-16). Counts at each dam are independent from other dam counts.

(A-17). The probability of a radio-tagged fish being detected, given that it is present, is 100%.

(A-18). All fish migrate within the detection range of the transducer.

(A-19). Sonar system has been properly calibrated.

(A-20). Upstream migrant detection probability is estimable, with an estimate of variance.

(A-21). An unbiased estimate of the proportion of fish detected moving downstream that are actively down-migrating is available, or included in the efficiency estimate.

(A-22). The minutes sampled within any hour are selected randomly.

5.4 Interdam Enumeration Method

Dams along the Columbia and Snake Rivers have provided an accurate and precise means to estimate minimum escapements within interdam regions (Beiningen, 1976). There are records of the number of adult chinook salmon at Lower Baker Dam from as early as 1926 (Pacific States Marine Fisheries Commission, 2008). Gangmark and Fulton (1952) estimated the number of sockeye salmon escaping to the entire Columbia basin in 1951 as the number counted at Rock Island Dam because there was no commercial fishery above it and no evidence of successful salmon spawning below. With more than fifteen dams on the Columbia and Snake Rivers, counts are made at each dam from the Bonneville Dam up to Wells Dam on the Columbia and to Lower Granite on the Snake River (USACE, 2007). Escapements are estimated as the difference in counts between successive dams, minus the commercial, tribal, and sport catches (Dauble and Watson, 1990, 1997).

Most Columbia and Snake River dams are monitored every day of the run between 4 am and 8 pm (USACE, 2007). Some dams are monitored more closely, like Priest Rapids and Wanapum dams, which are monitored 24 hours a day between April 15 and November 15 (GCPUD, 2008). Ice Harbor dam is monitored between 6 am and 4 pm during the month of March and then from 4 am to 8 pm through October (USACE, 2007)). Separate counts are kept of chinook, coho, sockeye, and pink salmon, along with other species either by a live attendant or by review of a time-lapse video recording (Hatch et al., 1994; USACE, 2007; GCPUD, 2008).

The U.S. Army Corps of Engineers (USACE, 2007) report escapement over each dam as the sum of the daily counts:

$$\hat{C}_j = \sum_{i=1}^{K} c_{ij}. \tag{5.1}$$

The sum of the daily counts is an unbiased estimator of the total number of salmon passing over a dam if 1) there are no errors in species identification (A-1), 2) there are no errors in counts (A-2), 3) daily counts are made during all hours of migration (A-3), and 4) counts are made every day of the run (A-4). These assumptions are reasonably met at dams like Priest Rapids and Wanapum, but not at Ice Harbor, where fish passage is only monitored during daylight hours. Most salmon species are known to migrate at night (Hatch et al., 1994; Daum and Osborne, 1995), so counts performed strictly during daylight are likely to undercount the actual number of migrating adults. One way of eliminating the need for assumption (A-3) is to monitor a random sample of the migration hours, and to expand the hourly average by a factor of 24:

$$\hat{C}_{ij} = \frac{24}{H_i} \sum_{h=1}^{H_i} c_{hij}. \tag{5.2}$$

If the aforementioned assumptions hold, the seasonal counts over dam j are:

$$\hat{C}_j = \sum_{i=1}^{K} \left(\frac{24}{H_i} \sum_{h=1}^{H_i} c_{hij} \right). \tag{5.3}$$

Unfortunately, fish migrations are not usually the same during daylight versus nighttime hours. Hatch et al. (1994), found that 13.56% of the chinook and 6.68% of the sockeye passage over Tumwater Dam occured at night (between 8 pm and 4 am). Daum and Osborne (1998) found that chum salmon on the left bank of the Chandalar River, Alaska preferentially migrated at night. Lilja et al. (2008) found that sockeye migrations in the Horsefly River, B.C. peaked at sunset. Thus, expanding the daytime estimate into night may severely bias the passage estimate. Instead, daily counts need to be stratified between the daytime and nighttime hours:

$$\hat{C}_j = \sum_{i=1}^{K} \left[\left(\frac{DH_i}{H_{Di}} \sum_{h=1}^{H_{Di}} c_{D,hij} \right) + \left(\frac{NH_i}{H_{Ni}} \sum_{h=1}^{H_{Ni}} c_{N,hij} \right) \right] \tag{5.4}$$

where DH_i is the number of daytime hours, H_{Di} is the number of daytime hours sampled, $c_{D,hij}$ is the passage count during the hth daytime survey hour of the ith day at the jth dam, NH_i is the number of nighttime hours, H_{Ni} is the number of nighttime hours sampled, and $c_{N,hij}$ is the passage count during the hth nighttime survey hour of the ith day at the jth dam. Here “daytime” refers to the hours of daylight, and “nighttime” refers to the hours of darkness, which will depend on latitude and run timing.

Equation (5.4) can only be unbiased for the total number of salmon migrating upstream if all salmon pass through the viewing area rather than ascending the dam through navigation locks (A-12), and if no salmon fall back over the dam (A-13). If fish pass upriver through navigation locks, then counts from the viewing area will undercount actual passage. If fish fall back down over the dam, then the number counted beyond the fish ladder will be too large. The positive bias caused by fallback can be exacerbated by reascension: a reascending fish may be counted as many times as it reascends the dam (Boggs et al., 2004). Equation (5.4) is only unbiased for the number of fish passing over dam j if there is no reascension (A-14).

Assuming that no natural mortality occurs during the season (A-6) or mortality due to passage through turbines (A-7), escapement is the number of fish passing one dam C_j, minus the number of fish passing the next dam, C_{j+1}, minus the combined commercial, tribal, and sport fishery harvests between the two dams, F_j. With an estimate of the number of fish harvested between dams j and $j+1$, escapement between dams j and $j+1$ can be estimated as:

$$\hat{N}_j = \hat{C}_j - \hat{C}_{j+1} - \hat{F}_{j,j+1}, \tag{5.5}$$

which is statistically unbiased for escapement as long as the assumptions hold.

If the methods used to estimate passage at each dam, C_j, and interdam catch, $F_{j,j+1}$, are independent, then the variance of the escapement estimator in Equation (5.5) is:

$$V(\hat{N}_j) = V(\hat{C}_j) + V(\hat{C}_{j+1}) + V(\hat{F}_{j,j+1}), \tag{5.6}$$

because the covariance terms are zero.

The variance Equation (5.5) requires unbiased estimates of F and $V(\hat{F})$. Methods estimating interdam sport and commercial harvest and it's variance are beyond the scope of this report, so we assume $\hat{F}$ and $\hat{V}(\hat{F})$ are known. The variance of passage counts at dam j is based on stratification of daily counts and is given by (Cochran, 1977, page 93):

$$V(\hat{C}_j) = \sum_{i=1}^{K} \left[H_{Di} \left(1 - \frac{DH_i}{H_{Di}}\right) \frac{S^2_{D,hij}}{DH_i} + H_{Ni} \left(1 - \frac{NH_i}{H_{Ni}}\right) \frac{S^2_{N,hij}}{NH_i} \right] \tag{5.7}$$

where $S^2_{D,ij} = \frac{1}{H_{Di}-1}\sum_{h=1}^{DH_i}(c_{D,hij} - \bar{C}_{D,ij})^2$ is the variance in hourly counts during daylight and $S^2_{N,ij} = \frac{1}{H_{Ni}-1}\sum_{h=1}^{NH_i}(c_{N,hij} - \bar{C}_{N,ij})^2$ is the variance in hourly counts during night for dam j.

The variance of the escapement estimator in Equation (5.5) is estimated by substituting the population variances, $S^2_{D,ij}$ and $S^2_{N,ij}$, with their estimates:

$$s^2_{D,ij} = \frac{\sum_{h=1}^{DH_i}(c_{D,hij} - \bar{c}_{D,ij})^2}{H_{Di} - 1}, \tag{5.8}$$

and

$$s^2_{N,ij} = \frac{\sum_{h=1}^{NH_i}(c_{N,hij} - \bar{c}_{N,ij})^2}{H_{Ni} - 1}. \tag{5.9}$$

Data for this example were taken from Boggs et al. (2004) and Manning and Smith (2005). Boggs et al. (2004) reported the USACE count of Spring/summer chinook salmon ascending the fish ladder at Bonneville and McNary Dams in 2001 as $\hat{C}_1 = 467,523$ and $\hat{C}_2 = 326,603$. Manning and Smith (2005) reported the sport catch of chinook salmon (including jacks) between Bonneville and McNary Dams from April to July as $\hat{F}_{1,2} = 749$. Assuming that the commercial and native fisheries took no additional fish, the escapement estimate for Spring/summer chinook between Bonneville and McNary Dams in 2001 is:

$$\hat{N}_{1,2} = 467,523 - 326,603 - 749 = 140,171 \text{ chinook}. \tag{5.10}$$

For this example, we assume that both dams are monitored for fish passage twenty-four hours per day during the entire run of Spring/summer chinook, so the variance of the fish ladder counts is zero. Manning and Smith (2005) did not include an estimate of variance for the sport catch data. Consequently, the estimate of variance for the chinook salmon escapement estimate is zero, although this underestimates the true variance due to the uncertainty in estimating catch data.

Fallback and Reascension Adjustment (Boggs et al., 2004)

Boggs et al. (2004) used radiotelemetry to quantify fallback rates and proportions, reascension rates and proportions, and the numbers of fish that pass upstream through navigational locks. They developed an adjustment factor for fishway viewing counts based on radiotelemetry data which incorporates the effects of fallback, reascension, and fish passing through navigation locks, thereby relaxing assumptions (A-12), (A-13), and (A-14).

The true number of radio-tagged salmon passing over the dam is calculated as the number of tagged fish passing through the ladder, LP, plus the number of tagged fish passing through the navigation locks, NLP, minus the number of tagged fish that fell back, FB, plus the number of radio-tagged fish that reascended, R. The adjustment factor, AF is the number of tagged fish that passed above the dam, divided by the total number of tagged individuals (not unique) counted passing through the fish ladder, TLP.

$$AF_j = \frac{LP_j + NLP_j - FB_j + R_j}{TLP_j}. \tag{5.11}$$

Assuming that passage proportions and fallback and reascension rates for untagged fish are the same as those for tagged fish (A-15), the adjustment factor can be applied to counts at each dam, resulting in an escapement estimator:

$$\hat{N}_j = \left(AF_j \times \hat{C}_j\right) - \left(AF_{j+1} \times \hat{C}_{j+1}\right) - \hat{F}_{j,j+1}. \tag{5.12}$$

Provided (A-15) holds, Equation (5.12) is unbiased for salmon escapement between dams j and $j+1$.

Boggs et al. (2004) assumed 100% detection in the radiotelemetry study (A-17), so the measurement error was assumed zero. This does not account for uncertainties due to range of radio-detection, tag failure, or undetected tags. Nor does a variance estimate of zero incorporate the variability in salmon migration routes. The variance term for the adjustment factor for dam j is included as $V(AF_j)$. Assuming, as before, that estimates made at each dam and for interdam harvest are independent (A-16), and that estimates of passage counts are independent of the radio-telemetry studies, the variance of Equation (5.12) is (Seber, 1982, page 9):

$$\begin{aligned} V(\hat{N}) &= V(AF_j)C_j^2 + V(\hat{C}_j)AF_j^2 + V(AF_j)V(\hat{C}_j) \\ &\quad + V(AF_{j+1})C_{j+1} + V(\hat{C}_{j+1})AF_{j+1}^2 + V(AF_{j+1})V(\hat{C}_{j+1}) \\ &\quad + V(\hat{F}_{j,j+1}). \end{aligned} \tag{5.13}$$

Equation (5.13) is estimated by:

$$\begin{aligned} V(\hat{N}) &= \hat{V}(AF_j)\hat{C}_j^2 + \hat{V}(\hat{C}_j)\widehat{AF}_j^2 - \hat{V}(AF_j)\hat{V}(\hat{C}_j) \\ &\quad + \hat{V}(AF_{j+1})\hat{C}_{j+1} + \hat{V}(\hat{C}_{j+1})\widehat{AF}_{j+1}^2 - \hat{V}(AF_{j+1})\hat{V}(\hat{C}_{j+1}) \\ &\quad + \hat{V}(\hat{F}_{j,j+1}). \end{aligned} \tag{5.14}$$

For the estimate of variance in Equation (5.14), $V(\hat{C}_j)$ and $V(\hat{C}_{j+1})$ are estimated by Equation (5.7) and $V(\hat{F}_{j,j+1})$ is estimated by a known $\hat{V}(\hat{F}_{j,j+1})$ (A-9). A variance estimate $\hat{V}(AF_i)$ has not yet been developed.

Fallback and Reascension Adjustment Example

This is a continuation of the Interdam Enumeration Example. Boggs et al. (2004) did not provide the number of tagged fish passing through the ladder, LP, the number passing through the navigation locks, NLP, or the number of radio-tagged fish that reascended, R, each dam. The reported adjustment factor for Bonneville Dam in 2001 was (Boggs et al., 2004, Table 9):

$$AF_1 = \frac{\text{No. of unique fish passing}}{\text{Total number of fish passages counted}} = 0.934, \tag{5.15}$$

and the adjustment factor for McNary Dam was (Boggs et al., 2004, Table 9):

$$AF_2 = 0.988. \tag{5.16}$$

Using the 2001 adjustment factors, the estimate of escapement for Spring/summer chinook between Bonneville and McNary Dams is:

$$\hat{N}_{1,2} = (0.934 \times 467,523) - (0.988 \times 326,603) - 749 = 113,233.7 \text{ chinook.} \tag{5.17}$$

5.5 Fence and Weir Counts

While passage counts generated at dams provide accurate escapement estimates for entire basins, many salmon runs need to be managed at the tributary level. Specific rivers can be isolated with weirs or fences in the river designed to obstruct and direct fish passage. Fish can be dipnetted over the weir, funneled into narrow channels where they can be counted (Ripley, 1949; Bue et al., 1998), or shunted into holding pens where they can be counted and released individually (Seiler et al., 2004; Volkhardt et al., 2005). Weirs are widely considered to be the most accurate form of escapement data (Cousens et al., 1982; Schubert, 1998) and are frequently used as the standard against which other types of escapement estimation are tested (Johnston et al., 1986; Bocking et al., 1988; Irvine et al., 1992; Miyakoshi and Kudo, 1999; Shardlow, 2001). In contrast, Crone and Bond (1974) found them to underestimate escapement.

Unlike dam passage counts, weir counts need no adjustment for fallbacks or reascensions. If the weir is operational and staffed 24 hours per day every day of the run (A-3) and (A-4), then the estimate of escapement is the total number of fish counted at the weir:

$$\hat{N} = \sum_{i=1}^{K} C_i, \tag{5.18}$$

Assuming that all salmon are correctly identified (A-1) and that counts are made without error (A-2) then Equation (5.18) is a complete census and the variance of Equation (5.18) is zero.

Often, weirs are not monitored 24 per day every day. Weir operations may be limited due to staff and funding limitations (for example Johnston et al. (1986) where the fence was operational only at night) or due to major flooding or freshet events that either wash the weir out or necessitate that the weir be temporarily removed to avoid being washed out (Skud, 1958; Shardlow et al., 1986; Labelle, 1994; Bue et al., 1998). Labelle (1994) proposed a method to account for the break in fence count data that assumes arrival and survival distributions and requires that some of the fish be tagged, though it does not appear to have been used. Another method is to assume that the distribution of fish arriving at the weir during non-operational hours is the same as during operational hours (A-10). Operational hours are treated as a random sample from 24 hours per day (A-5) and the operational days as a random sample from the total number of days in the run, K (A-11). In that case the estimator of escapement, based on finite sampling theory is:

$$\hat{N} = \frac{K}{k}\sum_{i=1}^{k}\left(\frac{24}{H}\sum_{h=1}^{H} c_{hi}\right) \tag{5.19}$$

Equation (5.19) is statistically unbiased for escapement, but is likely experimentally biased by the failure of Assumption (A-10). Often weirs are rendered inoperable due to freshets that spur large movements of fish upstream. Consequently, the number of salmon migrating past the weir site during monitored hours and days may be non-representative of non-monitored times. Therefore, weirs and fences are best suited to be used in tributaries where the risk of washout is minimal.

The variance of Equation (5.19) can be factored into two parts, corresponding to the variance contributed by taking a random sample of days in the run (Stage 1) and the variance contributed by taking a random sample of hours per day (Stage 2):

$$V(\hat{N}) = \underbrace{E_2[V_1(\hat{N}|2)]}_{\text{Stage 1}} + \underbrace{V_2(E_1[\hat{N}|2])}_{\text{Stage 2}} \tag{5.20}$$

The variance estimate for a two-stage sampling process is given by Cochran (1977, page 278) as:

$$\hat{V}(\hat{N}) = K^2\left[(1-\frac{k}{K})\frac{s_1^2}{k} + (1-\frac{H}{24})\frac{s_2^2}{Hk}\right] \tag{5.21}$$

where

$$s_1^2 = \frac{\sum_{i=1}^{k}(\bar{c}_i - \bar{\bar{c}})^2}{k-1}, \quad \bar{\bar{c}} = \frac{1}{k}\sum_{i=1}^{k}\bar{c}_i \tag{5.22}$$

is the sample variance among the daily means and

$$s_2^2 = \frac{\sum_{i=1}^{k}\sum_{h=1}^{H}(c_{ih} - \bar{c}_i)^2}{k(H-1)}, \quad \bar{c}_i = \frac{1}{H}\sum_{h=1}^{H}\bar{c}_{ih} \tag{5.23}$$

is the sample variance among hourly means within days.

Weir Count Method Example

The data for this example were taken from the Ballard Locks Sockeye Salmon Count during July 2008 and posted on the internet by WDFW (2008). We treated the daily counts from July 1 to July 30, 2008 as hourly counts from a wier, and assumed that every $H = 3$ counts were randomly selected from the 24 hours in each of $k = 10$ days. We further assumed that the 10 days were randomly selected from a run of $K = 30$ days. The data are shown in Table 5.1.

Using Equation (5.19), the weir count estimate of escapement is:

$$\hat{N} = \frac{30}{10} \sum_{i=1}^{10} \left(\frac{24}{3} \sum_{h=1}^{3} c_{hi} \right) = 376,224 \text{ sockeye.} \tag{5.24}$$

The variance estimate depends on the sampling variance of the days chosen out of the total number of run days, which is estimated as:

$$s_1^2 = \frac{\sum_{i=1}^{10} (\bar{c}_i - \bar{\bar{c}})^2}{10 - 1} = 139,237 \text{ sockeye.} \tag{5.25}$$

The variance estimate also depends on the sampling variance of the hours selected within each day, which is estimated as:

$$s_2^2 = \frac{\sum_{i=1}^{10} \sum_{h=1}^{3} (c_{ih} - \bar{c}_i)^2}{10(3-1)} = 25,948.03 \text{ sockeye.} \tag{5.26}$$

From the sampling variances of days and hours within days, the variance estimate for escapement is:

$$\hat{V}(\hat{N}) = (30)^2 \left[\left(1 - \frac{10}{30}\right) \frac{(139,237)}{10} + \left(1 - \frac{3}{24}\right) \frac{(25,948.03)}{3(10)} \right] = 9,035,355 \text{ sockeye.} \tag{5.27}$$

The standard error is $\hat{SE}(\hat{N}) = 3005.89$ and the coefficient of variation is $\hat{CV}(\hat{N}) = 0.00033$. Note that the sampling effort in hours per day need not be large in order to obtain precise estimates of escapement using passage counts.

Table 5.1: Data for the Weir Count Method example, constructed so that ten days were randomly selected from a run of thirty days, and three hours were randomly selected from each of the ten days. The counts are the total counts from each hour surveyed. Data taken from WDFW (2008).

Day (i)	Hour 1 (c_{i1})	Hour 2 (c_{i2})	Hour 3 (c_{i3})	Daily Mean ($\bar{c}_i$)
1	1103	1263	1167	1177.67
2	1367	916	498	927.00
3	745	600	619	654.67
4	886	842	564	764.00
5	518	654	787	653.00
6	439	394	311	381.33
7	251	278	275	268.00
8	230	320	245	265.00
9	149	97	53	99.67
10	30	53	22	35.00
Run Mean				522.53

5.6 Sonar Counts

A major drawback to the wier method is that weirs can be washed out during high flows, producing negatively biased estimates of escapement. Sonar is an in-river method of estimating salmon passage counts that is not affected by high flow events. Early applications of sonar to enumerate adult salmon in the Pacific Northwest used the Bendix counter, starting in 1978 in the Copper River, Alaska (Lewis, 2005), then later the single-beam echosounder (Burwen et al., 2005). Neither the Bendix counter nor the single-beam echosounder could provide information on direction of movement or the size of passing fish. Dual-beam echosounders were the first sonar applied to migrating salmon in-river to determine direction of travel (i.e. Eggers, 1994). Dual-beam echosounders were superseded by split-beam echosounders which provide the three-dimensional location of detected fish, give a more accurate estimate of fish size, and provide direction of migration (Johnston and Steig, 1995). Currently, two different types of sonar are being used to estimate salmon passage in rivers, both of which can determine the direction of movement of passing fish. One is the split-beam echosounder (i.e. Mulligan and Keiser, 1996) and the other is the dual-frequency identification sonar (DIDSON) (Holmes et al., 2005). Maxwell and Gove (2004, 2007) recommend DIDSON sonar above split-beam sonar because of it's higher resolution and less stringent placement requirements, but has a limited range without the addition of a long-range lens Sound Metrics Corp. (2009).

For the purpose of this report, we will assume that a transducer has been placed to ensonify the entire volume of water that migrating fish may pass through (i.e. Lilja et al., 2008), so that all fish are compelled to swim through the region of the river that the transducer can detect (A-18). This assumption is usually met through careful consideration of transducer placement, beam shape, and the construction of weirs that direct fish to swim through the acoustic beam. The details of transducer placement and specifications are numerous and beyond the scope of this report. They have been well documented in the literature, particularly in reviews by Ransom et al. (1998) and Maxwell and Smith (2007).

For sonar output to generate unbiased estimates of passing fish, each system must be properly calibrated (Simmonds, 1990). "Calibration" is the changing of settings in the sonar system so that echosounder output reports true values of a reference object placed in the transducer beam (Foote et al., 1987). The reference object is a tungsten carbide or copper sphere MacLennan and Simmonds (1992); Simmonds (1990). Calibration ensures that sonar systems are free from malfunction and work consistently within themselves so that any error (variability) in echosounder output is negligible. Calibration does not guarantee bias free estimates of escapement, but bias free estimates fish passage cannot be guaranteed without it (Foote et al., 1987). Descriptions of sonar calibration have been given by Foote et al. (1987) and Simmonds (1990).

As with dam counts, sonar counts must include the direction of migration. Let U_i be the number of upstream migrants and D_i be the number of downstream migrants of a particular species detected on run day i. A natural estimator of escapement would be the difference between the number of fish moving upstream and the number moving downstream:

$$\hat{N} = \sum_{i=1}^{K}(U_i - D_i). \tag{5.28}$$

The simple difference in upstream versus downstream counts is unbiased only if the detection probability of fish in each direction is 100% (A-2). The high-resolution images of DIDSON data can be reasonably assumed to detect each passing fish, even at densities exceding 2000 fish passing per hour (Maxwell and Gove, 2004). The lower-resolution output of split-beam echosounders cannot be assumed to have a 100% detection probability, and detection probability must be estimated. Upstream migrant detection probability can be estimated by comparing simultaneous sonar and alternate estimates. As with observer efficiency, the details of estimating detection probability are beyond the scope of this report, so we refer the reader to the examples given by Enzenhofer et al. (1998), and Hammarstrom and Hasbrouck (1998, 1999).

Not all objects detected moving downstream are out-migrating salmon (Stables and Kautsky, 2000). The estimate of the detection probability of downstream migrants must also account for debris (Nealson and Gregory, 2000). Non-fish debris should not be included in either the upstream or downstream estimates of salmon. Some salmon, particularly males, are washed out of the river in a moribund state, in increasing numbers as the season progresses (Chapman, 1943; Lilja et al., 2008). Since moribund and dead fish are included in the escapement, they should not be subtracted from the upstream estimate. Only those fish that are *actively* migrating should be subtracted. Observer efficiency estimates for downstream migrants should relate all downstream detections to only those fish actively migrating (A-21).

