EVALUATION OF VISUAL SURVEY PROGRAMS FOR MONITORING COHO SALMON ESCAPEMENT IN RELATION TO CONSERVATION GUIDELINES

by

Kendra Holt
B.Sc., University of Victoria, 2002

RESEARCH PROJECT
SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF RESOURCE MANAGEMENT

In the
School of Resource and Environmental Management

Project No. 410

SIMON FRASER UNIVERSITY

Fall 2006

APPROVAL

Name: **Kendra Holt**

Degree: **Master of Resource Management**

Title of Research Project: **Evaluation of Visual Survey Programs for Monitoring Coho Salmon Escapement in Relation to Conservation Guidelines**

Project Number: **410**

Supervisory Committee:

Chair: **Merran Hague**

____

Dr. Sean P. Cox
Senior Supervisor
Assistant Professor, School of Resource and Environmental Management

____

Dr. Randall M. Peterman
Supervisor
Professor, School of Resource and Environmental Management

Date Defended/Approved: October 4, 2006

ABSTRACT

Canada's Wild Salmon Policy (WSP) requires that quantitative survey designs be used to monitor annual trends in Pacific salmon escapement. Visual survey methods, in which periodic counts of spawning fish are made throughout a season, are often employed for this purpose. Coho salmon populations are difficult to monitor using visual survey methods due low probability of fish detection and high variability in the annual timing of fish presence in the survey area. I developed a Monte Carlo simulation procedure to evaluate the power of peak-count, mean-count, trapezoidal area-under-the-curve (AUC), and likelihood AUC methods to detect 30% declines in coho salmon escapement over 10 years, which is the magnitude of trend that would warrant listing a population as threatened under the Canadian Species at Risk Act (SARA). My results suggest that a simple mean-count method would be best suited for monitoring coho salmon abundance in relation to SARA and WSP guidelines.

Keywords

Oncorhynchus kisutch; coho salmon; escapement monitoring; visual surveys; survey design; simulation modelling; trend detection; power analysis; Wild Salmon Policy

ACKNOWLEDGEMENTS

I thank my senior supervisor, Sean Cox for his guidance, energy, and expertise, as well as my second committee member, Randall Peterman, for his thoughtful input and reviews. I also thank Dave Peacock and Joel Sawada from Fisheries and Oceans Canada, and the members of the Fisheries Science and Management Research Group at Simon Fraser University, for continually providing valuable comments that improved this research project. In particular, fellow students Merran Hague and Carrie Holt provided useful feedback on earlier drafts. For assistance in providing data and insight into visual survey methods, I thank Richard Bailey, Mike Chamberlain, Jim Irvine, Ian Matthews, Chuck Parken, Kent Simpson, and Joe Tadey from Fisheries and Oceans Canada. Finally, I thank my fellow students (and friends) in REM for the academic support they have provided while I've been at SFU, as well as all the great times away from SFU. Financial support was provided by Fisheries and Oceans Canada (via Joel Sawada and Dave Peacock).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

WSP	Wild Salmon Policy
CU	Conservation Unit (as defined in WSP)
DFO	Fisheries and Oceans Canada
SARA	Species at Risk Act

CHAPTER 1
INTRODUCTION TO VISUAL SURVEY METHODS

Under Canada's new Wild Salmon Policy (WSP), genetically and geographically similar spawning aggregations of Pacific salmon (*Oncorhynchus spp.*) will be grouped into conservation units (CUs), and for each CU, a monitoring plan will be developed to assess interannual trends in both the abundance and distribution of fish within the CU (DFO 2005). These monitoring plans will include documentation of the survey methods that will be used to monitor salmon escapement (the number of salmon returning to a stream to spawn), as well as the identification of two benchmarks (upper and lower) against which the status of the CU can be assessed. The annual status of a CU will be used to inform management decisions regarding harvest levels, enhancement activities, and habitat alterations.

Monitoring plans will vary among CUs; however, the WSP outlines four broad requirements: (i) cost-effectiveness, (ii) ability to build on existing stock assessment programs and partnerships with First Nations and enhancement groups, (iii) utilization of statistically-based survey methods that accurately assess interannual trends in abundance with a high level of confidence, and (iv) all data collected should be relevant to the provision of management advice. Visual survey methods, in which field personnel make periodic counts of spawner abundance throughout a spawning season, will likely play an important role in monitoring plans developed under the WSP. While more intensive survey methods such as enumeration fences and mark-recapture studies will be used to monitor a selected number of indicator streams within each CU, visual surveys provide a cost effective means of assessing the consistency of escapement trends across an entire CU. Furthermore, the large number of streams monitored using visual survey methods allow for an easy division of survey effort among Fisheries and Oceans Canada (DFO), First Nations communities, and local enhancement groups. The suitability of visual survey methods with regards to the first two requirements for WSP monitoring plans is thus well established. For my 699 research project, I evaluate the suitability of visual survey programs for monitoring coho salmon (*O. kisutch*) escapements with regard to the third requirement of WSP monitoring programs: that survey methods accurately assess interannual trends in abundance with a high degree of confidence.