Assuming that the sonar system has been calibrated (A-19) and that there is an estimate of the detection probability for both upstream, P_U ((A-20), and downstream migrants, P_D (A-21), the estimate of escapement is:

$$\hat{N} = \sum_{i=1}^{K} \left(\frac{U_i}{P_U} - \frac{D_i}{P_D} \right). \tag{5.29}$$

One drawback to using sonar to count fish is that species identification is often difficult. Errorless counts of migrating salmon derived from sonar are only possible where there is no species overlap in run timing (A-1). Eggers (1994) determined that target strength was insufficient to distinguish between sockeye and chinook salmon in the Kenai River, Alaska, because differences between the two species were not large enough to overcome variability in target strength within species. Lilja et al. (2000) found it impossible to distinguish salmonids from whitefish (*Coregonus lavaretus*) and pike (*Esox lucius*) in Finland based on target strength. Burwen and Fleischman (1998) used target strength and pulse widths at different frequencies to exclude sockeye from chinook counts on the Kenai River, Alaska. Using this multivariate approach, they were able to exclude 87% of the sockeye while including 88% of the chinook (Burwen and Fleischman, 1998, Table 2). With an estimate of either the actual number of one species or the relative abundances of the species, the mulitvariate approach can provide estimates of passage (Fleischman and Burwen, 2003).

Equation (5.29) assumes that fish passage data are recorded continuously. However, due to staffing, data processing, or memory space limitations, some rivers are monitored for less than 60 minutes each hour. On the Kenai River, Alaska, split-beam data is collected for 40 minutes each hour (Patrick Nealson, personal communication). If the distribution of salmon passage is random within the hour then the m minutes sampled constitute a random sample of the 60 minutes possible (A-22). Lilja et al. (2008) showed that the distribution of sockeye passages on the Horsefly River, British Columbia, were random within the hour, indicating that a sampling scheme of m consecutive minutes each hour can provide unbiased estimates of total salmon passage. The average counts per minute must be expanded by 60 to estimate the number of salmon that swam past in hour h, and the hourly counts must be summed over every hour of every day of the run. The estimate of escapement for the season is:

$$\hat{N} = \sum_{i=1}^{K} \sum_{h=1}^{24} \frac{60}{m} \sum_{j=1}^{m} \left(\frac{U_{jhi}}{P_U} - \frac{D_{jhi}}{P_D} \right), \tag{5.30}$$

where U_{jhi} is the number of upstream migrants counted in the jth minute of the hth hour of the ith day, and D_{jhi} is the number of downstream migrants counted in the jth minute of the hth hour of the ith day.

The numbers of upstream and downstream migrants are usually counted during the same sampling interval. Consequently, the estimates are not independent. If the upstream and downstream processes are independent, then the estimates will remain uncorrelated. The variance of the sonar escapement estimate can be estimated as the sum of the upstream and downstream variances.

Uncertainty in the escapement estimate contributed by estimating the number of upstream migrants has two sources: 1) the probabilistic sampling of the minutes recorded, and 2) the estimate of detection probability. Once a system has been calibrated in place, it will produce outputs that are consistent with itself, so there is no uncertainty associated with the calibration process. By conditioning on the ratio of detected migrants to detection probability, and assuming that days and hours within days are estimated independently, the variance of the upstream migrants is:

$$\hat{V}(\text{Upstream}) = \sum_{i=1}^{K}\sum_{h=1}^{24} 60^2 \left\{ \left(\frac{\bar{U}_{ih}}{P_U}\right)^2 [\hat{CV}^2(\bar{U}_{ih}) + \hat{CV}^2(P_U)] + \left(1 - \frac{m}{60}\right)\frac{s_{UR}^2}{m} \right\} \tag{5.31}$$

where $\bar{U}_{ih} = \frac{1}{m}\sum_{j=1}^{m} U_{ihj}$, $\hat{CV}^2(\bar{U}_{ih}) = \frac{\sum_{j=1}^{m}(U_{ihj}-\bar{U}_{ih})^2}{m(m-1)\bar{U}_{ih}^2}$, $s_{UR}^2 = \frac{1}{m-1}\sum_{j=1}^{m}(R_{ihj} - \bar{R}_{ih})^2$, $R_{ihj} = U_{ihj}/P_U$, and $\bar{R}_{ih} = \frac{1}{m}\sum_{j=1}^{m} R_{ihj}$. The estimate $\hat{CV}^2(P_U)$ is assumed knowable from the method used to estimate upstream detection probability. The variance contributed by the estimation of the downstream migrants has the same form as Equation (5.31) but with D_{ihj} in place of U_{ihj} and P_D instead of P_U.

The total variance estimate is:

$$\begin{aligned}\hat{V}(\hat{N}) \quad = \quad & \textstyle\sum_{i=1}^{K}\sum_{h=1}^{24} 60^2 \left\{ \left(\frac{\bar{U}_{ih}}{P_U}\right)^2 [\hat{CV}^2(\bar{U}_{ih}) + \hat{CV}^2(P_U)] \right. \\ & \left. + \left(\frac{\bar{D}_{ih}}{P_D}\right)^2 [\hat{CV}^2(\bar{D}_{ih}) + \hat{CV}^2(P_D)] + \left(1 - \frac{m}{60}\right)\left(\frac{s_{UR}^2}{m} + \frac{s_{DR}^2}{m}\right) \right\},\end{aligned} \tag{5.32}$$

which is based on the variances for a random sample (of hours) and an estimate of the ratio of two independent estimators. It requires that the upstream and downstream estimates are independent.

Sonar Count Example

The data for this example were generated from lognormal distributions in the freeware program R. We generated 40 upstream and 40 downstream counts to simulate sampling 40 minutes in an hour. We assumed a detection probability of $\hat{P} = 0.98$ with a $CV(\hat{P}) = 0.001$ for both upstream and downstream migrants. The data, means, and sampling variance estimates are shown in Table 5.6.

The estimate of escapement for the simulated hour is:

$$\hat{N} = \frac{60}{40}(995.90 - 53.06) = 1,573 \quad \text{salmon.} \tag{5.33}$$

Table 5.2: Simulated data for 40 minutes of sonar sampling with assumed equal upstream and downstream detection probability of $\hat{P} = 0.98$.

Minute (j)	U_j	U_j/P	D_j	D_j/P	Minute (j)	U_j	U_j/P	D_j	D_j/P
1	24.00	24.49	0.00	0.00	21	14.00	14.29	0.00	0.00
2	17.00	17.35	3.00	3.06	22	42.00	42.86	2.00	2.04
3	19.00	19.39	1.00	1.02	23	13.00	13.27	0.00	0.00
4	12.00	12.24	2.00	2.04	24	10.00	10.20	1.00	1.02
5	25.00	25.51	0.00	0.00	25	16.00	16.33	3.00	3.06
6	17.00	17.35	2.00	2.04	26	13.00	13.27	0.00	0.00
7	15.00	15.31	0.00	0.00	27	11.00	11.22	1.00	1.02
8	30.00	30.61	1.00	1.02	28	17.00	17.35	1.00	1.02
9	65.00	66.33	3.00	3.06	29	10.00	10.20	5.00	5.10
10	7.00	7.14	3.00	3.06	30	19.00	19.39	4.00	4.08
11	29.00	29.59	3.00	3.06	31	43.00	43.88	2.00	2.04
12	7.00	7.14	0.00	0.00	32	33.00	33.67	1.00	1.02
13	14.00	14.29	2.00	2.04	33	56.00	57.14	1.00	1.02
14	13.00	13.27	1.00	1.02	34	26.00	26.53	1.00	1.02
15	20.00	20.41	2.00	2.04	35	13.00	13.27	1.00	1.02
16	39.00	39.80	0.00	0.00	36	31.00	31.63	1.00	1.02
17	4.00	4.08	2.00	2.04	37	15.00	15.31	0.00	0.00
18	20.00	20.41	0.00	0.00	38	47.00	47.96	0.00	0.00
19	24.00	24.49	0.00	0.00	39	43.00	43.88	1.00	1.02
20	95.00	96.94	2.00	2.04	40	8.00	8.16	0.00	0.00
					Total	976.00	995.90	52.00	53.06
					Mean	24.40	24.90	1.30	1.33
					s^2	1.60	342.10	1.60	1.67

The variance estimate is:

$$\begin{aligned}\hat{V}(\hat{N}) &= 60^2 \Big\{ (24.90)^2 \, [\tfrac{1.60}{40(24.40)^2} + (0.001)^2] \\ &\quad + (1.33)^2 \, [\tfrac{1.60}{40(1.30)^2} + (0.001)^2] + \big(1 - \tfrac{40}{60}\big)\big(\tfrac{342.10}{40} + \tfrac{1.67}{40}\big)\Big\} \qquad (5.34) \\ &= 41{,}254.28 \quad \text{salmon.}\end{aligned}$$

The standard error on the sonar estimate of the escapement for this simulated hour is $\hat{SE}(\hat{N}) = 203.11$ with a coefficient of variation $\hat{CV}(\hat{N}) = 0.129$. The variance for an entire season would be the total variance from each hour of every run day. Equation (5.34) shows that most of the uncertainty in the sonar estimate of escapement is due to temporal sampling.

5.7 Discussion of Passage Count Methods

The methods described in this chapter are summarized in Table 5.3. Passage count methods are the most accurate and precise methods of estimating salmon escapement. They are statistically simple, design-based methods that are capable of producing precise escapement estimates. Passage counts require few assumptions, but do require special circumstances. Interdam estimates require that there be at least two dams. Weir counts require that river flows and debris loads are low enough to keep weirs in operation. Sonar counts require particular stream bed morphology and that the run timing of the species of interest does not overlap with the migration of other species. All passage count methods require probabilistic sampling of passage to provide unbiased estimates and variance estimates.

Table 5.3: Evaluation of Passage Count Methods for estimating salmon escapement

Name	Estimator	Variance	Assumptions	Empirical	Recommendation
Interdam Enumeration	Good	Good	Reasonable	Requires dams	Best Choice
Weir Counts	Good	N/A	Reasonable	Can get washed out	Best Choice
Sonar Counts	Good	N/A	Reasonable	For single species only	Best Choice

Chapter 6

PEAK COUNT METHODS

6.1 Introduction to Peak Count Methods

For Peak Count Methods, periodic counts are made of live spawners (e.g. McPherson et al., 1999), of live plus dead fish (e.g. Parken et al., 2003) or of live plus cumulative dead (e.g. Schubert, 1998) until the number of live fish detected starts to decline. The count with the largest number of live fish is reported as the peak count. Peak counts can also refer to the maximum number of redds detected during surveys (see Chapter 7) rather than individual fish.

While peak counts represent only an index of abundance, efforts have been made to relate the peak number of spawners to the total escapement. A StreamNet query on salmon escapements investigated between 1997 and 2007 showed 3,189 different studies using Peak Count Methods in Washington, Oregon, California, Idaho and Alaska. Knudsen (2000) found 1,017 U.S. streams for which the peak count was the standard escapement estimate used for management. Despite the fact that peak counts provide only an index of abundance, they are prevelant enough to warrent explenation.

6.2 Peak Count Notation

The following notation will be used in this chapter. Estimates are denoted with hat's, for example, escapement, N, is estimated by $\hat{N}$. Averages are denoted with bar's, for example, average expansion factor is denoted $\bar{\pi}$. Expected values and variances are denoted by $E()$ and $V()$, respectively.

β_0	Regression intercept parameter
β_1	Regression slope parameter
C_{max}	Peak (maximum) count, including live fish, carcasses and/or cumulative carcasses
d_i	Number of dead fish (carcasses) detected on survey occasion i
D_i	Number of dead fish present on survey occasion i
D_{max}	Number of dead fish detected at the time of the peak live count
k	Number of surveys conducted during the run
l_i	Number of live fish detected on survey occasion i
L_i	Number of live present on survey occasion i
L_{max}	Peak (maximum) number of live fish
m_j	Number of miles surveyed of reach j
M	Number of miles of habitat accessible to salmon
t_m	The time of the maximum or peak count
$MSE(\theta)$	Mean squared error of estimator θ
N	Escapement, total fish abundance
n	Number of years the peak count is calibrated against a full escapement enumeration
P_{Li}	Probability of detection of live fish on survey i
P_{Di}	Probability of detection of dead fish on survey i
π	Peak count expansion factor
r	Number of reaches surveyed
R	Number of reaches available for spawning to the desired run
σ^2	Variance of the regression response variable, estimated by MSE
Σ	Sum over all years of salmon escapement, range assumed infinite

6.3 Peak Count Assumptions

Below are assumptions used in one or more of the Peak Count Methods. Not all assumptions are required by any one method, and only (A-1) is applicable to all methods. The assumptions are discussed in more detail as they arise in descriptions of individual methods.

(A-1). There is no error associated with species identification.

(A-2). Observer efficiency is estimable, applying to both live fish and carcasses.

(A-3). All fish (live or dead) are counted without error, i.e. observer efficiency is 100%.

(A-4). The recorded peak corresponds to the true peak in abundance.

(A-5). No salmon arrive in a stream after a peak in abundance has been recorded.

(A-6). No salmon die until the peak in abundance has been recorded.

(A-7). The probability of detection for each fish (or carcass) is the same during any survey.

(A-8). Every fish is equally likely to be in any location

(A-9). The detection of each fish (or carcass) is independent of all others.

(A-10). The variance of the detection probability estimate is estimable.

(A-11). All carcasses remain in the stream until the day of the recorded peak.

(A-12). Live salmon and carcasses are enumerated independently.

(A-13). Carcass counts are independent between surveys.

(A-14). All carcasses remain in the stream until counted.

(A-15). Surveys are conducted daily.

(A-16). Observer efficiency remains the same from year to year.

(A-17). The proportion of the run in index areas is the same from year to year.

(A-18). The proportion of the run constituting the peak is the same each year.

(A-19). Survey areas are selected randomly.

(A-20). Fish do not stray, either as live fish or as carcasses.

(A-21). Observer efficiency remains the same between surveys

(A-22). The proportion of the run that constitutes the peak comes from the same distribution every year

(A-23). There is an unbiased estimate of escapement to use as a basis for calibration with peak counts

(A-24). The variance of the alternative escapement estimator is estimable.

(A-25). Regression error is distributed Normally with mean=0, and variance = σ^2.

(A-26). The escapement is known without error.

6.4 Uncalibrated Peak Count Methods

Many peak count methods do not provide an estimate of escapement, but rather a qualitative means of comparing abundances from year to year. I provide descriptions of “Peak Indexing” methods in Sections 6.4.1-6.4.5 because they are prevalent, though without many unachievable assumptions, uncalibrated peak counts do not provide accurate estimates of escapement.

6.4.1 Peak Live Counts

Counts of live spawners are often reported as a peak count, i.e. the largest count in a series of abundance counts. Gangmark and Fulton (1952) referred to the unadjusted peak count as the "minimum standing crop" of salmon, acknowledging that it was being increased by new arrivals and decreased by mortalities daily, and therefore did not constitute a total estimate of escapement. Very few studies include an estimate of detection probability, further increasing the negative bias of the raw peak live counts as an estimate of escapement.

One way to reduce the negative bias of the Peak Live Count is to include an estimate of observer efficiency or detection probability, P_{Li} (A-2). With an estimate of detection probability for every survey, the true number of salmon present on any given survey will be estimable, thereby ensuring that the reported peak is the true peak (A-4). The more frequently surveys are conducted, the more likely the recorded peak is the true peak.

With an estimate of observer efficiency, the Peak Live Count estimate of escapement is:

$$\hat{N} = MAX\left(\frac{l_i}{P_{Li}}\right) \quad i = 1, \ldots, k, \tag{6.1}$$

and is only unbiased for escapement if all salmon arrive prior to and survive until the recorded peak abundance (i.e. (A-5) and (A-6)). Often the number of live fish in a given river fluctuates over the course of the run, especially during periods of very high or very low flow. If more than one peak is present (e.g. Crone and Bond (1976) Sashin Creek coho, Ames (1984) Perry Creek chum, Bue et al. (1990) Herring Creek pink salmon, Bocking et al. (1988) Black Creek coho), then either assumption (A-5) or (A-6) has been violated, possibly both. Even with only one distinct peak in abundance, (A-6) is often violated due to the length of the run. In Washington State for example, the peak live count occurs at about the time that carcasses begin to outnumber live fish (Ames et al., 1974). In such cases, the Peak Count method will estimate at best about half of the total escapement (Liao, 1994).

A complete variance estimate for Equation (6.1) is not possible. The uncertainty contributed to the estimate by estimating the observer efficiency and the binomial sighting process can be accounted for within the peak count. The peak count is a single observation from the largest extreme distribution (LEV) because it is the kth order statistic of the k total adjusted counts (Meeker and Escobar, 1998, page 86). A different sample of days would yield a different peak value, but with only one observation, the uncertainty in the estimate cannot be estimated.

A partial variance estimate for Equation (6.1) requires two new assumptions. The first assumption is that the detection probability is the same for all fish (A-7). While the positions of some fish may make them more or less visible, the fish are likely interchangeable. That is to say that every fish is equally likely to be in any given location (A-8) and that the probability of being in the harder or easier to see location is the same for all fish. The second assumption is that the detection of each fish is independent of all other fish (A-9). While the location of some fish may render others less likely to be seen (i.e. if one fish is directly below another), the probability of each fish to be found in those positions is, as before, the same for all fish. With these two assumptions, (A-8) and (A-9), the number of live fish recorded can be considered the realization of a binomial trial with P_{Li} probability of detection. Therefore, the partial variance of the Live Only Estimator is estimated as:

$$\hat{V}(\hat{L}_i) = \frac{l_i(1-\hat{P}_{Li})}{(\hat{P}_{Li})^2} + \hat{CV}^2(\hat{P}_{Li}) \left[\frac{l_i(1-\hat{P}_{Li})}{(\hat{P}_{Li})^2} + \left(\frac{l_i}{\hat{P}_{Li}} \right) \right], \tag{6.2}$$

where the variance of the detection probability estimate, $\hat{V}(\hat{P}_{Li})$, is assumed known (A-10). The proof is identical to that shown in Section 2.5 for the variance of the Binomial Sighting Process described by Liao (1994).

6.4.2 Peak Live and Dead Counts

One way to reduce the negative bias of the Peak Live Counts is to include the number of carcasses detected on the day of the peak live count, D_m. This will also be an undercount unless an adjustment is made for the detection probability of carcasses (A-2) such that the estimate of escapement is:

$$\hat{N} = MAX\left(\frac{l_i}{P_{Li}} + \frac{d_i}{P_{Di}}\right) \quad i = 1, \ldots, k. \tag{6.3}$$

The peak dead adjustment alleviates the assumption that live fish must remain in the spawning area until the day of the peak count (A-6), but requires the additional assumption that carcasses remain in the stream until the day of the peak live count (A-11). If carcasses or live fish leave (or are removed from) the stream before the peak count, the estimate of escapement will be too small. The later the peak occurs in a season, the more negatively biased the estimator will be because fish and carcasses will have more opportunities to be removed from the stream.

A complete variance estimate for Equation (6.3) is not possible because the peak count represents a single observation from the largest extreme value distribution. Therefore, uncertainty associated with the observed peak count cannot be incorporated into the variance estimate. A partial variance estimate for Equation (6.3) requires the assumptions that the probability of carcasses detection is the same for all carcasses (A-7) and that detection of all carcasses is independent (A-9). The number of carcasses recorded can be described as the realization of a binomial trial with P_{Di} probability of detection. Assumptions supporting the dead binomial framework are tenable for the same reasons as for live only counts. Assuming live fish and carcasses are enumerated independently (A-12), the variance estimate is:

$$\hat{V}(C_{max}) = \hat{V}(\hat{L}_i) + \hat{V}(\hat{D}_i), \tag{6.4}$$

where $\hat{V}(\hat{L}_i)$ is given in Equation (6.2). Assuming that the variance of the carcass detection probability estimate is estimable, $\hat{V}(\hat{D}_i)$ is given by:

$$\hat{V}(\hat{D}_i) = \frac{d_i(1 - P_{Di})}{P_{Di}^2} + \hat{CV}^2(P_{Di})\Big[\frac{d_i(1 - P_{Di})}{P_{Di}^2} + \left(\frac{d_i}{P_{Di}}\right)^2\Big] \tag{6.5}$$

Equation (6.4) underestimates the true variance because the uncertainty associated with the selection of the peak count as the kth order statistic cannot be incorporated.

6.4.3 Peak Live and Cumulative Dead Counts

Another adjustment to reduce the negative bias in peak counts is to include the cumulative number of carcasses up to the day of the peak live count, $\sum_{i=1}^{t_m} D_i$. The cumulative dead adjustment requires that carcasses remain in the stream until counted (A-14), which is more reasonable than (A-11), especially when the surveys are conducted frequently. Carcasses can still be lost to decay, flooding and scavenging, but the opportunity for these losses is reduced if surveys are frequent. Similar to the Live Only and Live and Dead Count methods, the Cumulative Count adjustment requires that no fish arrive in the stream after the day of peak spawning (A-5). Also, without an estimate of observer efficiency, Live and Cumulative Dead Counts will underestimate escapement (A-2). If the aforementioned assumptions hold, every fish of the run is accounted for, either as a carcass or a live fish and escapement is estimated as:

$$\hat{N} = MAX\left(\frac{l_i}{P_{Li}} + \sum_{j=1}^{t_m} \frac{d_i}{P_{Di}}\right) \quad i = 1, \ldots, k. \tag{6.6}$$

For Peak Live plus Cumulative Dead counts, the variance estimate includes the variances of all carcass counts up through the peak live count:

$$\hat{V}(\hat{N}) = \hat{V}(\hat{L}_{t_m}) + \sum_{j=1}^{t_m} \hat{V}(\hat{D}_j), \tag{6.7}$$

where $\hat{L}_{t_m}$ is the adjusted count of live fish on the day of the peak count, and the counts of carcasses for each survey are independent (A-13), i.e. carcasses are not repeatedly counted over time.

Equation (6.7) is biased for the variance of the escapement estimator in Equation (6.6) because the uncertainty associated with the peak count as the kth order statistic cannot be included.

6.4.4 Relative Peak Counts

For many years in Washington State, counts of live spawners in standard index areas were made until the peak count was identified, i.e. the counts began to decrease (Deschamps, 1969; Ames et al., 1974). Index area peak counts could not be used “to project a total escapement to a river system, since the indexes represent[ed] an unknown portion of the total available spawning area” (Ames et al., 1974). Since the index areas were not randomly selected, no valid inference could be drawn from them. Instead, the counts were compared to the previous ten-year average for each index area, and “the overall escapement to each system [could] then be described in terms of the average value for recent years” (Ames et al., 1974). The index comparison is only valid if three things remain constant from year to year: 1) observer efficiency (A-16), 2) proportion of the run in index areas (A-17), and 3) proportion of the run present during the peak (A-18).

If two years with the same true peak abundance have different observer efficiencies, then the year with the higher observer efficiency would have a higher escapement estimate. For example, observer efficiency in Cathead Creek was estimated as 0.825 in 1990, 0.246 in 1991, and 0.685 in 1992 (Bue et al., 1998, Table 1). An actual peak count of 200 fish would appear to be 165 in 1990, 49 in 1991, and 137 in 1992. Not including observer efficiency would result in the conclusion that escapement was low in 1991 compared to 1990 or 1992.

Years with higher or lower proportions of the run in index areas or at peak spawning will be inaccurately believed to have higher or lower escapements. Deschamps (1969) reported that while there were more coho and chinook than usual in several of the rivers on the Washington coast, the index areas actually had fewer fish because greater flows allowed the fish to migrate further up tributaries than normal, and out of the index reaches.

The proportion of the run in the peak may also change from year to year. Deschamps (1969) noted that coho in a Wynoochee River tributary spawned over the course of four months, which was a longer time than usual. For the same total escapement, a longer duration of spawning would necessitate a smaller peak.

Due to wide variation in observer efficiency and proportions of runs in index areas and in the peak each year, relative peak counts cannot be used for escapement estimates.

6.4.5 Fish/mile Index

For many years in Oregon State, counts of live spawners and carcasses have been made in standard index areas until the peak is achieved (Beidler and Nickelson, 1980). For the jth index area, the peak total is divided by the length of the reach for an estimate of the average number of fish per mile:

$$\hat{F}_j = \frac{L_{max,j} + D_{max,j}}{m_j},$$

assuming that all live fish and carcasses present are detected, $P_{Li} = 1$ and $P_{Di} = 1$.

Fish/mile is a measure of salmon density at the time of the peak rather than abundance. It includes no method to estimate the relationship between peak counts and escapement and serves only as an index of abundance (Nickelson, 1981). Like relative peak counts, fish-miles cannot provide absolute estimates of escapement because index areas are not selected randomly. If the surveyed areas are selected randomly out of all possible areas (A-19), then the average fish-per-mile estimate can be made by dividing the total fish counted by total miles surveyed. The estimate of the total peak count, and therefore escapement, is the average fish-per-mile estimate multiplied by the total number of miles in the stream or system:

$$\hat{N} = M \frac{\sum_{j=1}^{r} C_{max,j}}{\sum_{j=1}^{r} m_j} = M \frac{C_{\bullet}}{m_{\bullet}}, \tag{6.8}$$

where $C_{max,j} = L_{max,j} + D_{max,j}$. The Fish/mile Method allows for differences in peak timing between reaches, but requires the additional assumption that fish do not stray from one reach to another (A-20). For nearby reaches this may be violated by roving males seeking out new ripe females. Carcasses are also likely to "stray" by being washed out of one reach and into another reach downstream (e.g. Schubert, 2000). Straying can introduce bias to the estimate in either direction at an unknown degree.

The variance of Equation (6.8) is based on simple random sampling for a total derived from a ratio and is approximated by (Cochran, 1977, page 153):

$$V(\hat{N}) = \frac{M^2}{r}\left(1 - \frac{m_\bullet}{M}\right)\left[\frac{\sum_{j=1}^{R}(C_{max,j} - \frac{C_\bullet}{m_\bullet}m_j)^2}{R-1}\right]. \tag{6.9}$$

The middle product term is the finite population correction which becomes zero when every mile of the system has been surveyed, such that $m_\bullet = \sum_{j=1}^{r} m_j = M$. The rightmost product term is the population variance of the peak fish per mile counts since the numerator is equivalent to $\sum_{j=1}^{R}(C_{max,j}/m_j - \frac{C_\bullet}{m_\bullet})^2$. The variance given in Equation (6.9) does not incorporate uncertainty in peak counts due to observer efficiency or the binomial sighting process. Consequently, the true variance does not go to zero as $m_\bullet \to M$.

The variance is estimated by (Cochran, 1977, page 155):

$$\hat{V}(\hat{N}) = \frac{M^2}{r}\left(1 - \frac{m_\bullet}{M}\right)\left[\frac{\sum_{j=1}^{r}(C_{max,j} - \frac{m_j\hat{N}}{M})^2}{r-1}\right]. \tag{6.10}$$

The variance estimate is unbiased for Equation (6.9) but underestimates the true escapement because it does not incorporate the uncertainty due to observer efficiency or the binomial sighting process.

6.5 Expanded Peak Count Methods

Due to the near-impossibility of meeting assumption criteria, uncalibrated peak counts can only be used as an index of abundance. To get a true estimate of escapement, the relationship between peak counts and escapement needs to be quantified. The relationship between peak count and escapement can be estimated four different ways. The simplest way is to estimate the escapement-to-peak ratio for a number of years and take the average over years to estimate the true relationship. This is called the "Mean of the Ratios Estimator" and is described in Section 6.5.1. A less biased estimator is the "Ratio of the Means Estimator," which is calculated by comparing the mean escapement to the mean peak count and is described in Section 6.5.2. The third estimator is called the "Least Squares Estimator" and is based on the regression through the origin of escapement against peak count, as described in Section 6.5.3. Alternately the peak count can be regressed against an errorless estimate of escapement and the established relationship is inverted to predict the escapement that caused a new peak count. This fourth method is called "Inverse Prediction" and is described in Section 6.5.4.

6.5.1 Mean of the Ratio Method (McPherson et al., 1999)

The Mean-of-the-Ratio estimate of escapement is the product of the peak count, C_{max}, and an expansion factor, β:

$$\hat{N} = C_{max} \times \beta, \tag{6.11}$$

where $C_{max} = L_m$ if only live spawners are counted (e.g. McPherson et al., 1999), $C_{max} = L_m + D_m$ if live spawners and carcasses are both tallied on each survey (e.g. Parken et al. 2003), or $C_{max} = L_m + \sum_{j=1}^{t_m} D_j$ if live spawners plus cumulative dead are recorded (e.g. Schubert, 1998). In order for Equation (6.11) to be unbiased for escapement, the counts must capture the true peak in abundance (A-4). This can be done if surveys are conducted daily (A-15) and if observer efficiency remains constant between surveys (A-21), so that the peak is not missed when only visibility improved. Often, counts are made until the number of live fish begins to decline (Ames et al., 1974). The latter portion of the run goes unmonitored, and there is potential for an initial peak to be mistaken as the true peak and the true peak missed.

Differences in observer efficiency may also hinder observer's ability to accurately identify the true peak. Often visibility is reduced or surveys are hindered during freshets, but it is during freshets that salmon preferentially migrate upstream (Deschamps, 1969). When there are more fish in the river, the probability of detection would decrease. Fewer fish might be reported and the peak count may be assigned during a day when visibility was higher but abundance was lower.

The issue of observer efficiency between years (A-16) confounds the problem by interfering with the assumption that the ratio of peak count to escapement is drawn from the same distribution every year. Differences in observer efficiency indicate that ratios are from different distributions, but expansion factors are estimated as if they were all from the same "population" of ratios.

The proportion of a run that makes up the peak from year to year must also come from the same distribution (A-22), such that annual expansion factors represent samples from the same population. Therefore all counts must be of the same life states, for example, all Live Only counts, as opposed to mixed Live Only and Live and Dead counts.

The estimate of the expansion factor, $\hat{\beta}$, requires an alternate estimate of escapement, $\hat{N}_{alt}$ (A-23), as well as the peak count, C_{max}:

$$\hat{\beta} = \frac{\hat{N}_{alt}}{C_{max}}. \tag{6.12}$$

McPherson et al. (1999) calculated expansion factors for chinook salmon in the Taku River in Alaska using a mark-recapture estimate of escapement. Note that McPherson et al. (1999) denoted the expansion factor as π but I have denoted it β to show how the method relates to other peak methods. The expansion factor of McPherson et al. (1999) was the mean of the expansion factors for five years ($n = 5$):

$$\bar{\beta} = \frac{1}{n}\sum_{i=1}^{n} \hat{\beta}_i = \frac{1}{n}\sum_{i=1}^{n} \frac{\hat{N}_{alt(i)}}{C_{max(i)}}. \tag{6.13}$$

The estimate of the expansion factor will only be an unbiased estimate if the alternate estimate of escapement is unbiased (A-23) and if the peak counts are made without error (A-3).

Parken et al. (2003) used the Bootstrap procedure to estimate the variance of the escapement estimate in Equation (6.11). McPherson et al. (1999) assumed that the variance of the alternative escapement estimator could be estimated (A-24), which was true in their study because they used mark-recapture methods. They also assumed that the detection probability of live fish was unity so there was no error in the peak count (A-3), and estimated the variance of each expansion factor estimate as:

$$\hat{V}(\hat{\beta}) = \frac{1}{(C_m)^2} V(\hat{N}_{alt}), \tag{6.14}$$

though in reality detection probabilities are almost never unity and visual counts are almost always underestimates (Bevan, 1961). The estimate of variance of the mean expansion factor, $\bar{\beta}$, includes both the sampling variance and the average process variance and is given by (McPherson et al., 1999, Table A3):

$$\hat{V}(\bar{\beta}) = \frac{\sum_{i=1}^{k}(\hat{\beta}_i - \bar{\beta})^2}{k-1} + \frac{\sum_{i=1}^{k} V(\hat{\beta}_i)}{k} \tag{6.15}$$

where $\hat{\beta}_i$ is the expansion factor estimated for year i, and $V(\hat{\beta}_i)$ is given by Equation (6.14). The variance of the Mean of the Ratios Escapement estimate is estimated as:

$$\hat{V}(\hat{N}) = (C_m)^2 \hat{V}(\bar{\beta}), \tag{6.16}$$

where $\hat{V}(\bar{\beta})$ is given by Equation (6.15). Equation (6.16) almost certainly underestimates the true variance because it does not incorporate the uncertainty contributed by the estimation of the peak counts.

6.5.2 Ratio of the Means Method

The second way to estimate the relationship between peak counts and total escapement is to estimate the ratio of the means (rather than the mean of the ratios) by:

$$\hat{\beta} = \frac{\frac{1}{n}\sum_{i=1}^{n} \hat{N}_{alt,i}}{\frac{1}{n}\sum_{i=1}^{n} C_{max,i}} = \frac{\bar{\hat{N}}_{alt}}{\bar{C}_{max}}. \tag{6.17}$$

The estimate of the expansion factor given in Equation (6.17) is positively biased on the order of $1/n$, which becomes negligible for samples larger than 30 (Cochran, 1977, pages 31, 160-162).

The variance of the expansion factor estimate is derived from finite sampling theory and is estimated by (Cochran, 1977, page 23):

$$\hat{V}(\hat{\beta}) = \frac{1}{n\bar{\hat{N}}^2} \frac{\sum_{i=1}^{n}(C_i - \hat{\beta}\hat{N}_{alt,i})^2}{(n-1)}, \tag{6.18}$$

where the finite population correction factor is ignored. Equation (6.18) underestimates the true variance on the order of $1/n^2$ (Cochran, 1977, page 162-163), which is negligible for more than a few years worth of data. Equation (6.18) further underestimates the true variance because it does not incorporate the uncertainty contributed by the estimation of peak counts.

*6.5.3 Calibration Regression (*or *Least Squares Method)*

Linear regression relates a known “predictor” or “independent” variable x, with an estimated or measured “response variable,” y, in the general form $y = \beta_0 + \beta_1 x + \epsilon$. When estimating salmon escapement, the peak count can be considered the predictor variable and the estimated escapement is the response variable. The relationship is estimated using linear regression with the peak count is “calibrated” to the total escapement, hence the name “calibration regression.”

When escapement is zero, the peak count must also be zero, and if counts are made without error (A-3), or have an unbiased estimate of observer efficiency (A-2), then a peak count of zero implies that escapement is zero. Thus a regression of escapement against peak counts may be constrained to go through the origin so that only the slope, β_1, must be estimated. For a given peak count, the average estimated escapement is modeled as:

$$N_i = \beta_1 C_{max,i} + \epsilon,$$

where ϵ is the random variation, distributed normally with mean zero and variance σ^2 (A-25). Other error distributions are also possible, but a normally distributed error is standard (Kutner et al., 2004).

The slope coefficient is estimated to minimize the vertical sum of squares of differences at each point from the estimated line through the origin:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} C_{max,i} N_i}{\sum_{i=1}^{n} C_{max,i}}. \tag{6.19}$$

In a regression through the origin, β_1 is an alternative to the expansion factor β of the "Ratio of the Means Estimator," and the prediction is valid if and only if peak counts are distributed normally (A-25) (Fuller, 2006, page 77).

The estimate of escapement for a new peak count C_{new}, is:

$$\hat{N} = \hat{\beta}_1 C_{new} \tag{6.20}$$

In the case of estimating salmon escapement, we are interested in the single new value of the escapement based on the peak, rather than the mean possible values. The variance now has two sources. One source is the uncertainty in estimating the location of that distribution since it is based on an estimate of β_1 rather than on a known value. The second source is the variance in the distribution of escapements at a given peak count, which is assumed to be σ^2. The variance of the predicted value of the escapement is derived by conditioning on the two stages of sampling, where the first stage refers to sampling with respect to peak counts, C_{max}, and the second stage refers to sampling with respect to escapement, N:

$$\begin{aligned}
V(N_{new}) &= V_2\left(E_1(N_{new}|2)\right) + E_2\left(V_1(N_{new}|2)\right) \\
&= V_2\left(E_1(\hat{\beta}_1 C_{new}|2)\right) + E_2\left(V_1(\hat{\beta}_1 C_{new}|2)\right) \\
&= V_2\left(C_{new} E_1(\hat{\beta}_1|2)\right) + E_2\left(C_{new}^2 V_1(\hat{\beta}_1|2)\right) \\
&= V_2(C_{new}\beta_1) + E_2\left(\frac{(C_{new})^2\sigma^2}{\sum_{i=1}^{T}(C_{max,i})^2}\right) \qquad (6.21) \\
&= V_2(N_{new}) + E_2\left(\frac{\sigma^2(C_{new})^2}{\sum_{i=1}^{n}(C_{max,i})^2}\right) \\
&= \sigma^2 + \frac{\sigma^2(C_{new})^2}{\sum_{i=1}^{n}(C_{max,i})^2} \\
&= \sigma^2\left(1 + \frac{(C_{new})^2}{\sum_{i=1}^{n}(C_{max,i})^2}\right)
\end{aligned}$$

From Equation (6.21) it is clear that the two sources of variance are independent, and the total variance is the sum of both the process error (σ^2) and the sampling error (from estimating β). Beidler and Nickelson (1980) regressed peak count against fish-days and found that though the correlation was high (R^2 up to .90), the variance was so large that estimates were unusable. For a live and dead peak count of 150 chinook, the 95% confidence interval was from 1000-6000 spawner days (Beidler and Nickelson, 1980).

6.5.4 Inverse Prediction Method

Calibration regression requires that peak counts be measured with absolute certainty. This is almost never true in practice (Bevan, 1961), though the degree of uncertainty may be unknown. If escapement estimates can be made without error (A-26), then peak counts can be regressed against the escapements (c.f. Section 6.5.3). Both a slope and an intercept must be estimated because a peak count of zero does not necessarily mean that escapement was zero, as there is a possibility that present fish are undetected. Once β_0 and β_1 have been estimated, predictions can be made about the unobserved escapement value that produced a newly observed peak count. The prediction is made by inverting the model equation and solving for the "predictor." There is also a technique called "inverse regression" which regresses the predictor variable against the response variable, but it is inappropriate for estimating salmon escapement from known escapements and uncertain peak counts because the predictor variable must be measured without error for a valid regression (Kutner et al., 2004).

For given values of escapement, N, the mean peak count is modeled as:

$$C_{max} = \beta_0 + \beta_1 N + \epsilon,$$

where the error, ϵ, is distributed normally with mean zero and variance σ^2 (A-25). The intercept and slope coefficients are estimated using simple linear regression, or ordinary least squares, and the peak count estimate is:

$$\hat{C}_{max} = \hat{\beta}_0 + \hat{\beta}_1 N. \tag{6.22}$$

To estimate the escapement from a new peak count, a new peak count, C_{new}, is substituted into Equation (6.22) and solve for the predicted escapement, N_{new}:

$$\hat{N}_{new} = \frac{\hat{C}_{new} - \hat{\beta}_0}{\hat{\beta}_1}, \quad \text{where} \quad \hat{\beta}_1 \neq 0. \tag{6.23}$$

Kutner et al. (2004) gives the variance of the inverse predicted escapement estimate as:

$$\begin{aligned} V(\hat{N}_{new}) &= V\left(\tfrac{\hat{C}_{new} - \hat{\beta}_0}{\hat{\beta}_1}\right) \\ &= \tfrac{V\left(\hat{C}_{new}\right)}{\hat{\beta}_1{}^2} \\ &= \tfrac{\sigma^2}{\hat{\beta}_1{}^2}\left(1 + \tfrac{1}{n} + \tfrac{(N_{new} - \bar{N})^2}{\sum_{i=1}^{n}(N_i - \bar{N})^2}\right), \end{aligned} \tag{6.24}$$

which is estimated by:

$$\hat{V}(\hat{N}_{new}) = \frac{MSE}{\hat{\beta}_1{}^2}\left(1 + \frac{1}{n} + \frac{(N_{new} - \bar{N})^2}{\sum_{i=1}^{n}(N_i - \bar{N})^2}\right). \tag{6.25}$$

Inverse Prediction Method Example

The data for this example are taken from (McPherson et al., 1999, Table A2) for chinook salmon spawning in the Taku River, British Columbia. The data are presented in Table 6.1.

Table 6.1: Data for the Inverse Prediction Method example, taken from (McPherson et al., 1999, Table A2) for chinook salmon spawning in the Taku River, British Columbia. Escapement was estimated using mark-recapture techniques.

Year	Peak Count	Escapement
1989	8,986	40,329
1990	12,077	52,142
1994	9,929	—
1995	7,971	33,805
1996	18,576	79,019
1997	13,201	114,938
Average	12,162.2	64046.6

A linear regression fit to the data provides estimates of β_0 (7077 with s.e.=3966) and β_1 (0.0794 with s.e.=0.056). Assuming a new peak count of 9929 (Table 6.1), the estimate of escapement is:

$$\hat{N}_{new} = \frac{9,929 - 7,077}{0.0794} = 35,922 \text{ chinook.} \tag{6.26}$$

It is worth noting that the new peak count should fall within the range of the peak counts used to estimate the slope and intercept. If the new peak count is too small, in this case less than 7,077, the escapement estimate will be negative and hence biologically irrelevant.

The variance estimate requires an estimate of the mean squared error of the regression. The fitted peak values are subtracted from the observed peak values, resulting in “residuals”:

$$-1,292.9265 \quad 860.0735 \quad -1,789.8945 \quad 5,224.9308 \quad -3,002.1833.$$

The residuals are squared, and the mean of the squared residuals is the estimated mean squared error: 8,385,623. The variance of the escapement estimate is:

$$\hat{V}(\hat{N}_{new}) = \frac{8,385,623}{(0.0794)^2}\left(1 + \frac{1}{5} + \frac{(35,922 - 64046.6)^2}{\sum_{i=1}^{5}(N_i - 64046.6)^2}\right) = 1,833,312,429. \quad (6.27)$$

The estimate of standard deviation is $\hat{SE}(\hat{N}) = 42,817$ and the coefficient of variation is $\hat{CV}(\hat{N}) = 1.1919$.

Measurement Error Bias Correction

Escapement estimates used to establish the peak-escapement relationship are not known without error. In such cases $\hat{\beta}_1$ is negatively biased for the slope in Equation (6.22) (Fuller, 2006, page 3) and the estimate of escapement given in Equation (6.23) will be positively biased.

Fuller (2006) described regression methods for cases when both the dependent and independent variables are measured with error. The classic "errors-in-variables" model, the ratio of the variance of the response variable to the predictor variable must be known (Fuller, 2006, page 30). In the case of salmon escapement, the ratio of peak count error to escapement estimate error, $\frac{V(C_{max})}{V(\hat{N})}$, is almost always unknown. Another bias correction exists for when the variance of only the predictor variable, i.e. the escapement estimate, is known. The variance of the escapement estimate can be incorporated into the estimate of β_1 as (Fuller, 2006, page 14):

$$\hat{\beta_1}^* = \frac{\sum_{i=1}^{n}(\hat{N}_i - \bar{\hat{N}})(\hat{C}_{max,i} - \bar{\hat{C}}_{max})}{(n-1)[s_N^2 - \overline{V(\hat{N})}]}, \quad (6.28)$$

where $s_N^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{N}_i - \bar{\hat{N}})^2$ is the sample variance among the escapement estimates and $\overline{V(\hat{N})} = \frac{1}{n}\sum_{i=1}^{n}\hat{V}(\hat{N}_i)$ is the mean variance estimate of the escapement estimates. Equation (6.28) can be used in place of $\hat{\beta}_1$ in Equation (6.23).

The estimate of β_1^* given in Equation (6.28) is only unbiased if escapement estimates are normally distributed and estimates of peak counts and escapements are independent. If surveys are simultaneous for both peak counts and mark-recapture estimates, detection biases (e.g. visibility) will affect both estimates and the errors will be correlated, causing a bias of unknown direction and magnitude.

When the “known” escapements must be estimated, Equation (6.25) will underestimate the true variance of the predicted escapement, because the uncertainty contributed by the estimate variance is not incorporated in the predicted estimate’s variance.

6.6 Discussion of Peak Count Methods

Peak counts are widely used, but can provide only an index of salmon escapement.

The Uncalibrated Peak Methods (Section 6.4) do not provide true estimates of escapement and should not be used (Table 6.2). Uncalibrated peak methods do not include a scaling from peak counts to total escapement. To equate peak counts with escapement requires assumptions that are rarely met in practice. Consequently, raw count methods, the relative peak count method, and the fish-per-mile method are not recommended for estimating salmon escapement.

The Mean-of-the-Ratios estimator provides a biased estimate of escapement, and requires no fewer assumptions and no less data than the Ratio of the Means estimator. The Ratio-of-the-Means estimator is less biased than the Mean of the Ratios estimator but still requires several years of data to estimate the relationship between peak counts and escapements, and has a large variance. The Mean-of-the-Ratio Method remains a statistically valid technique that can be used to estimate salmon escapement, and so is recommended with caution (Table 6.2).

The Calibration Regression Method is the most direct peak count method to estimate escapement and, because it estimates only the slope of the relationship, has the smallest variance. Unfortunately, empirical efforts to calibrate peak counts to escapement have shown that the variability in the relationship from year to year prohibits the Calibration Regression method from being useful for management (Beidler

Table 6.2: Evaluation of Peak Count methods for estimating salmon escapement

Name	Estimator	Variance	Assumptions	Empirical Experience	Recommendation
Raw count, live only	Poor	Poor	Unreasonable	Index only	Do Not Use
Raw count, live and dead	Poor	Poor	Unreasonable	Index only	Do Not Use
Raw count, live and cum. dead	Fair	Fair	Unreasonable	Index only	Do Not Use
Relative Peak	Poor	None	Unreasonable	Index only	Do Not Use
Fish/mile	Poor	Poor	Unreasonable	Index only	Do Not Use
Means of Ratios	Poor	Poor	Unreasonable	Not much testing	Do Not Use
Ratios of Means	Fair	Poor	Unreasonable	Imprecise, data-hungry	Cautionary
Calibration Regression	Good	Good	Unreasonable	Imprecise, data-hungry	Do Not Use
Inverse Prediction	Good	Good	Reasonable	Imprecise, data-hungry	Cautionary

and Nickelson, 1980). Calibration regression also requires that the peak count is known without error, which is unreasonable given different observer efficiencies. Consequently, calibration regression should not be used to estimate salmon escapement (Table 6.2).

The Inverse Prediction Method can allow for error in the predictor variable and is therefore more applicable for estimating salmon escapement than calibration regression. The variance is larger than that of Calibration Regression because inverse prediction estimates more parameters and incorporates extra uncertainty associated with predicting (rather than estimating) an escapement. When the error of the predictor variable must be estimated, the variance of the inverse prediction will underestimate the true uncertainty. Of the peak count methods, Inverse Prediction is the most statistically valid method and is recommended with caution (Table 6.2).

Chapter 7

REDD COUNT METHODS

7.1 Introduction to Redd Count Methods

Mature female salmonoids dig nests, called "redds," in which they bury eggs. The redds are often distinguishable from the surrounding substrate as lighter circular or ovoid patches on the stream bed (Chapman, 1943; Dauble and Watson, 1990).Surveys that count salmon redds are conducted aerially from fixed-wing aircraft (Meekin, 1967; Ames and Phinney, 1977; Dauble and Watson, 1997), from the ground (e.g. Meekin, 1967), by boat (e.g. Gallagher and Gallagher, 2005), with snorkels (e.g. Al-Chokhachy et al., 2005), or using SCUBA equipment (e.g. Swan, 1989). Redds are counted individually when sparse, or in 10's or 50's where more abundant (Welsh, 1983; Dauble and Watson, 1997). Redd counts are used as an index to compare relative abundances (Dauble and Watson, 1997) and to examine population trends (Welsh, 1983), but can also be used to identify characteristics of spawning habitat (Geist et al., 2002). While it is generally argued that salmonoid populations cannot be reliably determined from redd count data (Welsh, 1983; Dauble and Watson, 1990; Holecek and Walters, 2007), redd counts are the primary method used to estimate chinook escapement in the Pacific Northwest (Figure 1.1).

7.2 Redd Count Notation

The following notation will be used in this chapter. Estimates are denoted with hat's, for example, escapement, N, is estimated by $\hat{N}$. Variances are denoted by $V()$.