In general, the ability of visual survey programs to detect trends in escapement depends on consistency in both the ratio of estimated abundance to true abundance (hereafter referred to as "observer efficiency") and the timing of fish presence in the survey area among years (hereafter referred to as "run timing"). Visual surveys for coho salmon are known to be especially problematic due to high inconsistency in both of these factors. The camouflage colouration of coho salmon on the spawning grounds makes them difficult to detect, and their run timing dynamics can be highly variable between years. The shape of the curve describing coho salmon run timing dynamics frequently deviates from a normal distribution to a highly skewed or bimodal distribution (Fraser et al. 1983, Holtby et al. 1984).

I present my research in two chapters. In Chapter 2, I review existing literature on alternative methods used to derive escapement estimates or indices from visual survey data, and highlight the key sources of uncertainty for each method. In Chapter 3, I present a Monte Carlo simulation study that evaluates the ability of four alternative visual survey methods (peak-count, mean-count, trapezoidal area-under-the-curve (AUC), and maximum likelihood AUC with a beta-distributed run timing model) to detect trends in coho salmon escapement to a single stream given realistic levels of variability in run timing dynamics and observer efficiency. I use a range of scenarios about "true" population dynamics and survey designs to show that a simple mean-count method could be applied broadly to monitor coho salmon escapement in relation to rate-based benchmarks of population change, such as percent change in escapement over time. The mean-count method uses the mean number of salmon counted over all surveys for a given stream within a single year to index escapement.

CHAPTER 2
LITERATURE REVIEW

Introduction

To ensure both short- and long-term benefits from Pacific salmon fisheries, escapement goals are set for individual stocks and spawning populations are monitored annually in relation to these goals. Accurate and precise estimates of salmon escapement are regarded as an essential component of all salmon management plans. Escapement estimates are considered accurate when the estimated value is close to the true or accepted value, and precise when further sampling and calculations produce the same or similar result. Depending on the salmon population, escapement estimates can be used to define stock recruitment relationships, forecast recruitment of future generations, make in-season management decisions, and set optimal harvest rates.

Because time and budget constraints make it impossible to monitor all salmon populations on the Pacific Coast, populations are grouped into stock complexes and escapement monitoring is conducted on only a subset of populations. These monitored populations are used as indicators of trends experienced over the entire stock complex and are used to make annual harvest decisions. A wide range of methods are available for monitoring indicator populations, ranging from intensive mark-recapture programs and enumeration fences to less intensive visual survey programs (Cousens et al. 1982).

Visual survey methods have been commonly used to monitor Pacific salmon since the 1950's (e.g. Bevan 1961, Cousens et al. 1982, Bue et al. 1998). Sampling protocols typically involve field observers travelling a predetermined length of stream multiple times throughout a spawning season and counting the number of fish they observe on each visit. Various modes of travel are used, including aerial overflights, snorkelling, rafting, and walking. The relatively low cost and short time commitment required to complete each visual survey are significant advantages to monitoring programs operating on limited resources. Monitoring programs based on visual surveys are frequently referred to as extensive survey programs because they allow for a wide range of spawning populations to be monitored with less certainty than more effort-intensive methods. As a result, visual survey escapement estimates are often treated as relative indices, as opposed to absolute estimates of escapement. Estimation methods have been developed to produce absolute estimates of escapement using visual survey data (Ames and Phinney 1977, English et al. 1992, Hilborn et al. 1999); however, as I will show, the level of effort required to produce accurate and precise estimates of absolute escapement can be higher than necessary for accurate trend detection. Confidence in visual survey estimates generally decreases as the frequency of counts decreases (Hill 1997, Bue et al. 1998, Korman et al. 2002), and thus, extensive survey programs must find a balance between program cost and the level of certainty in escapement estimates required for management decisions.