A_i	The area of the redd patch at the time of the ith survey, i=1,2,...,k
C_i	Redd counts (redds detected) on the ith survey, i=1,2,...,k
C_j	Redds detected along the jth transect, j=1,2,...,n
$C_\bullet$	Total redds constructed in a given run
D	Density of redds
F	Number of female spawners in a given run
G	The ratio of true redds per female
k	Number of surveys conducted during a given run
L_j	Length of river included in transect study
M	Number of male spawners in a given run, including jacks (precocious males)
n	Number of transects
N	Number of adult spawners in a given run, i.e. escapement
P	Probability of detecting a redd that is present
R	Male-to-female ratio
T	Proportion of detected redds containing eggs; percentage of true redds
W	Width of transects

7.3 Redd Count Assumptions

In my descriptions of Redd Count Methods we have explicitly assumed that the species responsible for constructing observed redds were identified without error. Salmonid species frequently coincide and there may be some uncertainty in redd identification (e.g. Gallagher and Gallagher, 2005). Chinook salmon build the largest redds, making their redds easy to identify. Other species, like coho in the Strait of Juan de Fuca, can be distinguished based on run timing and the general lack of other species (Haymes, 2007). Therefore, redd count methods are best used when species are not likely to be confused with other salmonids.

Below are listed the assumptions required by one or more of the Redd Count

Methods. Not all assumptions are required by any one method, and only (A-1) is universal to all methods. The applicable assumptions are referred to and discussed in the descriptions of individual techniques.

(A-1). There is no error associated with redd species identification, i.e. all redds are of known species.

(A-2). All redds remain distinguishable until the last redd has been built.

(A-3). Redds do not overlap, i.e. no superimposition.

(A-4). Every redd present in study area (or transect) is detected, i.e. detection probability is 100%.

(A-5). Redds are counted during the entire season.

(A-6). Redds are counted once, at most.

(A-7). Proportion of superimposition remains constant between years.

(A-8). Probability of redd detection remains constant between years.

(A-9). Duration of redd visibility is known.

(A-10). The proportion of true redds, T, is known or can be accurately estimated.

(A-11). Each female deposits eggs in one and only one redd.

(A-12). Estimates of true redd proportion and number of redds are made independently.

(A-13). The male-to-female ratio, R, is known or can be accurately estimated.

(A-14). The male-to-female ratio remains constant over the course of a run.

(A-15). The variance of the true redd proportion estimator, $\hat{T}$, is estimable.

(A-16). The variance of the male-to-female ratio estimator, $\hat{R}$, is estimable.

(A-17). Estimates of true redd proportion, T, and male-to-female ratio, R, are made independently.

(A-18). Average redd size remains constant over the experiment.

(A-19). There are no gaps between redds.

(A-20). No false redds are recorded.

(A-21). Redd area is measured without error.

(A-22). The variance of the average redd size estimator is estimable.

(A-23). Estimates of the number of redds and the male-to-female ratio are made independently.

(A-24). Redd detectability by divers is constant over a season/run, i.e. the width of the strip transect is fixed.

(A-25). Transects are placed randomly at each study site.

(A-26). Area of a study site is measured without error.

(A-27). There is one and only one male for every female.

(A-28). The relationship between a female salmon and the area of the redd(s) she builds is known without error.

(A-29). Redds measured are a random sample of the redds present.

(A-30). The probability of redd detection remains constant during a run.

(A-31). The probability of redd detection and male-to-female ratio are estimated independently.

7.4 Ideal Redd Count Method of Estimating Salmon Escapement

The ideal escapement estimator generated from redd counts is:

$$\hat{N} = \frac{\hat{T}}{\hat{G}} \left(\sum_{i=1}^{k} \hat{C}_i \right) (1 + \hat{R}), \tag{7.1}$$

where $\hat{C}_i$ is the number of redds estimated to be present on the ith survey, T is the proportion of redds with eggs in them (true redds), G is the number of redds built per female, and R is the ratio of males to females. The number of redds detected is adjusted to estimate the true number of redds built, then scaled to the number of females. The number of males is estimated from number of estimated females. T, G, and R must be estimated, though in practice these parameters are often assumed known.

Female salmonids have been known to dig "trial" or "false" redds where a female digs a redd but does not deposit any eggs (Briggs, 1953; Welsh, 1983; Holecek and Walters, 2007). The number of detected redds must be adjusted to estimate the number of true redds that have been constructed (A-10). The proportion, T, of the detected redds assumed to be true redds can be estimated by sampling a random selection of redds for the presence of eggs (Holecek and Walters, 2007).

The proportion of false redds varies considerably from species to species and from system to system. Briggs (1953) found that 54% of coho redds and 68% of chinook redds in Prarie Creek, CA were false redds. In contrast, the Washington Department of Fisheries used an estimate of 5% trial redds for chinook in the Snohomish River system, WA (Ames and Phinney, 1977), so that $\hat{T} = 1 - 0.05 = 0.95$. The proportion of trial redds may also vary between year due to environmental factors. Meekin (1967) found trial redds exposed to air below Chief Joseph Dam. It was speculated that the unusually low river flow of that year forced chinook salmon to abandon incomplete nests for deeper water. The proportion of redds that contain eggs (i.e. are true redds, the complement of false redds) must be estimated specifically for every stream every year. The direction and magnitude of any bias associated with using a constant true redd proportion are unknown.

Most redd count methods assume that each female salmon builds one and only one true redd (A-11), such that $\hat{G} = 1$. If the redd-to-female ratio is larger than one, the estimate of escapement will be positively biased. There is evidence that coho salmon build between one and four redds. Gallagher and Gallagher (2005) found that escapement estimates based on the assumption of one redd per female were much larger than estimates based on mark-recapture and area-under-the-curve studies of the same population, although errors were too large for the difference to be considered significant. On the other hand, if the redd-to-female ratio is smaller than one, the estimate of escapement will be negatively biased.

The perceived redd-to-female ratio may be skewed by redd superimposition. Chapman (1943) observed 30 chinook salmon working on what he called a single redd. Likely, they were not all females as males attend ripe females on the redds very closely (Briggs, 1953). If the male-to-female ratio at the time was 2.1:1 as estimated by Briggs (1953) there may have been 10 females to at one redd, for a redd-to-female ratio estimate of $\hat{G} = 0.1$. Without a reliable way to estimate the number of true redds built per female, the direction and magnitude of the bias in the escapement estimate associated with redd enumeration is unknown. However, cases where females build more than one redd are considered rare (Haymes, 2007) and applying redd counts only at low densities help ensure that superimposition does not lower the perceived number of redds per female.

The second most influential parameter in the redd count methods is the male-to-female ratio, R. The male-to-female ratio is used to estimate the number of males from the estimated number of females and is therefore responsible for half the escapement estimate.

The sex ratio has been known to change over time (Briggs, 1953). In a study comparing 105 years of chinook, pink, coho, and sockeye data from seven streams, Morbey (2000) found that male salmon arrived at the spawning grounds significantly earlier than females 90% of the time. Consequently the male-to-female ratio tends to be very large at the beginning of the run. Killick (1955) found that males tended to outlive females, so the male-female ratio was inflated at the end of the run. An estimate of the sex ratio taken early or late in the run may lead to an overestimate of escapement. Conversely, a sex ratio estimate taken in the middle of the run will lead to an underestimate of escapement. Because of a changing male-to-female ratio in sockeye in the Adams River, B.C., Killick (1955) discouraged the use of redd count methods. Salmon must be sampled randomly throughout the run in order to obtain an unbiased estimate of the average male-to-female ratio (A-13).

The parameter R is often estimated from the male-to-female ratio of a random sample of live fish (Ames and Phinney, 1977; Gallagher and Gallagher, 2005) or of carcass recoveries (Hahn, WDFW, personal communication). There is evidence suggesting that the sex ratio of carcasses is not representative of the population (Shardlow et al., 1986) due to behavioral differences between male and female salmon. Female salmon defend their redds until death and are more likely to be detected as carcasses than male salmon which sink into inaccessible pools or drift downstream out of the study area.

The variance of the ideal escapement estimate given in Equation (7.1) is provided by Goodman (1960):

$$\begin{aligned} V(\hat{N}) &= E^2(\hat{N}) \Big[CV^2(\hat{T}) + CV^2(\hat{G}) + CV^2(\hat{R}) + CV^2(\hat{C}_\bullet) \\ &+ CV^2(\hat{T})CV^2(\hat{G}) + CV^2(\hat{T})CV^2(\hat{R}) + CV^2(\hat{T})CV^2(\hat{C}_\bullet) \\ &+ CV^2(\hat{G})CV^2(\hat{R}) + CV^2(\hat{G})CV^2(\hat{C}_\bullet) + CV^2(\hat{R})CV^2(\hat{C}_\bullet) \Big], \end{aligned} \tag{7.2}$$

and is estimated by:

$$\begin{aligned}
\hat{V}(\hat{N}) &= \hat{N}^2 \Big[\hat{CV}^2(\hat{T}) + \hat{CV}^2(\hat{G}) + \hat{CV}^2(\hat{R}) + \hat{CV}^2(\hat{C}_\bullet) \\
&- \hat{CV}^2(\hat{T})\hat{CV}^2(\hat{G}) - \hat{CV}^2(\hat{T})\hat{CV}^2(\hat{R}) - \hat{CV}^2(\hat{T})\hat{CV}^2(\hat{C}_\bullet) \\
&- \hat{CV}^2(\hat{G})\hat{CV}^2(\hat{R}) - \hat{CV}^2(\hat{G})\hat{CV}^2(\hat{C}_\bullet) - \hat{CV}^2(\hat{R})\hat{CV}^2(\hat{C}_\bullet)\Big].
\end{aligned} \tag{7.3}$$

Coefficients of variation for the true redd proportion $CV(\hat{T})$, the redd-to-female ratio $CV(\hat{G})$, and the male-to-female ratio $CV(\hat{R})$ are derived from the method of estimating their respective parameters. The coefficient of variation for the total redd estimate, $CV(\hat{C}_\bullet)$, depends on the method used to enumerate redds. The redd enumeration methods are described in Section 7.5.

Ideal Redd Count Method Example

The data for this example were taken from four different sources. The proportion of true redds was taken from Ames and Phinney (1977), the number of females per redd was taken from Haymes (2007), the total redd count was taken from Hahn and Thompson (2007), and the male-to-female ratio was taken from Briggs (1953). The parameter estimates, their assumed variances, and the source for each estimate are shown in Table 7.1.

Table 7.1: Parameter estimates, their assumed variances and coefficients of variation, and source for each estimate used in the Ideal Redd Count Method example.

Parameter	Estimate	Variance	CV	Source
T	0.95	0.04	0.04	Ames and Phinney (1977)
G	1	0	0	Haymes (2007)
$C_\bullet$	134	0	0	Hahn and Thompson (2007)
R	2.1	0.5	0.11	Briggs (1953)

Using Equation (7.1), the Ideal Redd Count estimate of escapement is:

$$\hat{N} = \frac{0.95}{1}\,(134)\,(1+2.1) = 395 \quad \text{salmon.} \tag{7.4}$$

It is worth noting that using a male:female ratio of 2.0 leads to an escapement estimate of 382 salmon and a male:female ratio of 1.5 leads to an escapement estimate of 318 salmon.

Using Equation(7.3), the variance estimate for the escapement is:

$$\begin{aligned} \hat{V}(\hat{N}) &= (395)^2\,[(0.04)^2 + (0)^2 + (0.11)^2 + (0)^2 \\ &\quad -(0.04)^2(0)^2 - (0.04)^2(0.11)^2 - (0.04)^2(0)^2 \\ &\quad -(0)^2(0.11)^2 - (0)^2(0)^2 - (0.11)^2(0)^2]\,. \\ &= 2303.89 \end{aligned} \tag{7.5}$$

The standard error is $\hat{SE}(\hat{N}) = 48.00$ and the coefficient of variation is $\hat{CV}(\hat{N}) = 0.1216$.

7.5 Redd Enumeration Methods

There are a myriad of ways to estimate the total number of redds, from aerial counts or photographs (Ames and Phinney, 1977; Visser et al., 2002), to marking redd locations with flags (Gallagher and Gallagher, 2005), and underwater videography (Dauble et al., 1999). All redd counts are used in some form of Equation (7.1) to relate the number of redds found during a season to the escapement.

7.5.1 Peak Redd Count Method

In some systems redds remain visible long after spawning is complete (A-2), and if every redd present is detected, (A-4), the peak count is considered a complete census of all redds constructed during the season (e.g. Dauble and Watson, 1997). The peak count is a total redd enumeration, and escapement is estimated from the peak count as described in Section (7.4), although it ignores all but a single C_i value from Equation (7.1).

Assumption (A-2) necessitates (A-3) that no redds be superimposed, or built on top of one another. When densities are low, redds tend not to overlap due to the territorial nature of female salmon (Haymes, 2007). At high spawning densities redds may overlap, making it impossible to distinguish individual redds (Hahn, WDFW, personal communication).

While redds may remain visible until the end of the run, they disappear over time due to algal growth (Dauble and Watson, 1990). Redds can also be scrubbed clean by high flows rendering them indistinguishable from the substrate (Haymes, WDFW personal, communication). When the assumption of a redd census cannot be relied on as a complete enumeration, the peak redd count can be related to total escapement in exactly the same manner as the Expanded Peak Count Methods described in Section 6.5. This relaxes assumptions of 100% detection (A-4) and no superimposition (A-3), but requires that the probability of redd detection (A-8) and proportion of superimposition (A-7) remain constant from year to year.

In practice, peak redd counts are either reported raw, as in many StreamNet queries, or they are compared to an independent estimate of escapement. Dauble and Watson (1990) acquired peak counts of Columbia River chinook salmon redds from aerial surveys and estimated escapement using the Interdam Enumeration Method (see Section 5.4). Dauble and Watson (1990) regressed escapement estimates against peak redd counts, assuming that redd counts were made without error (A-4) and (A-6). The escapement estimates were more likely to be error-less and inverse prediction would have been more appropriate (see Sections 6.5.3 and 6.5.4).

7.5.2 Redd Area-Under-the-Curve Method

In 1977 Russ Orrell sent a memo to Jim Ames (published as Appendix II in Ames and Phinney (1977)) describing the method used to estimate chinook escapement in the Skagit, Stillaguamish, and Snohomish River systems that year. The Washington Department of Fisheries (WDF) abandoned the use of the Factor 5 Method (Section 2.7.4), which was found to be unreliable due to variation in run timing and duration, and adopted the use of this new method to estimate escapement in the Skagit River (Ames and Phinney, 1977). The Orrell-Ames and Phinney Method equates the total number of redds detected during the run to female escapement, and does not depend on the magnitude or timing of the peak number of redds.

Aerial surveys are conducted on intervals equal to the amount of time redds remain visible in the study system (A-9). For instance, WDF estimated that redds in the Skagit system remained visible for 21 days, so aerial surveys were conducted every 21 days from the arrival of the first spawners (Ames and Phinney, 1977). For each survey, the number of detected redds was recorded, and assuming there was no redd overlap (A-3) and 100% probability of detection (A-4), the total number of redds detected for the season was the sum of the records from each flight, $\sum_{i=1}^{k} C_i = \hat{C}_\bullet$.

The average duration of redd visibility varies between systems and is probably not constant from year to year. Using a fixed value of redd visibility may introduce bias in either number of redds, because some redds will fade before they are counted. Conversely, if the average duration is longer than estimated, survey counts will overestimate total redds because individual redds may be counted more than once. Efforts are being made to estimate the duration of redd visibility for selected indicator streams (Hahn, WDFW, personal communication), which will likely reduce this bias. But even a stream-specific estimate is likely to be biased unless the estimate can be calibrated to environmental conditions governing the length of redd visibility.

There is no variance estimate possible for the sum of redd counts using the redd AUC method because survey days are chosen systematically. An estimate of the variance using Equation (7.3) and assuming $CV(\hat{C}_\bullet) = 0$ will underestimate the true variance of the escapement estimate. If survey days are chosen randomly, a variance estimator would be possible. See Section 2.7.2 for more details.

7.5.3 Redd Census

The assumption that all redds are visible for a known length of time (A-9) can be relaxed, if it can be assumed that each redd is only counted once (A-6). The position of each redd can be recorded on a map using a global positioning system (GPS) (Dauble et al., 1999; Geist et al., 2002) or flagged in river (Gallagher and Gallagher, 2005; Hahn and Thompson, 2007) to prevent it being counted multiple times. Visser et al. (2002), for example took aerial photographs of redds and used landmarks to record each redd in a geographical information system (GIS). When each redd can be individually recognized, surveys can be conducted more than once every redd life to ensure that no redds are overlooked (A-4), and the estimate of the number of redds built in a season is $\hat{C}_{\bullet}$.

Marking the position of each redd can help prevent undercounting due to redd overlap or complete superimposition (A-3) (Hahn and Thompson, 2007). Redd superimposition varies within salmon species and more so in years of high escapement (Hobbs, 1937; Neave, 1953; McNeil, 1964*a*; Dauble and Watson, 1997; Al-Chokhachy et al., 2005). If redds are laid directly on top of one another, recording positions using a GIS may still undercount the true number of redds. Hobbs (1937) excavated what was believed to be 16 different qunnat (chinook) salmon (*O. tschawytscha*) redds in New Zealand and found only 8 unique patches—the other 8 had been dug directly on top of the initial redds. If redds are marked with flags (strips of plastic wrapped around oblong stones) that are disturbed when another redd is constructed, then direct superimposition can be included in the estimate, assuming flags are not washed away. At low densities, the territorial behavior of female salmon helps prevent redd overlap, so redd count methods are best used when escapement is not large (Hahn, WDFW, personal communication).

The variance associated with a complete census is zero. Therefore, the estimate of variance of the escapement estimate in Equation (7.1) becomes:

$$\begin{aligned}\hat{V}(\hat{N}) &= \hat{N}^2 \Big[\hat{CV}^2(\hat{T}) + \hat{CV}^2(\hat{G}) + \hat{CV}^2(\hat{R}) \\ &- \hat{CV}^2(\hat{T})\hat{CV}^2(\hat{G}) - \hat{CV}^2(\hat{T})\hat{CV}^2(\hat{R}) - \hat{CV}^2(\hat{G})\hat{CV}^2(\hat{R})\Big].\end{aligned} \tag{7.6}$$

7.5.4 Strip Transect Method (Swan, 1989)

Several studies have indicated that some salmon, particularly chinook, spawn in water that is too deep for the redds to be detected from ground or air surveys (Swan, 1989; Dauble and Watson, 1990). Under these circumstances, the only way to estimate redd abundance is to use underwater surveys (Welsh, 1983). Swan (1989) conducted a study using SCUBA divers and a sled to count and map redd locations along strip transects in the Hanford Reach of the Columbia River. Redds were found as deep as 9.14 meters.

Location of redds within the width, W, of each transect were recorded on a map during each survey. At the end of the season, maps were overlaid to prevent double counting of redds (A-6) and the total number of redds for each transect, $C_j = \sum_{i=1}^{k} C_{ij}$, was recorded. The total number of redds recorded is unbiased for the total number of redds dug within the transect if the study commences prior to the start of spawning and is continued until all spawning has ceased (A-5), and if divers detect all redds within the transect 100% (A-4).

Swan (1989) estimated the average number of redds per transect, $\bar{C}$, to be the total number of redds in all transects, $\sum_{j=1}^{n} C_j$, divided by the number of transects, n:

$$\bar{C} = \frac{1}{n}\sum_{j=1}^{n} C_j. \tag{7.7}$$

Swan (1989) expanded the average number of redds per transect, $\bar{C}$, by the number of transects that would have fit in the study area, which was the length of the river in the study site, L, divided by the width of each transect, W:

$$\hat{C}_{\bullet} = \bar{C} \times \frac{L}{W}. \tag{7.8}$$

The estimate given in Equation (7.8) does not allow for transects to have differing lengths and thus cannot be applied to transects along wider or narrower reaches of a river. A better estimator of redd density is essentially the Ratio of the Means Estimator described in Section 6.5.2. In the case of strip transects, the estimated density of redds in the study area, $\hat{D}$, is the number of redds detected in all transects, divided by the total area of the transects:

$$\hat{D} = \frac{\sum_{j=1}^{n} C_j}{W \sum_{j=1}^{n} L_j}, \tag{7.9}$$

which is unbiased as an estimate of density if the transects are placed randomly in the study area (A-25). The estimate of total redds for the season is the estimated redd density, $\hat{D}$, times the total area of the study site, A_{site}:

$$\hat{C}_{\bullet} = \hat{D} A_{site}, \tag{7.10}$$

where A_{site} is assumed known without error (A-26). In a well-prepared study, it should be possible to calculate the relevant area of river bed *a priori*.

Swan (1989) gave no correction for trial redds. He assumed one redd per female (A-11) and a sex ratio of 50:50 (A-27). Based on Swan's (1989) assumptions, the ideal escapement estimator in Equation (7.1) reduces to:

$$\hat{N} = \left(\hat{C}_{\bullet}\right)(1+1) = 2\hat{C}_{\bullet}, \tag{7.11}$$

which may be biased in either direction, depending on the true values of the unestimated parameters. However, the number of redds estimated using the strip transect method is unbiased for the total number of redds in the study area and can be used in Equation (7.1) for an unbiased estimate of escapement, provided that transects are placed randomly as per Assumption (A-25) (Cochran, 1977, page 23).

Swan (1989) did not provide a variance estimate for the number of redds detected or for escapement. The estimate of the variance of the total number of redds, $\hat{V}(\hat{C}_{\bullet})$, is given by (Cochran, 1977, page 155) as:

$$\hat{V}(\hat{C}_\bullet) \cong \frac{(A_{site})^2}{n(n-1)} \left(1 - \frac{W\sum_{j=1}^{n} L_j}{A_{site}}\right) \sum_{j=1}^{n} (C_j - \hat{D}WL_j)^2, \tag{7.12}$$

which can be used, along with $\hat{C}_\bullet$ as a coefficient of variation, in Equation (7.3) for an unbiased estimate of the variance.

7.6 Redd Area Methods

7.6.1 Welsh (1983) Method

The Welsh (1983) Method is used to estimate the number of redds when the densities are so great that individual redds are difficult to distinguish. The boundaries of the spawning region are defined as the outer edges of the redd patch. The area within these boundaries is divided by the (assumed estimable) average area of a single redd to estimate the number of redds within the boundary. This method was never intended for use in estimating escapement, but only as a way "to determine large annual changes in population size" (Welsh, 1983).

Assuming that redds do not overlap (A-3) and that the average redd size remains constant for the whole study (A-18), the number of redds present at survey time i is equal to the redd area, A_i, divided by the average area of each redd, $\bar{A}$:

$$C_i = \frac{A_i}{\bar{A}}, \tag{7.13}$$

where

$$\bar{A} = \frac{1}{C_\bullet} \sum_{j=1}^{C_\bullet} A_j. \tag{7.14}$$

The Redd Area Method will underestimate the total area cleared for redds and the total number of redds if redds overlap. Since the Welsh method is used when redds are so close together that they cannot be individually identified, it cannot be assumed that they do not overlap. The extent of redd superimposition is difficult to quantify and most studies quantify redd superimposition by the quantity or proportion of eggs disturbed, rather than on area of overlap (e.g. Hobbs, 1937). Without a reliable way to quantify the area of redd superimposition, redd area methods of estimating redd number will be negatively biased to an unknown degree.

Assuming that there are no gaps between redds (A-19), the total area, A_i, can be measured as the area within the outer bounds of a spawning area, resulting in an estimate of total area, $\hat{A}_i$. It is reasonable to assume that there are no gaps between redds because if gaps existed, then individual redds would be distinguishable. If individual redds were distinguishable, then the Welsh (1983) Method would not be necessary. Using an estimate of the average area per redd, $\hat{\bar{A}}$, the number of redds present at time i can be estimated as:

$$\hat{C}_i = \frac{\hat{A}_i}{\hat{\bar{A}}}. \tag{7.15}$$

Provided that the area measured for each redd patch was only recorded once (A-6), the sum of the $\hat{C}_i$ is $\hat{C}_\bullet$ and can be used as in Equation (7.1) to estimate escapement. Welsh (1983) did not account for false redds (A-20), so his estimate of true redd proportion would be $T = 1$, with no variance. Similarly, Welsh (1983) did not provide an estimate of the ratio of redds to females, and thereby assumed the ratio to be 1:1 with no variance.

The variance of redd counts, $\hat{C}_i$, is not zero, because counts were estimated from redd area. Assuming that total redd area was measured without error (A-21), the variance of the redd counts estimates depended only on the estimate of average redd area, $\hat{\bar{A}}$, and its variance, $V(\hat{\bar{A}})$:

$$V(\hat{C}_i) = (\hat{A}_i)^2 \frac{V(\hat{\bar{A}})}{(\bar{A})^4}. \tag{7.16}$$

The variance of the sum of the redds, $\hat{C}_\bullet$, is estimated by:

$$\hat{V}(\hat{C}_\bullet) = \left(\sum_{i=1}^{k} \hat{A}_i \right)^2 \frac{\hat{V}(\hat{\bar{A}})}{(\hat{\bar{A}})^4}, \tag{7.17}$$

The variance estimate in Equation (7.17) can be used in Equation (7.3) to estimate the variance of the escapement estimate.

7.6.2 The Redd Area Method (Gallagher and Gallagher, 2005)

Rather than use redd area to estimate the number of redds and then equate the number redds to female abundance, Gallagher and Gallagher (2005) used redd area to estimate the number of females. They recorded the location and size of every coho redd detected in seven streams and rivers in Northern California. Based on an unpublished study that showed female coho make between one and four redds, they divided the redds into three size classes and assigned a proportion of a female to each size class. This method relaxes assumption (A-11) that each female builds one and only one redd, but requires additional assumptions about the average area a female will clear out for trial and true redds (A-28). It also requires an estimate of the male-to-female ratio (A-13).

Gallagher and Gallagher (2005) assumed that coho redds smaller in area than 2.0 square meters "represented" one quarter of a female coho. Redds between 2.1 and 5.0 square meters corresponded to half a female, and that all redds greater than 5.0 square meters corresponded to one female coho (Gallagher and Gallagher, 2005). In mathematical notation, the function of the area of the ith redd, $f(A_i)$, returns either 1/4, 1/2, or 1 depending on the area of the redd, as shown:

$$f(A_i) = \begin{cases} 1 & 5.1 \text{ m}^2 < A_i \\ \\ 1/2 & 2.1 < A_i < 5.0 \text{ m}^2 \\ \\ 1/4 & A_i < 2.0 \text{ m}^2 \end{cases}$$

where $i = 1, 2, ..., \hat{C}_{\bullet}$. The Redd-Area Function requires that each redd have distinct boundaries (A-3), so that the area can be measured without error (A-21). It also requires that no redd be counted more than once (A-6), which Gallagher and Gallagher (2005) ensured by placing flags on each identified redd.

Gallagher and Gallagher (2005) found that the detection probability of redds smaller than 1.5 m^2 was not significantly different for redds larger than 1.5 m^2. Since relationships in the area-to-female function were linear, adjusting the total output by observer efficiency was equivalent to adjusting each individual input. The total number of females was estimated as the sum of the function across all redd areas, divided by the estimate of detection probability, $\hat{P}$, which must be the average of the entire study in order to be unbiased (A-13):

$$\hat{F} = \frac{\sum_{i=1}^{\hat{B}} f(A_i)}{\hat{P}}. \tag{7.18}$$

The direction and magnitude of bias in the estimated number of females cannot be determined because of the arbitrary assignment of redd area to female salmon. If redds share boundaries or are overlapping such that multiple redds are classified as a single redd larger than 5.0 m^2, then the number of females will be underestimated. If many females build three redds that are between 2.0 and 5.0 m^2, then the number of females will be overestimated.

Under the assumptions of Gallagher and Gallagher (2005), the proportion of true redds T, the redd-to-female ratio G, and the number of redds built $\hat{C}_\bullet$ are irrelevant. Instead, escapement, N, is estimated by the estimated number of females, $\hat{F}$, times the quantity 'one plus the estimated sex ratio,' $(1 + \hat{R})$:

$$\hat{N} = \hat{F}(1 + \hat{R}), \tag{7.19}$$

where $\hat{F}$ is given by Equation (7.18) Due to potential violations in the assumption of redd superimposition and the arbitrary classification of redd areas to female abundance, the magnitude and direction of the bias of the Redd Area Method of estimating escapement is unknown.

Gallagher and Gallagher (2005) incorporated uncertainty associated with species identification in their estimate of variance, though they did not include the formula.

Issues of species identification are beyond the scope of this report, so we assume that redds are associated with the correct species and without error (A-1). The variance of the Redd Area Estimator includes uncertainty in the estimate of detection probabilities and variability in both sex ratio and redd area. While the numbers and areas of redds made by female salmon vary, the arbitrary female-to-area relationship does not, so the sampling variance associated with the redd size assignment is zero. The variance of female counts then depends strictly on the numbers of redds in each of the three size classes and is estimated by:

$$\hat{V}\left(\sum_{i=1}^{\hat{C}_\bullet} f(A_i)\right) = \frac{\sum_{i=1}^{\hat{C}_\bullet}(f(A_i) - F\bar{(A)})^2}{\hat{C}_\bullet}, \tag{7.20}$$

assuming that the redds detected, $\hat{B}$, are a random sample of redds present, B, which is reasonable if a constant sampling effort is maintained for the duration of the run (A-29) and detection probability does not change during the run (A-30).

Variability in the escapement estimate contributed by the estimate of observer efficiency does not manifest itself as a binomial error as in Section 2.5 because the study site was surveyed many times, with the location of each detected redd recorded. Instead, detection probability variability is incorporated using the Delta Method and the variance of Equation (7.19) can be estimated by:

$$\begin{aligned} \hat{V}(\hat{N}) &= \hat{V}(\hat{P})\left(\frac{\sum_{i=1}^{\hat{C}_\bullet} f(A_i)(1+\hat{R})}{\hat{P}^2}\right)^2 \\ &+ \hat{V}(\textstyle\sum_{i=1}^{\hat{C}_\bullet} f(A_i))\left(\frac{1+\hat{R}}{\hat{P}}\right)^2 \\ &+ \hat{V}(\hat{R})\left(\frac{\sum_{i=1}^{\hat{C}_\bullet} f(A_i)}{\hat{P}}\right)^2 \\ &+ 2\hat{Cov}(\hat{P}, \textstyle\sum_{i=1}^{\hat{C}_\bullet} f(A_i))\left(\frac{\sum_{i=1}^{\hat{C}_\bullet} f(A_i)(1+\hat{R})}{\hat{P}^2}\right)\left(\frac{1+\hat{R}}{\hat{P}}\right) \\ &+ 2\hat{Cov}(\hat{P}, \hat{R})\left(\frac{\sum_{i=1}^{\hat{C}_\bullet} f(A_i)(1+\hat{R})}{\hat{P}^2}\right)\left(\frac{\sum_{i=1}^{\hat{C}_\bullet} f(A_i)}{\hat{P}}\right) \\ &+ 2\hat{Cov}(\textstyle\sum_{i=1}^{\hat{C}_\bullet} f(A_i), \hat{R})\left(\frac{1+\hat{R}}{\hat{P}}\right)\left(\frac{\sum_{i=1}^{\hat{C}_\bullet} f(A_i)}{\hat{P}}\right). \end{aligned} \tag{7.21}$$

If the estimates of P, $\sum_{i=1}^{\hat{C}_\bullet} f(A_i)$, and R are independent (A-17), (A-23), and

(A-31) Equation (7.21) is replaced by the formula for the exact variance of a product as given by Goodman (1960):

$$\begin{aligned}\hat{V}(\hat{N}) &= \hat{N}^2 \Big[C\hat{V}^2(\hat{P}) + C\hat{V}^2(\textstyle\sum_{i=1}^{\hat{C}_\bullet} f(A_i)) + C\hat{V}^2(\hat{R}) \\ &- C\hat{V}^2(\hat{P}) C\hat{V}^2(\textstyle\sum_{i=1}^{\hat{C}_\bullet} f(A_i)) - C\hat{V}^2(\hat{P}) C\hat{V}^2(\hat{R}) \\ &- C\hat{V}^2(\textstyle\sum_{i=1}^{\hat{C}_\bullet} f(A_i)) C\hat{V}^2(R) \Big] .\end{aligned} \tag{7.22}$$

7.7 Discussion of Redd Count Methods

In theory, Redd Count methods can provide unbiased estimate of escapement. In practice, Redd Count estimates are biased, largely due to the failure of assumptions.

Peak Redd Count Methods provide only an index of escapement (Table 7.2). Dauble and Watson (1990) compared 24 years of chinook salmon dam counts on the Columbia River (corrected for sport and commercial harvest) with peak redd counts and found that the average adult-to-redd ratio was 16:1 which ranged from 5:1 to 39:1. Dauble and Watson (1997) continued the comparison seven years later and found that while there was a correlation between redd counts and escapement of 0.70, the “wide range of fish-to-redd ratios indicates that aerial redd counts do not provide precise estimates of spawning populations in the Hanford Reach.” They concluded that aerial surveys could not provide an “absolute measure of the spawning population” (Dauble and Watson, 1997). Consequently, Peak Redd Count methods should not be used to estimate salmon escapement.

The Redd Census Method is the best method of estimating escapement when using redd counts. It is limited to streams where escapement is low enough to ensure that redds to not overlap and redds can be individually identified. Even when every redd is counted once, the Redd Census method is subject to strict assumptions about the proportion of true redds built, the number of redds per female, and the male-to-female ratio. Of the redd count methods, it is the most accurate and precise, but should be applied with careful consideration regarding the assumptions (Table 7.2).

Table 7.2: Evaluation of redd count methods for estimating salmon escapement

Name	Estimator	Variance	Assumptions	Empirical	Recommendation
Peak Redd	Poor	None	Unreasonable	Index only	Do Not Use
Redd Census	Fair	Good	Reasonable	Accuracy questionable	Cautionary
Welsh	Fair	Good	Conflicutal	Not much tested	Do Not Use
Strip Transect	Fair	Good	Reasonable	Not much tested	Cautionary
Redd Area	Unknown	Poor	Unknown	Imprecise	Do Not Use

The Welsh (1983) Method was intended to alleviate difficulties caused by redd superimposition by using redd area rather than redd number. However, when redds are superimposed, the extent of overlap is usually unknown. Since total redd area cannot be reliably measured, the Welsh Method cannot provide an unbiased estimate of redd number. Consequently, the circumstances for which the Welsh Method was designed violate the assumptions supporting the method (Table 7.2). The Welsh Method should not be used to estimate salmon escapement.

The Strip Transect Method is a legitimate method for estimating the total number of true redds, but should be applied with caution (Table 7.2). Underwater surveys ensure that redds are not overlooked and provide an opportunity to check the proportion of true redds. Flagging redds ensures that no redds are counted multiple times. The total number of redds in a system must be estimated so the strip transect method has a larger variance than the redd census method.

The direction and magnitude of bias in the Redd Area Method is unknown due to the unknown applicability of the female area function. Briggs (1953) estimated the average redd area of coho to be six square feet, or about 0.56 m^2. Under the categorization scale and arbitrary area-to-female ratio of Gallagher and Gallagher (2005), a female coho would have to build four redds this size to be counted as a single female. There is evidence that female coho build up to four redds, but four is at the extreme high end of the range. If the average coho builds only three redds of average size, then the female-area relationship described in Section 7.6.2 underestimates the true number of females. If a much larger redd is found, it is only counted as belonging to a single female, when it is more likely that multiple nests have been built in close proximity resulting in overlap. Because the delineation of redd area to female salmon is arbitrary, the Redd Area Method should not be used to routinely estimate salmon escapement, unless the area-to-female ratio is directly measured (Table 7.2).

Chapter 8

CARCASS COUNT METHODS

8.1 Introduction to Carcass Count Methods

Every spawning Pacific salmon dies. Since every spawner produces exactly one carcass, an estimate of the number of carcasses in a stream can serve as an estimate of escapement. Knudsen (2000) found that escapement estimates for only nine populations of salmon in the United States, all of them chinook, were regularly made using carcass counts. A StreamNet Query showed 187 distinct river sections that had ongoing escapement estimates derived from carcass counts from 1997 and 2007. Of these, one was for chum, four were for coho, and the rest were for chinook from tributaries to nine different rivers.

8.2 Carcass Count Notation

The following notation will be used in this chapter. Not every technique uses all of the notation, but in all cases, estimates are denoted with hats. Escapement, N, for example is estimated by $\hat{N}$. Totals are denoted with a dot in the subscript, for example, $N = \sum_{j=1}^{K} D_j = D_{\bullet}$. The C used in this chapter refers to carcasses rather than redd counts as in Chapter 7.

B_i^*	The number of carcasses “arriving” after the ith survey and surviving until the next survey.
C_i	The number of carcasses detected in the stream during the ith survey.
$Cov(x, y)$	Covariance of x and y.
$CV(x)$	Coefficient of variation of x.
D_i	The number of fish dying after the ith survey, $i = 1, \ldots, k$.
D_j	The number of fish dying in the stream on the jth day of the run, $j = 1, \ldots, K$.
$E(x)$	The expected value of x.
m_i	The number of marked carcasses detected during the ith survey.
M	Total number of marked carcasses released.
K	Number of days in run season.
k	Number of surveys conducted during the run, $1 \leq k \leq K$.
$\mathbf{P}$	The k x k diagonal matrix of detection probabilities.
p_i	The probability that a carcass present on day i will be detected, and removed from the system.
Δ	The k x K matrix of carcass retention probabilities.
Ξ	The K x k matrix defining linear interpolation of counts between surveys.
N	Total escapement.
n_i	The number of fish captured during survey i.
ϕ_i^*	Retention rate of carcasses present just after the ith survey.
R_i	The number of tagged fish released on survey i.
$R_{t_i,j}$	The probability that a fish dying on day j will be retained in the river until day i.
t_i	Time of the ith survey.
$V(x)$	Variance of x.

8.3 Carcass Count Assumptions

The following assumptions are required by one or more of the methods described in this chapter. Not every assumption is required by each method and no method requires all of the assumptions. Each assumption is discussed when introduced into a method.

(A-1). Carcasses are counted once.

(A-2). Surveys are conducted until there are no more live fish in the river.

(A-3). Carcasses are not removed from the study area except by surveyors.

(A-4). Carcasses are detected with 100% probability.

(A-5). All carcasses have the same probability of being tagged.

(A-6). All carcasses have the same probability of detection.

(A-7). All carcasses present during a survey have a common probability of "surviving" to the next.

(A-8). Survival of carcasses is constant over the course of the run.

(A-9). Tagging does not affect the detectability of a carcass.

(A-10). Carcasses mix randomly.

(A-11). Detection probabilities of carcasses are known for each survey without error.

(A-12). Carcass retention rates are known without error.

(A-13). All fish remain in the study area until they have died i.e. no fish are washed out of the study area in a moribund state.

(A-14). Each stream survey is instantaneous.

(A-15). Surveys are conducted every day of the run.

(A-16). Mortalities occurring between two surveys can be interpolated linearly between the two surveys.

8.4 Simple Carcass Counts

The total number of unique carcasses found can be used as an estimate of the minimum number of spawners present in a stream. To ensure that each carcass is only counted once, they are often "dead pitched," or are removed from the river and left on the shore above the high water mark (Frith and Nelson, 1994; Schubert, 1998). Alternately, or in addition, carcasses are cut in half (Johnston et al., 1986; Maselko et al., 2003) or have their tails removed (Bue et al., 1998) before they are replaced in or above the stream to signify that they have been counted. Usually streams are surveyed for carcasses once every seven to fourteen days (e.g. Frith and Nelson, 1994) from the time the first salmon arrive until there are no more live salmon in the river. In some cases surveys are performed daily (e.g. Bue et al., 1998).

Escapement estimates based on simple carcass counts are usually reported as the sum of counts from each of k surveys:

$$\hat{N} = \sum_{i=1}^{k} C_i \tag{8.1}$$

The estimate in Equation (8.1) assumes that every carcass is counted once (A-1). This requires that carcasses are dead-pitched to prevent multiple counting and tht surveys occur until all live fish have become carcasses (A-2). It also requires that fish are not removed from the stream except by surveyors (A-3), and that every carcass present is detected (A-4). Carcasses and moribund males are often washed out of the stream and out of the study area (Simpson, 1984; Johnston et al., 1986; Schubert,

2000) preventing them from being detected and counted. These losses can be reduced by a weir at the downstream end of the study area (Simpson, 1984), but weirs are often rendered inoperable during periods of high run-off, precisely when the rate of carcass loss is greatest (Johnston et al., 1986; Shardlow et al., 1986). Additional carcasses are lost to scavengers or decay, and live fish may be removed by predators such as bears or bald eagles (Dunwiddie and Kuntz, 2001; Gende et al., 2001). Even if carcasses are not removed from the study stream, they may not be detected by surveyors. Gangmark and Fulton (1952) could not use carcass counts to supplement their estimate of escapement because the carcasses settled into deep pools and out of their detection range.

Equation (8.1) would be unbiased if assumptions were met, but they are rarely met. Shardlow et al. (1986) asserted that simple carcass counts were useful only as an index of abundance and did not provide a reliable estimate of escapement. Typically, a simple carcass count will undercount escapement.

While the variance of a simple carcass count is zero, the mean squared error can be substantial.

8.5 Mark-Recapture of Carcasses

Due to the potential for handling stress and mortality on live salmon, some mark-recapture studies are done entirely on salmon carcasses (e.g. Shardlow et al., 1986; Conrad, 2000). Carcasses are found in a river or stream, tagged, and replaced. Later carcasses are recovered and the total number in the system is estimated in the same way as it is in live mark-recapture experiments. Carcass populations are open to immigration in the form of additional mortalities and to emigration due to carcass loss. Therefore, Shardlow et al. (1986) recommended the Jolly-Seber Method (Section 4.5.1) to estimate escapement based on the mark and recapture of carcasses, though closed population methods have also been applied to carcass mark-recapture counts.

8.5.1 Modified Pooled Petersen Method

Shardlow et al. (1986) tagged carcasses from the Quinsam Hatchery with neutrally-colored spaghetti tags and released them into the Quinsam and Campbell Rivers. They recovered both tagged and untagged carcasses on subsequent surveys and used the Pooled Petersen Estimator as described in Section 4.4.1 to estimate the total number of carcasses. This estimate included the hatchery carcasses added to the river, so they subtracted the number of hatchery carcasses from the Pooled Petersen estimate:

$$\hat{N} = \left[\frac{MC_{\bullet}}{m_{\bullet}}\right] - M \tag{8.2}$$

The estimator in Equation (8.2) is biased in the same way as Equation (4.2) described in Section 4.4.1 and can be improved by using the Chapman modification (Section 4.4.1). The Chapman modification only provides nearly-unbiased estimates if all carcasses have the same probability of being tagged (A-5) or the same probability of detection (A-6). Since Shardlow et al. (1986) introduced hatchery fish to the rivers, and only hatchery fish were tagged, their method required that all hatchery carcasses had the same probability of detection as wild carcasses. The probability of detection is only the same for salmon carcasses if there is no difference between an untagged wild fish and a tagged hatchery fish. Shardlow et al. (1986) expressed concern that hatchery fish may have different buoyancy or decomposition rates which would lead to differences in recoverability and therefore biased escapement estimates. Differences in placement of tagged carcasses can also lead to differences in detectability. Shardlow et al. (1986) recommended further research be conducted to test the effects of introducing hatchery fish and carcass placement. Differences in buoyancy and decay rate are not a concern when only wild carcasses from the stream are used and carcasses are replaced in the location from which they were removed (e.g. Frith and Nelson, 1994).

Due to losses of tagged and untagged carcasses, the true number of tagged carcasses in a stream at any given time is unknown, and less than the number released. Consequently, using M in Equation (8.2) will lead to an overestimate of escapement. Shardlow et al. (1986) showed through Monte Carlo simulation that this bias is increased during high washout rates. Thus, the higher the probability of a carcass being washed out, the larger the positive bias of the Carcass Pooled Petersen estimator. Law (1994) also showed via simulation that the Pooled Petersen estimator applied to carcasses consistently over estimated true population abundance.

The variance of the Chapman modified Carcass Pooled Petersen estimator is essentially the same as Equation (4.7):

$$\hat{V}(\hat{N}) = \frac{(M+1)(C_\bullet+1)(M-m_\bullet)(C_\bullet-m_\bullet)}{(m_\bullet+1)^2(m_\bullet+2)}. \tag{8.3}$$

The number of tagged carcasses, M, is assumed known without error, provided the population is closed to emigration (A-3). Since carcasses are lost, Equation (8.3) is positively biased for the variance of the escapement estimate.

8.5.2 Carcass Schaefer Method

The Schaefer Method is "less affected by the [loss of tags] caused by washout of tagged carcasses" than the Pooled Petersen Method (Shardlow et al., 1986). The development of the model is the same as described in Section 4.4.2, only here it is applied strictly to carcasses. Shardlow et al. (1986) released tagged chinook carcasses on two occasions in the Quinsam and Campbell Rivers (with four total distinct tag colors) and recovered carcasses on 18 and 21 occasions over the course of the season. For each river and release the estimate of escapement was:

$$\hat{N}_{ij} = \left[\frac{m_{ij}M_iC_i}{m_{i\bullet}m_{\bullet j}}\right] \tag{8.4}$$

where i indicates release period and j indicates recovery period. The total escapement estimate is the sum of escapements from each strata:

$$\hat{N} = \sum_i \sum_j \hat{N}_{ij} = \hat{N}_{\bullet\bullet} \tag{8.5}$$

The Shaefer Method requires that all fish have some non-zero probability of being tagged. If tagging does not commence at the first mortality, or if there are gaps in the pattern of tag releases, then the estimate given in Equation (8.5) will be negatively biased (Shardlow et al., 1986). The estimate may also be negatively biased if recoveries are not carried out until the end of carcass availability (A-2).

Like the Petersen Estimator, the Schaefer Estimator requires that the population remain closed (A-3). Law (1994) compared the performance of the Schaefer Estimator for varying survival rates, catch rates, and tag rates of live fish while allowing for carcass loss. He found that in all cases where carcasses were lost, the Schaefer Estimator was positively biased. The bias was reduced by increased catch rates, but no 90% confidence intervals overlapped the true estimate (Law, 1994).

Shardlow et al. (1986) performed a Monte Carlo simulation to estimate the precision of the Carcass Schaefer Estimator. They found that the coefficient of variation of the estimate given in Equation 8.5 was approximately equal to the inverse square root of the number of tags recovered, $m_{\bullet\bullet}$ (Shardlow et al., 1986):

$$CV(\hat{N}) = \frac{V(\hat{N})}{N} = \frac{1}{\sqrt{m_{\bullet\bullet}}}. \tag{8.6}$$

The estimate of variance is therefore:

$$\hat{V}(\hat{N}) = \frac{\hat{N}}{\sqrt{m_{\bullet\bullet}}}. \tag{8.7}$$

8.5.3 Carcass Jolly-Seber Method

Since the population of carcasses in a spawning area is increased by mortalities and reduced by decay and washout, an open population mark-recapture model is more appropriate than closed population models (Shardlow et al., 1986). Law (1994) found by simulation that the Jolly-Seber model provided the most precise and accurate estimates of abundance for varying ranges of survival, catch rates, and tagging rates, as compared to the Petersen and Schaeffer models.

The Jolly-Seber estimate of escapement using carcasses requires the same steps as it does when applied to live fish. Estimates of the periodic abundances of carcasses, $\hat{N}_i^*$, must be derived, as well as the "survival" probabilities over each sampling period, $\hat{\phi}_i^*$. When applied to carcasses, the survival probabilities are sometimes called "retention probabilities." Periodic abundances and retention probabilities must be combined to estimate the number of new recruits:

$$\hat{B}_i^* = \hat{N}_{i+1}^* - \hat{\phi}_i^*(\hat{N}_i^* - n_i + R_i), \tag{8.8}$$

where R_i is the number of previously tagged carcasses that are detected and returned to the stream during the ith survey and n_i is the total number of carcasses detected. In this case, the new recruits, or "births," are fresh carcasses, i.e. deaths. Like the Jolly-Seber estimate applied to live fish, the Jolly-Seber method applied to carcasses underestimates the true number of new recruits, and therefore underestimates escapement. Conrad (2000) and Fieberg (2002) used the uniform recruitment distribution correction suggested by Schwarz et al. (1993):

$$D_i = \hat{B}_i^* \left[\frac{\ln(\hat{\phi}_i^*)}{\hat{\phi}_i^* - 1} \right] \qquad i = 2, \ldots, k \tag{8.9}$$

and

$$D_1 = \hat{N}_2^* \left[\frac{\ln(\hat{\phi}_1^*)}{\hat{\phi}_1^* - 1} \right]. \tag{8.10}$$

The new recruits correction requires that all fish present during a given sampling interval have a common probability of survival (A-7).