Several estimation methods are available for deriving either a relative escapement index or an absolute escapement estimate from visual survey counts. In this literature review, I evaluate the utility of three commonly used estimation methods (peak-live-count, trapezoidal area-under-the-curve (AUC), and maximum likelihood AUC) for monitoring coho salmon escapement as part of an extensive survey program aimed at detecting long-term escapement trends. I chose to examine the peak-live-count and trapezoidal AUC methods because they are commonly employed by salmon management agencies along the Pacific coast. The maximum likelihood AUC method is of interest because it is relatively new (Hilborn et al. 1999) and allows uncertainty in count data, survey life, and observer efficiency to be taken into account when estimating escapement. Detailed reviews of other methods that I do not cover are provided in Cousens et al. 1982.

An important difference among the three estimation methods I evaluate is the level of escapement information provided by estimates. The peak-count method provides a relative estimate of escapement, while the two AUC methods attempt to provide absolute estimates of escapement. Absolute estimates of escapement are necessary for some monitoring purposes, such as determining the probability of escapement dropping below a minimum acceptable population size. However, when the primary purpose of escapement monitoring is to detect long-term trends in escapement, relative estimates of escapement are sufficient, as long as biases in escapement estimates remain constant over time. In this literature review, as well as in Chapter 3, I have chosen to focus on the latter of these two monitoring purposes: the detection of time trends in escapement.

In the first section of this literature review, I discuss key sources of uncertainty that are common to all visual survey monitoring methods. In the second section, I present a more detailed review of the peak-live-count, trapezoidal AUC, and maximum likelihood AUC methods. For each method, I highlight the advantages and disadvantages of its application to escapement monitoring, as well as how previous studies have assessed its performance relative to other methods. Finally, in my concluding section, I summarize my findings, identify areas for further research, and propose a fourth alternative for indexing escapement, a simple mean-count method that may reduce some of the uncertainty in estimated escapement trends.

Uncertainty in Visual Surveys

Sources of uncertainty in visual survey escapement estimates can be categorized as observation errors arising from measurement (i.e., counts) and process errors arising from natural variability in population dynamics (Hilborn and Mangel 1997). While the key sources of uncertainty in escapement estimates vary with estimation method, all methods require that assumptions be made about the ability of observers to see fish and the timing and shape of the curve describing fish presence in the survey area (hereafter referred to collectively as "run timing dynamics"). These two sources of uncertainty in escapement estimates represent observation and process errors, respectively. I will show that both of these factors are highly variable and difficult to predict. When possible, I refer to studies that have focused on coho salmon specifically, although when necessary, I draw from other species of Pacific salmon.

Observer Efficiency

Observer efficiency (i.e., the proportion of fish seen by observers) varies among streams, among years within streams, and among individual surveys conducted in a single stream in a single year. Factors influencing observer efficiency include observer experience, weather conditions, fish behaviour, life history (adult or jack), survey method, and physical stream characteristics such as turbidity, water level, and habitat type (Bevan 1961, Shardlow et al. 1987, Jones et al. 1998, Korman et al. 2002). As an additional complication, the accuracy of counts has been shown to decrease non-linearly with increasing fish density (Jones et al. 1998). The magnitude of variability in observer efficiency is difficult to quantify due to observation error in estimates of the "true" number of fish present on a given day; however, it is likely quite high. For coho salmon foot surveys in Clear Creek, Alaska estimates of observer efficiency ranged from 25% to 100% within a single spawning season (Hetrick and Nemeth 2003). Estimates of between-observer coefficients of variation for counts of pink salmon (*O. gorbuscha*) collected by foot surveys in Chaik Bay Creek, Alaska ranged from 30% to 49% (Jones et al. 1998). Predictive relationships describing observer efficiency as a function of physical stream conditions (e.g., river discharge, turbidity, visibility; Korman et al. 2002, Hetrick and Nemeth 2003), or individual observers (Jones et al. 1998) have been developed for some salmon populations.

Run Timing Dynamics

The run timing of salmon into spawning streams is influenced by both heritable genetic traits (Hansen and Jonsonn 1991, Smoker et al. 1998, Stewart et al. 2002) and localized environmental factors (Fukushima and Smoker 1997, Hodgson and Quinn 2002, Keefer et al. 2004), and therefore, can also vary both between streams and between years. Variability in run timing dynamics can be broken down into two components: (i) variability in the timing of fish arrival into survey areas and (ii) variability in the length of time fish remain alive in the survey area (referred to as "survey life"). When the entire spawning population enters a stream within a few days, the shape of the run timing curve describing fish abundance as a function of date will be narrower than when fish arrival is spread out over several days or weeks. Similarly, for years in which the mean survey life is shorter, the shape of the run timing curve will be narrower.