The total number of recruited carcasses is the escapement:

$$\hat{N} = \sum_{i=1}^{k-2} D_i^* + \hat{D}_{k-1} + \hat{D}_k. \tag{8.11}$$

Like the Jolly-Seber estimate applied to live fish, the carcass Jolly-Seber Method cannot estimate the number of fresh carcasses arriving (fish dying) prior to the last ($\hat{D}_k$) or penultimate ($\hat{D}_{k-1}$) surveys. Sykes and Botsford (1986) offered a correction for estimating abundance at the final survey time, allowing for an estimate of D_{k-1}. Their correction requires that survival does not change significantly over the course of the season (A-8), i.e. that the final survival probability can be estimated as the average of the previous estimates. Since stream life generally decreases over the course of the run, survival cannot be assumed to remain constant. A more reliable correction would be to extend surveys until there are no longer live fish in the river (A-2) and therefore no posibility of new carcasses.

The assumption of equal survival for all live fish in a sampling period in the Jolly-Seber model translates to common retention rates of all carcasses in a sampling period (A-7). If age-dependent mortality is present, i.e. if older carcasses are more likely to be removed from the population, then tagged carcasses will be less likely to be recaptured and survival will be underestimated (Manly, 1970). Consequently, the number of new recruits will be over-estimated, as will escapement. While carcasses are subject to both scavenging and decay, the real complication arises due to issues of equal catchability. The Jolly-Seber Method requires that all individuals be equally catchable (A-6), regardless of tag status (A-9). Age-dependence in catchability of carcasses will affect recruit and escapement estimates in the same manner as age-dependent survival. Fresh silver carcasses are easier to detect than decayed brown carcasses, leading to a theoretical positive bias in escapement (Sykes and Botsford, 1986). This discrepancy is exacerbated for smaller carcasses, which decay faster, leading to capture heterogeneity. To reduce age-induced heterogeneity between marked ("older", "decayed") and unmarked ("younger", "fresh") carcasses, Conrad (2000) and Fieberg (2002) omitted recaptures of carcasses that had previously been recaptured (such that carcasses were tagged on initial capture and noted at second capture but not thereafter). Sykes and Botsford (1986) concluded from both simulation and empirical study that using only fresh carcasses in the Jolly-Seber method made the escapement estimates robust to age-dependent detection probabilities and capture heterogeneity. Conversely, Law (1994) found in simulation that the Jolly-Seber estimates based strictly on fresh carcasses significantly underestimated escapement. Removing decayed carcasses from the river ameliorates the effects of age-dependence in the capture probabilities, but does not ensure that all carcasses have equal probability of detection. If carcasses do not mix randomly (A-10), then marked carcasses are more likely to be detected than carcasses that have not yet been captured (A-5). A team of SCUBA divers found that 54% of chinook carcasses in the Stillaguamish River were inaccessible to surveyors in chest waders with long handled gaffs (Hahn et al., 2003). All "inaccessible" carcasses were marked, and of those, only 6.6% were recaptured later by wading surveyors. Thus, because of unequal capture and recapture probabilities, the Carcass Jolly-Seber estimate may be negatively biased by about 50%. Sykes and Botsford (1986) simulated the effect of omitting all captures of carcasses from the bottom of Bogus Creek, CA and discovered only a small positive bias when compared to a wier count. However, they did not indicate what proportion of carcasses had originally been found on the bottom (as compared to floating or on the banks). The fraction omitted may have been a negligably small fraction of the total population, whereas in a larger river ignoring "inaccessible" carcasses may equate to ignoring a significant portion of the total escapement (Law, 1994). The Carcass Jolly-Seber Method also underestimates escapement on account of fish that exit the river while still alive, either having been removed by predators or drifted downstream in a morribund state.

The Jolly-Seber method provides an assymptotic variance estimator (Equation 4.40). However, if carcasses are removed after the first recapture, the variance formula cannot be used. Instead, the bootstrapping technique of Buckland (1980) is used (Conrad, 2000; Fieberg, 2002). Sykes and Botsford (1986) considered bootstrapped confidence intervals to be "more realistic" than "approximate analytical expressions". Bootstrapped standard errors tend to be larger than the analytical standard errors (Sykes and Botsford, 1986, Table 1).

For an example of how to use the Jolly-Seber Method, see Section 4.5.1.

8.6 Daily Dead Method (Skalski and Liao in review)

Liao (1994) and Skalski and Liao (in review) developed an estimate of daily mortalities based on periodic recoveries of dead fish, an assumed detection probability, and a loss rate of carcasses estimated from the "survival" of tagged carcasses.

Total escapement can be expressed as the total number of fish dying each day of the season:

$$N = \sum_{i=1}^{K} D_i = D_\bullet \tag{8.12}$$

Daily mortalities can be estimated as the number of carcasses found during each survey, C_i, adjusted by observer efficiency, p_i, and carcass retention rates, R_{t_ij}, where j is the day of mortality and t_i is the time of the ith survey. Assume that observer efficiency is known (A-11). Assume also that carcass retention rates are known without error (A-12), and that no moribund fish are washed out of the study area (A-13).

Skalski and Liao (in review) modeled the relationship between the expected number of carcasses detected during each survey, C_i, and the new carcasses available each day, D_i as:

$$E(\mathbf{C}) = \mathbf{P} \cdot \Delta \cdot \Xi \cdot \mathbf{D}, \tag{8.13}$$

where $\mathbf{C}$ is the k x 1 vector of detected carcasses from the k surveys, $\mathbf{P}$ is a k x k diagonal matrix of the detection probabilities of carcasses for each of the k surveys, Δ is a k x K matrix containing the probabilities of a carcass remaining in the stream from death until detection, Ξ is a K x k "link" matrix defining a linear relationship in the numbers of new carcasses occurring on days between surveys, and $\mathbf{D}$ is the k x 1 vector of daily deaths.

Solving Equation 8.13 provides an estimate of the numbers of fish dying on each day surveyed:

$$\mathbf{D} = [\mathbf{P} \cdot \Delta \cdot \Xi]^{-1}\mathbf{C}. \tag{8.14}$$

In the carcass retention matrix, Δ, all elements where $j > t_i$ are zero, because a fish dying on day j cannot be dead prior to day j. The first row of the carcass retention matrix has non-zero elements

$$\Delta_{11} = R_{t_1,1}, \qquad \Delta_{12} = R_{t_1,2}, \qquad \ldots \Delta_{1t_i} = R_{1,t_i} = 1,$$

assuming that each survey is instantaneous (A-14). It takes time to conduct surveys so that they are not instantaneous, but if they are performed quickly with regard to the retention rate (i.e. loss) of carcasses, then surveys can be considered instantaneous. The second row must account for both carcass retention rates and the probability that a carcass was not detected (and not removed) during the first sample. The second row elements are

$$\Delta_{21} = R_{t_2,1}(1 - p_{t_1}), \qquad \Delta_{22} = R_{t_2,2}(1 - p_{t_1}), \qquad \ldots \Delta_{2,t_1} = R_{t_2,t_1}(1 - p_1)$$

until the time of the second survey and then

$$\Delta_{2,t_1+1} = R_{t_2,t_1+1}, \quad \Delta_{2,t_1+2} = R_{t_2,t_1+2}, \quad \ldots, \Delta_{2,t_2} = R_{t_2,t_2} = 1.$$

This pattern continues such that the first element in the third row is

$$\Delta_{3,1} = R_{t_3,1}(1 - p_{t_1})(1 - p_{t_2}),$$

and so on.

If surveys are conducted every day, then the link matrix, Ξ is a K x K, identity matrix and survey day mortalities are the daily mortalities (A-15). If surveys are not conducted every day, then the number of carcasses that would be expected to be detected on unsurveyed days must be estimated. Skalski and Liao (in review) assumed that the number of mortalities occurring on days between surveys could be linearly interpolated using surveys preceding and following unsurveyed days (A-16). Essentially, the new-dead estimate for each of the days between surveys is an average of the survey values before and after, weighted by the temporal proximity of the unsurveyed day to each of the surveyed dates:

$$D_j = \frac{t_i -}{t_i - t_{i-1}} D_{t_{i-1}} + \frac{j - t_{i-1}}{t_i - t_{i-1}} D_{t_i} \tag{8.15}$$

for days j when $t_{i-1} \leq j \leq t_i$. The numbers of fish dying each day during the run are estimated (in matrix notation) by left-multiplying the K x k link matrix to the survey-day mortalities:

$$\hat{\mathbf{D}} = \Xi \mathbf{D}_{t_i}. \tag{8.16}$$

The elements the link matrix are:

$$\Xi_{i,c} = \begin{cases} \frac{j}{t_1} & \text{for } c = 1, \quad 1 \leq r \leq t_1 \\ \frac{t_i - j}{t_i - t_{i-1}} & \text{for } c = i-1, \quad 2 \leq i \leq k, \quad t_{i-1} \leq j \leq t_1 \\ \frac{j - t_{i-1}}{t_i - t_{i-1}} & \text{for } c = i, \quad 2 \leq i \leq k, \quad t_{i-1} \leq j \leq t_1 \\ 0 & \text{otherwise} \end{cases}$$

where i denotes the row number (out of K) and c denotes column number (out of k).

The estimated escapement is the sum of the daily mortalities:

$$\hat{N} = 1'\tilde{D}_j = 1'\Xi[\mathbf{P} \cdot \Delta \cdot \Xi]^{-1}\tilde{C}, \tag{8.17}$$

where $1'$ is the 1 x K vector of 1's. Assuming that observer efficiencies (A-11) and retention rates (A-12) are accurate, and that the linear interpolation accurately reflects the daily dead (A-16), then the Daily Dead Estimator is unbiased for escapement (Liao, 1994). If retention between surveys is not linear, then the direction and magnitude of any bias is unknown. That bias can be avoided by recovering carcasses daily (A-15).

Liao (1994) showed that the variance of the estimator in Equation (8.17) is

$$V(\hat{N}) = \mathbf{W} V(\tilde{C}|\tilde{D}) \mathbf{W}' \tag{8.18}$$

where $\mathbf{W} = \mathbf{1}'\Xi[\mathbf{P} \cdot \Delta \cdot \Xi]^{-1}$ and $V(\tilde{C}|\tilde{D})$ is the k x k variance-covariance matrix of the number of detected carcasses in k surveys. The diagonal elements of $V(\tilde{C}|\tilde{D})$ are given by:

$$\begin{aligned} V(C_{t_i}|\tilde{D}) &= (1 - p_{t_i})^2 \textstyle\sum_{h=1}^{t_{i-1}} R^2_{h,t_{i-1},t_i} V(C_{t_{i-1}h}|\tilde{D}) \\ &+ (1 - p_{t_i}) \textstyle\sum_{h=1}^{t_{i-1}} R_{h,t_{i-1},t_i}[1 - (1 - p_{t_i}) R_{h,t_{i-1},t_i}] E[C_{t_i k}|\tilde{D}] \\ &+ \textstyle\sum_{h=t_{i-1}+1}^{t_i} R_{t_i - h}(1 - R_{t_i - h}) D_h. \end{aligned} \tag{8.19}$$

Equation (8.19) can be computed recursively, beginning with $i = 1$.

The off-diagonal (covariance) elements of $V(\tilde{C}|\tilde{D})$ are given by:

$$Cov(c_{t_i}, c_{t_j}|\tilde{D}) = p_{t_j} p_{t_i} \Big[\prod_{l=1}^{i-j} (1 - p_{t_{i-l}}) \Big] \sum_{h=1}^{t_j} R_{h,t_j,t_i} \Big[V(C_{t_j h}|\tilde{D}) - E[C_{t_j h}|\tilde{D}] \Big] \quad (8.20)$$

where $E[C_{t_j h}|\tilde{D}] = (1 - p_{t_i}) R_{h,t_{i-1},t_i} E[C_{t_{i-1} h}|\tilde{D}]$.

The estimate of variance given by Equation (8.18) is based on the assumption that retention rates are known without error (A-12). Since the retention rates must be estimated, Equation (8.18) underestimates the true variance of the escapement estimate.

For an example of the Daily Dead Method, see (Liao, 1994, page 92).

8.7 Discussion of Carcass Count Techniques

Certain Carcass Count methods provide a theoretically unbiased escapement estimate, but have not been emirically tested. None of the Carcass Count methods are recommended as "Best Choice" methods.

The major assumption of the Simple Carcass Count and both closed population mark-recapture methods of estimating escapement based on carcass counts is that the population is closed to carcass loss. In reality, salmon bodies are removed from the study area by predators, scavengers, floods, and decay. For the Simple Carcass Count this results in an undercount of escapement. For both the Pooled Petersen and the Schaefer Methods, carcass loss results in an overestimate of escapement. Neither closed population method should be used to estimate salmon escapement (Table 8.1).

The Jolly-Seber mark-recapture method seems to be more applicable to salmon carcasses than the closed population mark-recapture models. It allows for carcass loss but relies on the assumption that all carcasses are equally catchable, which is usually not the case. The Carcass Jolly-Seber Method may be robust to age and size-dependent catchability if decayed carcasses are removed after second capture,

Table 8.1: Evaluation of carcass count methods for estimating salmon escapement

Name	Estimator	Variance	Assumptions	Empirical	Recommendation
Simple carcass count	Fair	Zero	Unreasonable	Negatively Biased, index only	Do Not Use
Carcass Petersen	Fair	Good	Unreasonable	Positively Biased	Do Not Use
Carcass Schaeffer	Fair	Good	Unreasonable	Positively Biased	Do Not Use
Jolly-Seber	Poor	Fair	Unreasonable	Not much tested	Cautionary
Daily Dead	Good	Fair	Unreasonable	Not much tested	Cautionary

though emperical evidence is limited, and simulations provide contradictory results. Great care must be taken to ensure that all carcasses are accessible to surveyors, lest significant portions of an escapement be missed and abundance be underestimated. Salmon deaths (new carcasses) occuring after the last survey cannot be estimated, contributing negative bias to the estimate unless a study is carried out until no live salmon are present. Unaccounted-for predation and morribund drifters contribute further negative bias. The Carcass Jolly-Seber Method is recommended only with caution for limited empirical evidence regarding its robustness to model biases and violations of assumptions (Table 8.1).

The Daily Dead Method includes carcass loss in the development of the estimator. In theory, the Daily Dead Estimator is unbiased for escapement and provides a variance estimate, though the variance estimate will be negatively biased in practice. The Daily Dead Method remains to be tested in the field, but may prove to be a valid option for estimating escapement using carcass counts, so it is recommended with caution (Table 8.1).

Chapter 9

DISCUSSION

There are numerous variations on the six types of escapement estimators. Below we summarize "the good" methods, "the bad" methods, and the "ugly" escapement estimation methods. We discuss five issues that managers need to consider when selecting an escapement estimation method, and suggest six areas for future research.

9.1 The Good

Passage counts (Chapter 5) are the most accurate and precise methods of estimating salmon escapement, as they are capable of providing complete censuses of migrating salmon populations (Table 9.1). They are most accurate when enumerating fish 24 hours per day every day of the run, but can still provide an unbiased estimate of escapement based on probabilistic sampling of the run. Interdam Enumeration (Section 5.4) can estimate escapement on a large-scale (i.e. watershed) level. Weirs (Section 5.5) are best used in small tributaries where the risk of weir failure due to washout or debris load is small. Sonar techniques (Section 5.6) are more resistant to flooding than weirs, but should only be used when species identification is accurate through size distribution or run timing. When species can be reliably identified, the DIDSON is generally recommended over split-beam transducers because it provides higher resolution data (Maxwell et al., 2007). However, DIDSON sonars have been limited to a range of about 15 meters at high frequencies (1.6 MHz) (Lilja et al., 2008). A high-resolution lens is available which extends the range of high-frequency (1.2 MHz) DIDSON sonars to about 30 meters (Deborah Burwen, personal communication).

Low-resolution DIDSON sonars (at 0.700 MHz) have a range of 80 meters, but offer low resolution (Sound Metrics Corp., 2009). Split-beam echosounders have ranges of more than 100 meters (Burwen and Fleischman, 1998; Xie et al., 2002). Both sonars require suitable river bottom profiles and weir structures in order to ensure that all fish pass through the ensonified volume of water.

For streams where live salmon are visible to surveyors, Area-Under-the-Curve methods (Chapters 2 and 3) can provide unbiased estimates of escapement with reasonable assumptions (Table 9.1). The Average Spawner Method (Section 2.7.4) provides the best estimate of fish-days for any number of survey days and an estimate of variance for two or more surveys. The Trapezoidal Method (Section 2.7.2) of estimating fish-days can provide an unbiased estimate of fish-days only if the run is surveyed every day. The estimate from either method will be more precise as the number of surveys increases.

The best methods to estimate average stream life uses timer tags developed by Shardlow et al. (2007) because timer tags directly measure time from tagging to death with no discrepancy between time of death and carcass discovery. The parametric estimate (Section 3.4.16) generated from timer tag data will be more precise than the non-parametric estimate, because it estimates fewer parameters. If the data do not follow a parametric distribution, then the non-parametric estimate (Section 3.4.15) should be used, as it will likely be more accurate. The tag-depletion method can also be valid for estimating stream life, but requires that salmon be surveyed every day.

In general, mark-recapture methods (Chapter 4) do not provide unbiased estimates of escapement due to difficulty in fulfilling assumptions. The Darroch Method (Section 4.4.3) as used by Underwood et al. (2007) is an exception (Table 9.1). Typically for the Darroch Method the assumption of a closed population is violated by the loss of carcasses, but Underwood et al. (2007) ensured a closed population by tagging and recapturing only live fish migrating through a closed section of the river. The Change-In-Ratio Method (Section 4.7) is another method that can provide an

Table 9.1: Summary of the recommended (“Good Alternative” or “Best Choice”) methods of estimating salmon escapement. All recommended estimators are rated either “Fair” or “Good,” have variance estimators that are either “Good” or “N/A” (as in a census) and have “Reasonable” assumptions.

Type	Name	Section	Implementation
AUC:	Trapezoidal	2.7.2	High observer efficiency
Fish-days	Average Spawner	2.7.4	High observer efficiency
AUC:	Tag-depletion	3.4.7	High observer efficiency
Stream life	Live Non-parametric	3.4.11	$\geq$10 recaptures
	Live and Dead Non-parametric	3.4.12	$\geq$10 recaptures, access to carcasses
	Live Parametric	3.4.13	$\geq$10 recaptures
	Live and Dead Parametric	3.4.14	$\geq$10 recaptures, access to carcasses
	Timer Tags, Non-parametric	3.4.15	Rivers or tributaries with access to carcasses
	Timer Tags, Parametric	3.4.16	Rivers or tributaries with access to carcasses
M/R:	Darroch	4.4.3	Closed migration route, $\geq$10 recaptures per stratum
	Change In Ratio (CIR)	4.7	Rivers with hatcheries that mark smolts before release
Passage:	Interdam Enumeration	5.4	Basin wide estimates where dams are present
	Weir Counts	5.5	Tributaries with low washout risk
	Sonar Counts	5.6	Single species

unbiased estimate of escapement. It requires reasonable assumptions, but is limited to streams with hatcheries downriver of wild spawning grounds.

9.2 The Bad

Peak Count methods (Chapter 6), whether applied to live fish or to redds, do not provide estimates of fish abundance. Even when there is a unique peak, the relationship between peak count and escapement for a given year is unknown. Peak-escapement relationships are inconsistent from year to year due to differences in run duration and arrival distributions. Any estimate generated by expanding a peak count using a conversion factor is likely to be so imprecise as to be uninformative.

For the most part, Mark-Recapture Methods (Chapter 4) are biased for salmon escapement because of the difficulty in fulfilling assumptions. Closed-population methods such as the Pooled Petersen (Section 4.4.1) and the Schaefer (Section 4.4.2) methods fail because spawning salmon populations are not closed. The number of marked fish which is assumed known, is not known. The Jolly-Seber Method (Chapter 4.5.1) fails because 1) it does not allow for the estimation of new arrivals at the start and finish of the experiment, 2) it does not account for fish that both arrive and die within the same period, and 3) it assumes that all fish present during a sampling interval have the same probability of survival. Adjustments have been made to estimate the numbers of new arrivals at the start and end of a study, and for those that arrive and die between surveys. No adjustment can be made for bias caused by mortality probabilities dependent on arrival time. The Manly-Parr Method (Section 4.5.4) requires no assumptions about mortality and is therefore more robust than the Jolly-Seber method, but is unable to account for new arrivals that die before the next survey.

9.3 The Ugly

Theoretically, redd counts (Chapter 7) provide unbiased estimates of escapement. The redds can be easily censused, though some redds may be missed if they are located in deep water. Of the redds detected, an estimate of the proportion of true redds (i.e. the proportion of counted nests that have eggs) is required. The assumption of one female per true redd seems reasonably met, though there have been studies testing this assumption, so the number of true redds can give an accurate estimate of the number of female spawners. In practice redd counts perform poorly when compared to other total escapement estimates, likely due to the use of a biased gender ratio. Gender ratios often change over the course of a run, and must be estimated throughout a season to provide an unbiased estimate. The male-female ratio is usually estimated using recovered carcasses, but the sex ratio of carcasses is typically lower than that of live fish (fewer males than females) because males are more likely to be washed downstream or sink into deep pools than females.

For the same reason, carcass count methods do not provide unbiased estimates of escapement. The Carcass Jolly-Seber Method (Section 8.5.3) and the Daily Dead Method (Section 8.6) may provide unbiased estimates of the total number of carcasses created, and the total escapement, but do not account for fish that are washed from the stream in a moribund state or removed alive by predators. If the number of moribund salmon washed downstream is negligible, predation rates are low, and all carcasses are accessible to surveyors, then the Carcass Jolly-Seber Method and the Daily Dead Method could theoretically be applicable. However, neither method has been rigorously tested. In addition, both methods estimate large numbers of parameters, resulting in large variance estimates.

9.4 Considerations for Managers

No single method of estimating salmon escapement is the best method for all circumstances. An escapement estimation method that works well in one river, may work poorly in another. It would be impossible to make a rule assigning a single method to a particular type of river or a specific salmon species. There are five issues managers should consider when selecting an escapement estimation method. The assumptions for any selected method should be reviewed with respect to the characteristics of the river or run to be estimated.

1. **Geographic Scale**

   The first consideration is geographic scale of the estimate. For basin-wide estimates of escapement, the Interdam Enumeration Method (Section 5.4) is the most practical. Facilities to count salmon passage over dams are already established in many places and provide large-scale estimates of salmon abundance. There are other options for estimating escapement to specific rivers or tributaries, i.e. on finer scales.

2. **River Dimensions**

   The depth and width of a river may prevent some methods from being effective. Constructing a weir across a wide river may be impractical, especially if periods of high flow will wash the weir out. On narrow or shallow rivers a weir may be the best way to estimate escapement, especially if multiple species migrate concurrently. River depth and width may limit the probability of seeing live fish on a single pass, or the probability of recovering carcasses (for stream life estimates) so Area-Under-the-Curve methods may not be applicable. On very large rivers, the only options for escapement estimation may be the Darroch Method of Mark-Recapture (Section 4.4.3) or Sonar (Section 5.6). The Darroch method can be used on any portion of the river where the migration route is closed, does not depend on water clarity, and can be used regardless of the number of species present, as long as at least ten tags are recovered from each stratum. Sonar can be used in rivers of any size, as long as only one species migrates past at a time (or the size distribution of species can be determined to distinguish between species), and there is a deployment suitable location.

3. **River Characteristics**

   River discharge, flow changes, water clarity, degree of plant overgrowth, and presence of deep pools and other fish hiding places should all be taken into account when selecting an escapement estimation method. A river with high discharge may not be suitable for a weir, or for methods requiring the recovery of carcasses (i.e. AUC stream life estimates). Large fluctuations in discharge may preclude certain equipment from being used, i.e. fish-wheels in a Mark-Recapture study. A river with clear water may be ideal for Area-Under-the-Curve methods (Chapters 2 and 3), which require periodic surveys but not continuous attention. Sonar (Section 5.6) may be the only means of detecting salmon in a low visibility, i.e. glacial, stream.

   Overgrowth, deep pools, and other fish hiding places influence the probability of live and carcass detections (observer efficiency), which may also influence the selection of methods. Observer efficiency is crucial to visual count methods, particularly Area-Under-the-Curve methods (Chapters 2 and 3). The AUC methods require a high live fish detection probability for the fish-days estimate, and a high probability of carcass recovery for the stream life estimate. “High” is relative and can be defined according to the purpose of a study (see “Level of Precision”). However it must be remembered that the lower the observer efficiency, the lower the precision of the escapement estimate, and that an unbiased estimate of observer efficiency is necessary for an unbiased estimate of escapement.

4. **Species Characteristics**

   Life-history and behavior of each estimated species must be taken into account. Chinook have been shown to delay or reverse migration when tagged, making them poor candidates for Mark-Recapture methods which rely on the assumption that tagging has no effect on behavior. Weirs used to count chum should be built high enough to prevent chum from jumping over (McNeil, 1966). Sockeye spawn in streams and lakes, so study area must be clearly defined, and method assumptions must be evaluated for both habitats. Coho tend to spawn in small streams with low gradients (Quinn, 2005), making them prime candidates for Area-Under-the-Curve methods (Chapters 2 and 3). Pink and sockeye are the most abundant salmon in the Pacific Northwest (Quinn, 2005), and may migrate in such large numbers that sonar and human counters may not be able to distinguish individuals well enough to count them, although such occurances are rare (Enzenhofer et al., 1998). In such cases, the Darroch Method of Mark-Recapture (Section 4.4.3) may be the only valid option.

5. **Level of Precision**

   Some escapement estimation methods are more precise than others and the manner in which a study is conducted will influence the level of uncertainty in the resulting estimate. Passage Counts (Chapter 5) can provide a census with no variability in the estimate, but when temporal sampling is introduced, so is uncertainty. It is left to managers and policy makers to determine necessary precision levels. Strongly influencial escapement estimates require greater precision; less influential estimates require less precision.

9.5 Areas of Further Research and Development

Knowledge pertaining to salmon escapement continues to expand as technologies and methodologies continue to advance. In this report we have considered the accuracy, precision, and assumption validity of escapement estimation methods used in the Pacific Northwest. Further research should focus on the following topics.

1. **Promote greater awareness of benefits and limitations of available escapement methods and tools.** Knowing the appropriate applications for a suite of recommended methods (see Table 9.1) will allow fisheries managers greater flexibility in selecting suitable escapement estimators. Advances in computing power and statistical software packages may aid escapement estimation by providing sophisticated computations quickly and efficiently. The package USER, for example, allows quick and easy computations for the Change-in-Ratio method (Skalski and Millspaugh, 2006). Methods that were previously too computationally demanding are more accessible to managers due to new computational tools and computing power.

2. **Further empirical testing.** Advances in technology and theory mean little if they cannot be applied in the field. Ultimately, the merits of any escapement estimation method depend on the accuracy and precision of the estimator rather than on how accurate and precise it can be. Assumptions for all methods used should undergo continued testing in the field, particularly the Carcass Jolly-Seber (Section 8.5.3) and the Daily Dead (Section 8.6) methods, which have little published information on their empirical performance.

3. **Better estimation of support parameters.** Accurate estimates of detection probability (or observer efficiency) are essential for unbiased escapement estimates, with the exception of mark-recapture techniques. Uncertainty in detection probability increases uncertainty in the escapement estimate. Care should be taken to ensure that the estimate of detection probability is as accurate and precise as possible. Other support parameters, such as the male-to-female, true-to-false redd, and redd-to-female ratios used

in Redd Count methods (Chapter 7), should be estimated regularly and with great care.

4. **Probabilistic survey date selection for Area-Under-the-Curve methods (Chapters 2 and 3).** Survey dates are usually chosen systematically, invalidating a complete estimate of the variance. Changing the selection process to a stratified random sample of the days in a run, i.e. stratified by week, would allow for a complete variance estimator, while still providing a pre-season survey schedule and a way to extend sampling to the end of a run.

5. **Better resolution and species identification for sonar.** Further research is currently underway to improve the resolution of DIDSON sonars. Improved resolution, coupled with length-frequency data, could potentially be used to distinguish between fish of different species. Reliable species identification would allow sonar passage counts to be effective in streams where salmonid runs overlap.

6. **Multi-objective sampling.** In addition to total escapement, researchers are often interested in other characteristics of a run. Some agencies desire to estimate the age distribution of spawning salmon, which requires scale samples. Usually scales are taken from fish carcasses (Hahn, WDFW, personal communication), so an escapement estimate using carcasses, such as stream life estimates for AUC (Chapter 3), would seem the most appropriate.

   It may be desirable to estimate different components of the escapement separately. Separate estimates for hatchery and wild origin fish may be necessary to assess the affects of the hatchery on the wild population. The Change-in-Ratio Method (Section 4.7) would be the most applicable escapement method. The number of escaping wild origin female salmon may be a more accurate indicator of the recovery and future success of a run. In that case, if females cannot be enumerated separately, an estimate of the sex ratio (and possibly hatchery-to-wild origin ratio) would be necessary.

We have not considered cost in my evaluation of methods, though cost (due to sampling, equipment, analysis, etc.) will certainly play a role in determining the feasibility of a method.

From a logistic perspective, there will be many criteria for the "best" escapement estimator, but statistical accuracy, precision, and assumption validity should always be considered. It is my hope that this review will allow managers to choose a method from among the "Best Choices."

BIBLIOGRAPHY

Adkison, M. D., and Z. M. Su. 2001. A comparison of salmon escapement estimates using a hierarchical Bayesian approach versus separate maximum likelihood estimation of each year's return. Canadian Journal of Fisheries and Aquatic Sciences **58**:1663–1671.

Al-Chokhachy, R., P. Budy, and H. Schaller. 2005. Understanding the significance of redd counts: a comparison between two methods for estimating the abundance of and monitoring bull trout populations. North American Journal of Fisheries Management **25**:1505–1512.

Alexandersdottir, M., 2005. Nisqually 2003 chinook salmon abundance estimates from change in ratio. Technical report, Northwest Indian Fisheries Commission.

Ames, J., A. Burgh, and C. Morrill, 1974. Report on the salmon escapement in the State of Washington, 1971. Technical report, Washington State Department of Fisheries, Management and Research Division.

Ames, J., and D. E. Phinney. 1977. 1977 Puget Sound summer-fall Chinook methodology: escapement estimates and goals, run size forecasts, and in-season run size updates. Washington Department of Fisheries Technical Report **29**.

Ames, J. J., 1984. Puget Sound chum salmon escapement estimates using spawner curve methodology.

Bailey, N. T. J. 1951. On estimating the size of mobile populations from recapture data. Biometrika **38**:293–306.

Bailey, N. T. J. 1952. Improvements in the Interpretation of Recapture Data. The Journal of Animal Ecology **21**:120–127.

Beidler, W. M., and T. E. Nickelson, 1980. An evaluation of the Oregon Department of Fish and Wildlife standard spawning fish survey system for coho salmon. Technical report, Oregon Department of Fish and Wildlife.

Beiningen, K. T., 1976. Fish runs. Pages E1–E65 *in* Investigative reports of Columbia River fisheries project. Pacific Northwest Regional Commission, Portland, Oregon.

Bevan, D. E. 1961. Variability in aerial counts of spawning salmon. Journal of Fishery Research Board Canada **18**:337–348.

Bocking, R. C. 1991. Abundance, age, size, sex and coded wire tag recoveries for chinook salmon escapements of Campbell and Quinsam rivers, 1989–90. Canadian Manuscript Report of Fisheries and Aquatic Sciences **2124**:109.

Bocking, R. C., J. R. Irvine, K. K. English, and M. Labelle. 1988. Evaluation of random and indexing sampling designs for estimating coho salmon (*Oncorhynchus kisutch*) escapement to three Vancouver Island streams. Canadian Technical Report of Fisheries and Aquatic Sciences **No. 1639**:95.

Boggs, C. T., M. L. Keefer, C. A. Peery, T. C. Bjornn, and L. C. Stuehrenberg. 2004. Fallback, reascension and adjusted fishway escapement estimates for adult chinook salmon and steelhead at Columbia and Snake River dams. Transactions of the American Fisheries Society **133**:pp. 930–947.

Briggs, J. C. 1953. The behavior and reproduction of salmonid fishes in a small coastal stream. California Fish and Game, Fish Bulletin **94**.

Buckland, S. 1980. A modified analysis of the Jolly-Seber capture-recapture model. Biometrics **47**:255–268.

Bue, B. G., S. M. Fried, S. Sharr, D. G. Sharp, J. Wilcock, and H. J. Geiger. 1998. Estimating salmon escapement using area-under-the curve, aerial observer efficiency, and streamlife estimates: the Prince William Sound pink salmon example. North Pacific Anadromous Fisheries Commission Bulletin **1**:240–250.

Burwen, D., S. J. Fleischman, S. L. Maxwell, and C. Pfisterer. 2005. A retrospective on hydroacoustic assessment of fish passage in Alaskan rivers. Journal of the Acoustical Society of America **117**:2366–2367.

Burwen, D. L., and S. J. Fleischman. 1998. Evaluation of side-aspect target strength and pulse width as potential hydroacoustic discriminators of fish species in rivers. Canadian Journal of Fisheries and Aquatic Sciences **55**:2492–2502.

Casella, G., and R. L. Berger. 2002. Statistical Inference. Second edition edition. Duxbury Advanced Series, Duxbury, Pacific Grove, California.

Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. University of California Publication of Statistics **1**.

Chapman, D. G., and C. O. Junge. 1956. The Estimation of the Size of a Stratified Animal Population. The Annals of Mathematical Statistics **Vol. 27**:pp. 375–389.

Chapman, W. M. 1943. The spawning of chinook salmon in the main Columbia River. Copeia **1943**:168–170.

Cochran, W. G. 1977. Sampling Techniques. Wiley series in probability and mathematical statistics-applied, John Wiley and Sons, New York.

Conrad, B., 2000. Estimate escapement of chinook salmon to the North fork of the Stillaguamish River–1999. Technical report, Northwest Indian Fisheries Commission Report.

Cormack, R. M. 1964. Estimates of survival from sighting of marked animals. Biometricka **51**:429–438.

Cousens, N., G. Thomas, C. Swann, and M. Healey. 1982. A review of salmon escapement estimation techniques. Canadian Technical Report of Fisheries and Aquatic Sciences **1108**:122.

Crone, R. A., and C. E. Bond. 1974. Life history of coho salmon, *Oncorhynchus kisutch*, in Sashin Creek, southeastern Alaska. Fishery Bulletin **74**:897–923.

Crosbie, S. F., and B. F. J. Manly. 1985. Parsimonious modeling of capture-mark-recapture studies. Biometrics **41**:385–398.

Darroch, J. N. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. Biometrika **48**:241–260.

Dauble, D. D., R. L. Johnson, and A. P. Garcia. 1999. Fall chinook salmon spawning in the tailraces of lower Snake River hyrdroelectric projects. Transactions of the American Fisheries Society **128**:672–679.

Dauble, D. D., and D. G. Watson, 1990. Spawning and abundance of fall chinook salmon (*Oncorhynchus tshawytscha*) in the Hanford Reach of the Columbia River, 1948–1988. Technical report, Pacific Northwest Laboratory.

Dauble, D. D., and D. G. Watson. 1997. Status of fall chinook salmon populations in the mid-Columbia River, 1948-1992. North American Journal of Fisheries Management **17**:283–300.

Daum, D. W., and B. M. Osborne, 1995. Enumeration of Chandalar River fall chum salmon using split-beam sonar, 1994. Technical report, United States Fish and Wildlife Service.

Daum, D. W., and B. M. Osborne. 1998. Use of Fixed-Location, Split-Beam Sonar to Describe Temporal and Spatial Patterns of Adult Fall Chum Salmon Migration in the Chandalar River, Alaska. North American Journal of Fisheries Management **18**:477–486.

Davidson, F. A., 1966. Population trends in the blueback and upper Columbia River fall chinook salmon. Technical report, Public Utility District of Grant County.

Deschamps, G., 1969. Progress report for 1959. Technical report, State of Washington Department of Fisheries Research Division.

Dunwiddie, P. W., and I. Kuntz, Robert C. 2001. Long-Term Trends of Bald Eagles in Winter on the Skagit River, Washington. The Journal of Wildlife Management **65**:290–299.

Eggers, D. M. 1994. On the Discrimination of Sockeye and Chinook Salmon in the Kenai River Based on Target Strength Determined with 420 kHz Dual-Beam Sonar. Alaska Fishery Research Bulletin **1**:125–139.

English, K. K., R. C. Bocking, and J. R. Irvine. 1992. A robust procedure for estimating salmon escapement based on the area-under-the-curve method. Canadian Journal of Fisheries and Aquatic Sciences **49**:1982–1989.

Enzenhofer, H. J., N. Olsen, and T. J. Mulligan. 1998. Fixed-location riverine hydroacoustics as a method of enumerating migrating adult Pacific salmon: comparison of split-beam acoustics vs. visual counting. Aquatic Living Resources **11**:61–74.

Fieberg, J., 2002. Estimated escapement of chinook salmon to the North fork of the Stillaguamish River - 2001. Technical report, Northwest Indian Fisheries Commission Report.

Fleischman, S. J., and D. L. Burwen. 2003. Mixture models for the species apportionment of hydroacoustic data, with echo-envelope length as the discriminatory variable. ICES Journal of Marine Science **60**:592–598.

Foote, K. G., H. P. Knudsen, G. Vestnes, D. N. MacLennan, and E. J. Simmonds, 1987. Calibration of acoustic instruments for fish density estimation: a practical guide. Technical report, International Council for the Exploration of the Sea, Cooperative Research Report No. 144.

Frith, H. R., and T. C. Nelson. 1994. Abundance, age, size, sex and coded wire tag recoveries for chinook salmon escapements of Campbell and Quinsam Rivers, 1993. Canadian Manuscript Report of Fisheries and Aquatic Sciences **2251**:ix+59 p.

Fukushima, M., and W. W. Smoker. 1997. Determinants of stream life, spawning efficiency, and spawning habitat in pink salmon in the Auke Lake system, Alaska. Canadian Journal of Fisheries and Aquatic Sciences **54**:96–104.

Fuller, W. A. 2006. Measurement Error Models. Wiley series in probability and statistics, John Wiley and Sons, Inc., Hoboken, New Jersey.

Gallagher, S. P., and C. M. Gallagher. 2005. Discrimination of chinook salmon, coho salmon, and steelhead redds and evaluation of the use of redd data for estimating escapement in several unregulated streams in Northern California. North American Journal of Fisheries Management **25**:284–300.

Gangmark, H. A., and L. A. Fulton, 1952. Status of Columbia River blueback salmon runs, 1951. Technical report, United States Fish and Wildlife Service Special Scientific Report, Fisheries.

GCPUD. 2008. Fish Counts. Grant County Public Utility District website, http://www.gcpud.org/resources/resFish/fishCounts.htm.

Geist, D. R., T. P. Hanrahan, E. V. Arntzen, G. A. McMichael, C. J. Murray, and Y. J. Chien. 2002. Physiochemical characteristics of the hyporheic zone affect redd site selection by chum salmon and fall chinook salmon in Columbia River. North American Journal of Fisheries Management **22**:1077–1085.

Gende, S. M., T. P. Quinn, and M. F. Willson. 2001. Consumption choice by bears feeding on salmon. Oecologia **127**:372–382.

Goodman, L. A. 1960. On the exact variance of products. Journal of the American Statistical Association **55**:708–713.

Hahn, P., M. C. Mizell, and C. Kraemer, 2003. Assessment of chinook salmon spawning escapement in the Stillaguamish River System, 2000. Technical report, Completion Report, Fish Management Program, Washington Department of Fish and Wildlife.

Hahn, P., and B. Thompson, 2007. Investigation of possible undercounting of chinook salmon redds in the Green-Duwamish River, Washington: 2006. Technical report, Washington Department of Fish and Wildlife.

Hammarstrom, S. L., and J. J. Hasbrouck. 1998. Estimation of the abundance of late-run chinook salmon in the Kenai River based on exploitation rate and harvest, 1996. Alaska Department of Fish and Game Fishery Data Series **No. 98-6**.

Hammarstrom, S. L., and J. J. Hasbrouck. 1999. Estimation of the abundance of late-run chinook salmon in the Kenai River based on exploitation rate and harvest, 1997. Alaska Department of Fish and Game Fishery Data Series **No. 99-8**.

Hankin, D. G. 1982. Estimating escapement of Pacific salmon: marking practices to discriminate wild and hatchery fish. Transactions of the American Fisheries Society **111**:286–298.

Hatch, D. R., J. K. Fryer, M. Schwartzberg, D. R. Pederson, and A. Wand. 1998. A computerized editing system for video monitoring of fish passage. North American Journal of Fisheries Management **18**:694–699.

Hatch, D. R., M. Schwartzberg, and P. R. Mundy. 1994. Estimation of Pacific salmon escapement with a time-lapse video recording technique. North American Journal of Fisheries Management **14**:626–635.

Haymes, J., 2007. Review of natural escapements, exploitation rates, and natrual production and exploitation rate monitoring methodologies for Washington Pacific Salmon Treaty Coho Managment Units. Technical report, Washington Department of Fish and Wildlife, 46 pp.

Hilborn, R., B. G. Bue, and S. Sharr. 1999. Estimating spawning escapements from periodic counts: a comparison of methods. Canadian Jounal of Fisheries and Aquatic Sciences **56**:888–896.

Hill, R. A. 1997. Optimizing aerial count frequency for the area-under-the-curve method of estimating escapement. North American Journal of Fisheries Management **17**:461–466.

Hoaglin, D. C., F. Mosteller, and J. W. Tukey. 1983. Understanding Robust and Exploratory Data Analysis. Wiley series in probability and mathematical statistics, John Wiley and Sons, Inc., New York, New York.

Hobbs, D. F. 1937. Natural reproduction of quinnat salmon, brown and rainbow trout in certain New Zealand waters. New Zealand Marine Department Fisheries Bulletin **6**:104 p.

Holecek, D. E., and J. P. Walters. 2007. Spawning characteristics of adfluvial rainbow trout in a North Idaho stream: Implications for error in redd counts. North American Journal of Fisheries Management **27**:1010–1017.

Holmes, J., G. Cronkite, and H. J. Enzenhofer. 2005. Feasibility of deploying a dual-frequency identification sonar (DIDSON) system to estimate salmon spawning ground escapement in major tributary systems of the Fraser River, British Columbia. Canadian Technical Report of Fisheries and Aquatic Sciences **2592**:66 pp.

Iglewiez, B., 1983. Robust Scale Estimators and Confidence Intervals for Location. Pages 404–431 *in* D. C. Hoaglin, F. Mosteller, and J. W. Tukey, editors. Understanding Robust and Exploratory Data Analysis. John Wiley and Sons, Inc., New York, New York.

Irvine, J. R., R. C. Bocking, K. K. English, and M. Labelle. 1992. Estimating coho salmon (*Oncorhynchus kisutch*) spawning escapements by conducting visual surveys in areas selected using stratified random and stratified index sampling designs. Canadian Journal of Fisheries and Aquatic Sciences **49**:1972–1981.

Irvine, J. R., J. F. T. Morris, and L. M. Cobb. 1993. Area-under-the-curve salmon escapement estimation manual. Canadian Technical Report of Fisheries and Aquatic Sciences **1932**:84.

Jacobs, S., 2002. Calibration of estimates of coho spawner abundance in the Smith River Basin, 2001. Technical report, Monitoring Program Report Number OPSW-ODFW-2002-06, Oregon Department of Fish and Wildlife, 14 pp.

Johnston, N., J. R. Irvine, and C. J. Perrin. 1986. A comparative evaluation of fence count, mark-recapture and Bendix sonar estimates of salmon escapements in the Keogh River, a variable-flow Coastal B.C. Stream. Canadian Technical Report of Fisheries and Aquatic Sciences **1453**:44.

Johnston, S. V., and T. W. Steig. 1995. Evolution of fixed-location hydroacoustic techniques for monitoring upstream migrating adult salmonids (*Oncorhynchus* spp.) in riverine environments. ICES International Syumposium on Fisheries and Plankton Acoustics.

Jolly. 1965. Mark recapture models with parameters constant in time. Biometrics **38**:301–321.

Jones III, E. L., T. J. Quinn II, and B. W. Van Alen. 1998. Observer accuracy and precision in aerial and foot survey counts of pink salmon in a southeast Alaska stream. North American Journal of Fisheries Management **18**:832–846.

Junge, C. O. 1963. A quantitative evaluation of the bias in population estimates based on selective samples. International Commission for Northwest Atlantic Fisheries Special Report **4**:pp. 26–28.

Kendall, W. L. 1999. Robustness of closed capture-recapture methods to violations of the closure assumption. Ecology **80**:2517–2525.

Kendall, W. L., K. H. Pollock, and C. Brownie. 1995. A Likelihood-Based Approach to Capture-Recapture Estimation of Demographic Parameters under the Robust Design. Biometrics **51**:293–308.

Killick, S. R. 1955. The chrolological order of Fraser River salmon during migration, spawning, and death. International Pacific Salmon Fisheries Commission Bulletin **VII**.

Knudsen, E. E., 2000. Managing Pacific salmon escapements: the gaps between theory and reality. Pages 237–272 *in* E. E. Knudsen, C. R. Steward, D. D.

MacDonald, J. E. Williams, and D. W. Reiser, editors. Sustainable Fisheries Management: Pacific Salmon. Lewis Publishers, Boca Raton, Florida.

Kutner, M. H., C. J. Nachtsheim, and J. Neter. 2004. Applied Linear Regression Models. Fourth edition. McGraw-Hill/Irwin, New York, NY.

Labelle, M. 1994. A likelihood method for estimating Pacific salmon escapement based on fence counts and mark-recapture data. Canadian Journal of Fisheries and Aquatic Sciences **51**:pp. 552–566.

Lady, J. M., 1996. Release-recapture models for estimating the stream residence time of spawning salmon. Master's, University of Washington.

Lady, J. M., and J. R. Skalski. 1998. Estimators of stream residence time of Pacific salmon (*Oncorhynchus* spp.) based on release-recapture data. Canadian Journal of Fisheries and Aquatic Sciences **57**:241–246.

Law, P. M. W. 1994. Simulation study of salmon carcass survey capture-recapture methods. California Fish and Game **80**:pp. 14–28.

Lee, E. T. 1992. Statistical Methods for Survival Data Analysis. John Wiley& Sons, Inc., New York, New York.

Lewis, B., 2005. The role of sonar in Copper River salmon escapement monitoring and commercial fisheries management. *in* R. Kopchak, editor. Copper River Salmon Workshop Series. EcoTrust Fisheris, Anchorage, Alaska.

Liao, S., 1994. Statistical models for estimating salmon escapement and stream residence time based on stream survey data. Dissertation, University of Washington, 190 pp.

Lilja, J., T. J. Marjomki, R. Riikonen, and J. Jurvelius. 2000. Side-aspect target strength of Atlantic salmon (Salmo salar), brown trout (Salmo trutta), whitefish

(Coregonus lavaretus), and pike (Esox lucius). Aquatic Living Resources **13**:355–360.

Lilja, J., T. Ridley, G. M. W. Cronkite, H. J. Enzenhofer, and J. A. Holmes. 2008. Optimizing sampling effort within a systematic design for estimating abundant escapement of sockeye salmon (*Oncorhynchus nerka*) in their natal river. Fisheries Research **90**:118–127.

MacLennan, D. N., and E. J. Simmonds. 1992. Fishery Acoustics. Chapman Hall, New York.

Manly, B. F. J. 1969. Some Properties of a Method of Estimating the Size of Mobile Animal Populations. Biometrika **56**:407–410.

Manly, B. F. J. 1970. A simulation study of animal population estimation using the caprure-recapture method. Journal of Applied Ecology **7**:13–39.

Manly, B. F. J., T. L. McDonald, S. C. Amstrup, and E. V. Regehr. 2003. Improving size estimates of open animal populations by incorporating information on age. BioScience **53**:666–669.

Manly, B. F. J., and M. J. Parr. 1968. A new method of estimating population size, survivorship, and birth rate from capture-recapture data. Transactions of the Society of British Entomology **18**:81–89.

Manning, T., and S. Smith, 2005. Washington State Sport Catch Report 2001. Technical report, Washington Department of Fish and Wildlife, Available online at http://wdfw.wa.gov/fish/harvest/2001sport.pdf.

Manske, M., and C. J. Schwarz. 2000. Estimates of stream residence time and escapement based on capture-recapture data. Canadian Jounal of Fisheries and Aquatic Sciences **57**:241–246.

Maselko, J., A. Werheimer, and J. Thedinga, 2003. Selection and application of a mark-and-recapture technique for estimating pink salmon escapments. Technical report, NOAA, Report Number NMFS-AFSC-137.

Maxwell, S. L., D. Degan, A. V. Smith, L. McKinley, and N. E. Gove. 2007. The feasibility of using a split-beam sonar to estimate salmon passage on the Nushagak River as a potential replacement for an echo-counting Bendix sonar. Alaska Department of Fish and Game, Fishery Manuscript **No. 07-11**:Anchorage.

Maxwell, S. L., and N. E. Gove. 2004. The feasibility of estimating migratin salmon passage rates in turbid rivers using dual frequency identification sonar (DISON) 2002. Alaska Department of Fish and Game Regional Information Report **No. 2A04-05, Anchorage**.

Maxwell, S. L., and N. E. Gove. 2007. Assessing a dual-frequency identification sonars' fish-counting accuracy, precision, and turbid river range capability. Journal of the Acoustical Society of America **122**:3364–3377.

Maxwell, S. L., and A. V. Smith. 2007. Generating River Bottom Profiles with a Dual-Frequency Identification Sonar (DIDSON). North American Journal of Fisheries Management **27**:1294–1309.

McNeil, W. J. 1964*a*. A method of measuring mortality of pink salmon eggs and larvae. U. S. Fish and Wildlife Service Fishery Bulletin **63**:575–588.

McNeil, W. J. 1964*b*. Redd superimposition and egg capacity of pink salmon spawning beds. Journal of Fishery Research Board Canada **21**:1385–1396.

McNeil, W. J., 1966. Distribution of spawning pink salmon in Sashin Creek, southeastern Alaska, and survival of their progeny. Technical report, United States Fish and Wildlife Service, Special Scientific Report, Fisheries 538.

McPherson, S. A., D. R. Bernard, R. J. Yanusz, P. A. Milligan, and P. Timpany, 1999. Spawning abundance of chinook salmon in the Taku River in 1998. Technical report, Alaska Department of Fish and Game, Division of Sport Fish.

Meeker, W. Q., and L. A. Escobar. 1998. Statistical Methods for Reliability Data. Wiley Series in Probability and Statistics, John Wiley And Sons, Inc., New York.

Meekin, T. K., 1967. Observations of exposed fall chinook redds below Chief Joseph Dam during periods of low flow. Technical report, Douglas County Public Utility District.

Miyakoshi, Y., and S. Kudo. 1999. Mark-recapture estimation of escapement of masu salmon *Oncorhynchus masou* with a comparison to a fence count. North American Journal of Fisheries Management **19**:1108–1111.

Miyakoshi, Y., T. Takami, K. Takeuchi, H. Omori, M. Nagata, and J. R. Irvine. 2003. Sampling of masu salmon, *Oncorhynchus masou*, spawners on the spawning grounds: is carcass sampling effective as a mark-recovery method? Fisheries Management and Ecology **10**:273–275.

Morbey, Y. 2000. Protandry in Pacific salmon. Canadian Journal of Fisheries and Aquatic Sciences **57**:1252–1257.

Mulligan, T. J., and R. Keiser. 1996. A split-beam echo counting model for riverine use. ICES Journal of Marine Science **53**:403–406.

Nealson, P. A., and J. Gregory. 2000. Hydroacoustic differentiation of adult Atlantic salmon and aquatic macrophytes in the River Wye, Wales. Aquatic Living Resources **13**:331–339.

Neave, F. 1953. Principles affecting the size of pink and chum salmon populations in British Columbia. Journal of Fishery Research Board Canada **9**:450–491.

Neilson, J. D., and G. H. Geen. 1981. Enumeration of spawning salmon from spawner residence time and aerial counts. Transactions of the American Fisheries Society **110**.

Nickelson, T. E., 1981. Coastal coho production factors. Technical report, Oregon Department of Fish and Wildlife Federal Aid Project AFS-74-2. Job final report.

Pacific States Marine Fisheries Commission, 2008. StreamNet: Fish data for the Northwest. Online at http://www.streamnet.org/.

Parken, C. K., R. E. Bailey, and J. R. Irvine. 2003. Incorporating uncertainty into area-under-the-curve and peak count salmon escapement estimation. North American Journal of Fisheries Management **23**:78–90.

Paulik, G. J., and D. S. Robson. 1969. Statistical calculations for change-in-ratio estimators of population parameters. The Journal of Wildlife Management **33**:1–27.

Perrin, C. J., and J. R. Irvine. 1990. A review of survey life estimates as they apply to the area-under-the-curve method for estimating the spawning escapement of Pacific salmon. Canadian Technical Report of Fisheries and Aquatic Sciences **1733**:49.

Pollock, K. H. 1982. A capture-recapture design robust to unequal probability of capture. Journal of Wildlife Management **46**:pp. 752–757.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs **107**:3–97.

Quinn, T. P. 2005. The Behavior and Ecology of Pacific Salmon and Trout. University of Washington Press, Seattle, Washington, 278 pp.

Rajwani, K. N., and C. J. Schwarz. 1997. Adjusting for missed tags in salmon escapement surveys. Canadian Jounal of Fisheries and Aquatic Sciences **54**:800–808.

Ransom, B. H., S. V. Johnston, and T. W. Steig. 1998. Review on monitoring adult salmonid (*Oncorhynchus* and *Salmo* spp.) escapement using fixed-location split-beam hydroacoustics. Fisheries Research **35**:33–42.

Rawding, D., Y.-W. Lee, P. Hahn, and Y. Cheng, 2008. Chinook stock assessment: tools for 21st Century management. Technical Report Project 07-1234-07, Washington Department of Fish and Wildlife Stock Assessment Unit, Olympia.

Reynolds, J. H., C. A. Woody, N. E. Gove, and L. F. Fair. 2007. Efficiently estimating salmon escapement uncertainty using systematically sampled data. American Fisheries Society Symposium: Sockeye salmon evolution, ecology, and management **54**:121–129.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin Fisheries Research Board Canada **191**:382.

Ripley, W. E. 1949. Tagging salmon with blowgun darts. Copeia **1949**:97–100.

Robson, D. S., and H. A. Regier. 1964. Sample size in Petersen mark-recapture experiments. Transactions of the American Fisheries Society **93**:215–26.

Rose, G., S. Gauthier, and G. Lawson. 2000. Acoustic surveys in the full monte: simulating uncertainty. Aquatic Living Resources **13**:367–372.

Rosenberger, J. L., and M. Gasko, 1983. Comparing location estimators: trimmed means, medians, and trimean. Pages 297–338 *in* D. C. Hoaglin, F. Mosteller, and J. W. Tukey, editors. Understanding Robust and Exploratory Data Analysis. John Wiley& Sons, New York, New York.

Schaefer, M. B. 1951. Estimation of size of animal populations by marking experiments. U.S. Fish and Wildlife Service Fishery Bulletin **52**:191–203.

Schroeder, S. L., 1973. Effects of density on the spawning success of chum salmon (*Oncorhynchus keta*) in an artificial spawning channel. Master's thesis, University of Washington, 78 pp.

Schubert, N. D. 1998. The 1994 Fraser River sockeye salmon (*Oncorhynchus nerka*) escapement. Canadian Technical Report of Fisheries and Aquatic Sciences **2201**:62.

Schubert, N. D. 2000. The 1994 Stellako River sockeye (*Oncorhynchus nerka*) escapement: evaluation of pooled Petersen and stratified mark-recapture estimates of a known population. Canadian Technical Report of Fisheries and Aquatic Sciences **2303**:56 pp.

Schwarz, C. J., and A. N. Arnason. 1996. A general methodology for the analysis of capture-recapture experiments in open populations. Biometrics **52**:860–873.

Schwarz, C. J., R. E. Bailey, J. R. Irvine, and F. C. Dalziel. 1993. Estimating salmon spawning escapement using capture-recapture methods. Canadian Jounal of Fisheries and Aquatic Sciences **50**:1181–1197.

Seber, G. A. F. 1965. A note on the multiple-recapture census. Biometricka **52**:249.

Seber, G. A. F. 1982. The Estimation of Animal Abundance and Related Parameters. Second edition. MacMillan, New York, NY.

Seiler, D., G. Volkhardt, L. Peterson, L. Fleischer, S. Neuhauser, P. Hanratty, and L. Kishimoto, 2004. 2003 juvenile salmonid production evaluation and adult escapement: intensively monitored watersheds (IMW) annual report. Technical report, Washington Department of Fish and Wildlife, Accessed online at http://www.wdfw.wa.gov/fish/wild_salmon_monitor

/publications/imw_salmon_monitoring_2003.pdf.

Shardlow, T. F. 2001. Using Underwater Still Photography to Estimate Salmon Escapement in Streams. North American Journal of Fisheries Management **21**:699–702.

Shardlow, T. F. 2004. Using time-lapsed video to estimate survey life for area-under-the-curve methods of escapement estimation using redd residence time for stream life. North American Journal of Fisheries Management **24**:1413–1420.

Shardlow, T. F., K. D. Hyatt, and S. R. Pearson. 2007. A new microcontrolled fish tag for accurately estimating life span and survey life of spawning salmon. Aquatic Living Resources **20**:197–203.

Shardlow, T. F., T. Webb, and D. Lightly. 1986. Chinook salmon escapement estimation on the Campbell and Quinsam Rivers in 1984: accuracy and precision of mark/recapture techniques using tagged salmon carcasses. Canadian Technical Report of Fisheries and Aquatic Sciences **1507**:52 p.

Simmonds, E. J. 1990. Very accurate calibration of a vertical echo sounder: a five-year assessment of performance and accuracy. Rapports et Proces-Verbaux des Reunions Conseil International pour l'exploration de la mer **189**:183–191.

Simpson, K. 1984. The accuracy of mark-recapture estimates of escapement. Canadian Technical Report of Fisheries and Aquatic Sciences **1326**:209–225.

Skalski, J. R., and J. J. Millspaugh. 2006. Application of multidimensional change-in-ratio methods using Program USER. Wildlife Society Bulletin **34**:433–439.

Skud, B. E. 1958. Relation of adult pink salmon size to time of migration and freshwater survival. Copeia **1958**:170–176.

Smith, J. J., M. R. Link, and B. D. Cain. 2005. Development of a long-term monitoring project to estimate abundance of Chinook salmon in the Copper River, Alaska, 2001-2004. Alaska Fishery Research Bulletin **11**:118–134.

Solazzi, M. F., S. L. Johnson, and L. A. VanDyke. 1983. Research and development of Oregon's anadromous fish stocks-coho. Oregon Department of Fish and Wildlife Fish Research Project **AFC-102-3**:Portland.

Sound Metrics Corp., 2009. Dual-frequency IDentification SONar (DIDSON). Online at http://www.soundmetrics.com/index.html.

Stables, T. B., and G. A. Kautsky. 2000. Evaluation of a stimulus-response method for distinguishing out-migrant salmonids from drifting debris for sonar counts in the Trinity River, California. Aquatic Living Resources **13**:341–347.

Su, Z. M., M. D. Adkison, and B. W. Van Alen. 2001. A hierarchical Bayesian model for estimating historical salmon escapement and escapement timing. Canadian Journal of Fisheries and Aquatic Sciences **58**:1648–1662.

Svoboda, P., and B. Harrington-Tweit. 1983. Estimating chum salmon escapements in the Nisqually River system from spawning ground survey data for the years 1975/76-1982/83. Nisqually Indian Tribe Technical Report **8**:19 pp.

Swan, G. A. 1989. Chinook salmon spawning surveys in deep waters of a large, regulated river. Regulated Rivers: Research and Management **4**:355–370.

Sykes, S. D., and L. W. Botsford. 1986. Chinook salmon, *Oncorhynchus tshawytscha*, spawning escapement based upon multiple mark-recapture of carcasses. Fisheries Bulletin **84**:261–270.

Thomason, G. J., and J. D. Jones. 1984. Southeastern Alaska pink salmon (*Oncorhynchus gorbuscha*) stream life studies, 1983. Alaska Department of Fish and Game Information Leaflet **No. 236**:15.

Underwood, T. J., J. A. Gordon, M. J. Millard, B. R. Lubinski, and S. P. Klosiewski. 2007. Appropriateness of the Darroch estimator for monitoring adult chum salmon on the middle Yukon River, Alaska. North American Journal of Fisheries Management **27**:303–314.

USACE. 2007. Columbia and Snake River Projects. United States Army Corps of Engineers Annual Fish Passage Report **2007**.

Visser, R., D. D. Dauble, and D. R. Geist. 2002. Use of Aerial Photography to Monitor Fall Chinook Salmon Spawning in the Columbia River. Transactions of the American Fisheries Society **131**:1173–1179.

Volkhardt, G., S. Neuhauser, P. Hanratty, L. Kishimoto, and L. Peterson, 2005. Appendix A: 2004 IMW juvenile salmonid production evaluation and adult escapement. Technical report, Washington Department of Fish and Wildlife, Accessed online at http://www.wdfw.wa.gov/fish/wild_salmon _monitor/publications/2004_imw_appendixa_(revised).pdf.

WDFW, 2008. Ballard Locks Sockeye Salmon Counts June-July 2008. Washington Department of Fish and Wildlife records posted online at http://wdfw.wa.gov/fish/sockeye/counts.htm.

Welsh, T. L., 1983. Redd Counting. Page 49 *in* W. S. Platts, W. F. Megahan, and G. W. Minshall, editors. Methods for evaluating stream, riparian, and biotic conditions. U. S. Intermountain Forest and Range Experiment Station, U.S. Department of Agriculture, Forest Service General Technical Report INT-138, Ogden, Utah.

Wolter, K. M. 1985. Introduction to Variance Estimation. Springer-Verlag, New York, 427 pp.

Xie, Y., T. J. Mulligan, G. M. W. Cronkite, and A. P. Gray. 2002. Assessment of potential bias in hydroacoustic estimation of Fraser River sockeye and pink salmon at Mission, B.C. Pacific Salmon Commission Technical Report **11**:42 pp.

Appendix A

TRAPEZOIDAL ESTIMATOR OF FISH-DAYS

This appendix shows the derivation of the alternate form of the trapezoidal estimator of fish-days (see Equation 2.9)

$$\begin{aligned}
\hat{F} &= \frac{1}{2}\sum_{i=1}^{r-1}(t_{i+1}-t_i)(\hat{L}_{i+1}+\hat{L}_i) \\
&= \frac{1}{2}[(t_2-t_1)(\hat{L}_2+\hat{L}_1)+(t_3-t_2)(\hat{L}_3+\hat{L}_2)+(t_4-t_3)(\hat{L}_4+\hat{L}_3)+... \\
&\quad ...+(t_{r-1}-t_{r-2})(\hat{L}_{r-1}+\hat{L}_{r-2})+(t_r-t_{r-1})(\hat{L}_r+\hat{L}_{r-1})] \\
&= \frac{1}{2}[t_2\hat{L}_2+t_2\hat{L}_1-t_1\hat{L}_2-t_1\hat{L}_1+t_3\hat{L}_3+t_3\hat{L}_2-t_2\hat{L}_3-t_2\hat{L}_2+... \\
&\quad ...+t_4\hat{L}_4+t_4\hat{L}_3-t_3\hat{L}_4-t_3\hat{L}_3+t_{r-1}\hat{L}_{r-1}+t_{r-1}\hat{L}_{r-2}... \\
&\quad ...-t_{r-2}\hat{L}_{r-1}-t_{r-1}\hat{L}_{r-2}+t_r\hat{L}_r+t_r\hat{L}_{r-1}-t_{r-1}\hat{L}_r-t_{r-1}\hat{L}_{r-1}] \\
&= \frac{1}{2}[t_2\hat{L}_1-t_1\hat{L}_2-t_1\hat{L}_1+t_3\hat{L}_2-t_2\hat{L}_3+t_4\hat{L}_3-t_3\hat{L}_4+... \\
&\quad ...+t_{r-1}\hat{L}_{r-2}-t_{r-2}\hat{L}_{r-1}-t_{r-1}\hat{L}_{r-2}+t_r\hat{L}_r+t_r\hat{L}_{r-1}-t_{r-1}\hat{L}_r] \\
&= \frac{1}{2}[(t_2-t_1)\hat{L}_1+(t_3-t_1)\hat{L}_2+(t_4-t_2)\hat{L}_3+(t_5-t_3)\hat{L}_4+... \\
&\quad ...+(t_{r-1}-t_{r-3})\hat{L}_{r-2}+(t_r-t_{r-2})\hat{L}_{r-1}+(t_r-t_{r-1})\hat{L}_r] \\
&= \frac{1}{2}(t_2-t_1)\hat{L}_1+\frac{1}{2}\sum_{i=2}^{r-1}(t_{i+1}-t_{i-1})\hat{L}_i+\frac{1}{2}(t_r-t_{r-1})\hat{L}_r \\
\hat{F} &= \frac{1}{2}\sum_{i=2}^{r-1}(t_{i+1}-t_{i-1})\hat{L}_i.
\end{aligned}$$

The final step holds if and only if $\hat{L}_1=0$ and $\hat{L}_r=0$.

Appendix B

TRAPEZOIDAL FISH-DAYS VARIANCE

The following is the derivation for the variance of the trapezoidal estimator of fish-days. Recall that the trapezoidal estimator is $\hat{F} = \frac{1}{2}\sum_{i=2}^{k-1}(t_{i+1} - t_{i-1})\hat{L}_i$, where $\hat{L}_{ihj} = \frac{l_{ihj}}{P_{ihj}}$. For simplicity, $(t_{i+1} - t_{i-1})$ is denoted τ_i.

The trapezoidal variance is presented in four sections in keeping with the four stages of the experiment that contribute variability: 1) estimating observer efficiency, 2) the binomial sighting process (where fish are detected with a common probability), 3) spatial subsampling, and 4) temporal subsampling. Taylor series expansions are carried out only to the first order term.

B.1 Uncertainty from estimating observer efficiency

This section draws upon the use of the Delta Method for the variance of the inverse of a random variable as estimated by the variance of the random variable divided by a variable raised to the fourth power:

$$V\left(\frac{1}{\hat{P}}\right) = V(\hat{P})\left(\frac{1}{\hat{P}^2}\right)^2 = \frac{V(\hat{P})}{P^4}. \tag{B.1}$$

The following property about random variable l will also be used:

$$E(l^2) = V(l) + E^2(l)$$

.

$$\begin{aligned}
E_4E_3E_2V_1(\hat{F}|2,3,4) &= E_4E_3E_2V_1(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{l_{ihj}}{\hat{P_{ihj}}}\right)|2,3,4) \\
&= E_4E_3E_2(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2V_1\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{l_{ihj}}{\hat{P_{ihj}}}\right)|2,3,4) \\
&= E_4E_3E_2(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})^2\sum_{j=1}^{n_h}l_{ihj}^2V_1\left(\frac{1}{\hat{P_{ihj}}}\right)|2,3,4) \\
&= E_4E_3E_2(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})^2\sum_{j=1}^{n_h}l_{ihj}^2\left(\frac{V(\hat{P}_{ihj})}{P_{ihj}^4}\right)|3,4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})^2\sum_{j=1}^{n_h}E_2\left(l_{ihj}^2\right)\frac{V(\hat{P}_{ihj})}{P_{ihj}^4}|3,4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})^2 \\
&\quad \textstyle\sum_{j=1}^{n_h}\left(V(l_{ihj})+E_2(l_{ihj})^2\right)\frac{V(\hat{P}_{ihj})}{P_{ihj}^4}|3,4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})^2 \\
&\quad \textstyle\sum_{j=1}^{n_h}\left(L_{ihj}P_{ihj}(1-P_{ihj})+L_{ihj}^2P_{ihj}^2\right)\frac{V(\hat{P}_{ihj})}{P_{ihj}^4}|4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})^2 \\
&\quad \textstyle\sum_{j=1}^{n_h}\frac{V(\hat{P}_{ihj})L_{ihj}[(1-P_{ihj})+L_{ihj}P_{ihj}]}{P_{ihj}^3}|4)
\end{aligned}$$

Let $x_{ihj}=\frac{V(\hat{P}_{ihj})L_{ihj}[(1-P_{ihj})+L_{ihj}P_{ihj}]}{P_{ihj}^3}$.

$$\begin{aligned}
E_4E_3E_2V_1(\hat{F}|2,3,4) &= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})^2\sum_{j=1}^{n_h}x_{ihj}|4) \\
&= E_4(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})E_3\left\{\left(\frac{N_h}{n_h}\right)\sum_{j=1}^{n_h}x_{ihj}\right\}) \\
&= E_4(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})\sum_{j=1}^{N_h}x_{ihj}) \\
&= \tfrac{k}{4K}\textstyle\sum_{i=2}^{K-1}\tau_i^2\sum_{h=1}^{H}(\frac{N_h}{n_h})\sum_{j=1}^{N_h}x_{ihj} \\
&= \tfrac{k}{4K}\textstyle\sum_{i=2}^{K-1}(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}(\frac{N_h}{n_h}) \\
&\quad \textstyle\sum_{j=1}^{N_h}\frac{V(\hat{P}_{ihj})L_{ihj}[(1-P_{ihj})+L_{ihj}P_{ihj}]}{P_{ihj}^3}
\end{aligned}$$

B.2 Uncertainty from the binomial sighting process

$$\begin{aligned}
E_4E_3V_2E_1(\hat{F}|2,3,4) &= E_4E_3V_2E_1(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{l_{ihj}}{\hat{P_{ihj}}}\right)|2,3,4) \\
&= E_4E_3V_2(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{l_{ihj}}{P_{ihj}}\right)|3,4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2V_2\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{l_{ihj}}{P_{ihj}}\right)|3,4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}\left(\frac{N_h}{n_h}\right)^2\sum_{j=1}^{n_h}\frac{1}{P_{ihj}^2}V_2(l_{ihj})|3,4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}\left(\frac{N_h}{n_h}\right)^2\sum_{j=1}^{n_h}\frac{1}{P_{ihj}^2}(L_{ihj}P_{ihj}(1-P_{ihj}))|4) \\
&= E_4E_3(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}\left(\frac{N_h}{n_h}\right)^2\sum_{j=1}^{n_h}\frac{L_{ihj}(1-P_{ihj})}{P_{ihj}}|4) \\
&= E_4(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}\frac{N_h}{n_h}E_3\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{L_{ihj}(1-P_{ihj})}{P_{ihj}}\right)|4) \\
&= E_4(\tfrac{1}{4}\textstyle\sum_{i=2}^{k-1}\tau_i^2\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{N_h}\frac{L_{ihj}(1-P_{ihj})}{P_{ihj}}) \\
&= \tfrac{k}{4K}\textstyle\sum_{i=2}^{K-1}\tau_i^2\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{N_h}\frac{L_{ihj}(1-P_{ihj})}{P_{ihj}} \\
&= \tfrac{k}{4K}\textstyle\sum_{i=2}^{K-1}(t_{i+1}-t_{i-1})^2\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{N_h}\frac{L_{ihj}(1-P_{ihj})}{P_{ihj}}
\end{aligned}$$

B.3 Uncertainty from spatial subsampling

$$\begin{aligned}
E_4V_3E_2E_1(\hat{F}|2,3,4) &= E_4V_3E_2E_1(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{l_{ihj}}{\hat{P_{ihj}}}\right)|2,3,4) \\
&= E_4V_3E_2(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\left(\frac{l_{ihj}}{P_{ihj}}\right)|3,4) \\
&= E_4V_3(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}L_{ihj}|4) \\
&= E_4(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i^2V_3\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}L_{ihj}\right)|4) \\
&= E_4(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}(t_{i+1}-t_{i-1})^2V_3\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}L_{ihj}\right))
\end{aligned}$$

The form of the variance contributed by spatial sub-sampling depends on the sampling pattern used to select surveyed stream sections (reaches).

B.4 Uncertainty from temporal subsampling

$$\begin{aligned}
V_4E_3E_2E_1(\hat{F}|2,3,4) &= V_4E_3E_2E_1(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\frac{l_{ihj}}{\hat{P_{ihj}}}\right)|2,3,4)\\
&= V_4E_3E_2(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}\left(\frac{l_{ihj}}{P_{ihj}}\right)|3,4)\\
&= V_4E_3(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}L_{ihj}|4)\\
&= V_4(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_iE_3\left(\sum_{h=1}^{H}\frac{N_h}{n_h}\sum_{j=1}^{n_h}L_{ihj}\right)|4)\\
&= V_4(\tfrac{1}{2}\textstyle\sum_{i=2}^{k-1}\tau_i\left(\sum_{h=1}^{H}\sum_{j=1}^{N_h}L_{ihj}\right))\\
&= \tfrac{1}{4}V_4(\textstyle\sum_{i=2}^{k-1}(t_{i+1}-t_{i-1})\left(\sum_{h=1}^{H}\sum_{j=1}^{N_h}L_{ihj}\right))
\end{aligned}$$

The form of the variance due to temporal sampling depends upon the sampling scheme used to select surveyed days.