Interannual variability in environmental factors has been shown to influence the timing of arrival of salmon into spawning streams. Discharge levels in the Columbia River explain much of the among-year variation in migration timing of spring-summer chinook (*O. tshawytscha*) salmon (Keefer et al. 2004). Hodgson and Quinn (2002) showed that among 129 summer sockeye (O. *nerka*) salmon populations ranging from Washington to Alaska, peak in-river temperature, as well as migration distance and latitude, contributed significantly to spawning date. In years with peak temperatures exceeding an upper threshold of 19°C, sockeye tended to enter freshwater either prior to or after the warmest period.

Several anecdotal accounts suggest that the influence of environmental factors on arrival timing dynamics may be particularly strong for coho salmon, which spawn later in the season than other species and frequently display pulsed and bimodal run timing distributions. Coho often enter spawning streams during periods of increased water flow (Neave 1943, Fraser et al. 1983, Holtby et al. 1984), and the tendency of coho to migrate past enumeration fences during rainfall events is well known. Holtby *et al.* (1984) observed coho milling around the mouth of a coastal spawning stream, presumably waiting for stream conditions to become suitable before entry. He found that in years in which flow was constant, coho were constantly entering the stream, whereas in years with infrequent freshets, entry timing was pulsed.

The second component of run timing dynamics, survey life, is alternatively referred to as "stream life" and "redd residence time" within the literature. Stream life is used when the survey area covers the entire stream, whereas redd residence time is used when counts are restricted to spawning fish associated with redd nest sites. I have chosen to use the generic term survey life, which encompasses both of these metrics. As with observer efficiency and arrival timing, average survey life can vary both among years and among streams within a given year (Perrin and Irvine 1990, English et al. 1992, Irvine et al. 1992). For example, Irvine et al. (1992) found that for two Vancouver Island, BC streams located within 80 km of each other (Black Creek and French Creek), the estimated mean annual redd residence times for coho salmon in 1988 were 8.0 and 16.7 days, respectively, compared with 15.0 and 15.5 days in 1989. Over a period of three years, mean residence times in Black Creek ranged from 8.0 to 16.6 days, while those in French Creek ranged from 13.3 to 16.7 days.

Variation among individual fish further complicates estimation of survey life. For example, early arriving fish tend to have longer redd residence times than fish that arrive later in the spawning season (Neilson and Geen 1981, Neilson and Banford 1983, Perrin and Irvine 1990). Van den Berghe and Gross (1986) found that life history, body size, and adult spawning density all contributed significantly to explaining variation in coho salmon stream life. Stream discharge and stream temperature have also been shown to influence stream life of pink salmon in the Auke Lake system, Alaska (Fukushima and Smoker 1997).

In summary, there are several sources of uncertainty arising from the measurement of fish abundance in spawning streams, as well as natural variability in run timing dynamics, that contribute to errors in visual survey escapement estimates. The escapement estimation methods I review in the next section vary in the extent to which they take these uncertainties into account when analysing visual survey data, producing escapement estimates with variable levels of precision and accuracy.

Escapement Estimation Methods from Visual Survey Data

Peak-Live-Count (Peak-count)

In the peak-live-count method (hereafter referred to as simply "peak-count"), the highest count value observed over all surveys conducted in a year is used to index escapement (Bevan 1961). The peak-count value is expected to be biased low because only a portion of the total escapement is in the stream and available for observation on a given day (Bevan 1961). In order to use peak-count indices for annual management of salmon stocks, an absolute estimate of escapement is often obtained by multiplying the peak-count value by an expansion factor that accounts for fish not present in the stream at the time of counting and fish that were present but not observed (Jones et al. 1998, Parken et al. 2003). Alternatively, peak-count values themselves have been used as escapement targets for a single stream or composite of streams (Shaul et al. 2003 for coho salmon).

The accuracy and precision of escapement indices developed using the peak-count method are dependent on annual consistency in the ratio of observed peak-count to total spawner escapement. There are thus two major sources of uncertainty in escapement indices developed using the peak-count method: (i) interannual variability in observer efficiency, and (ii) interannual variability in run timing dynamics. As already shown, both of these factors can be highly variable between years.

The advantage of the peak-count method over more complex AUC methods is that it can be used with as little as one survey count per year, although more commonly, two or three surveys are conducted near the historic date of peak spawning abundance (Jones et al. 1998, Parken et al. 2003). A significant disadvantage of the peak-count method is the requirement to schedule surveys to overlap with the date of peak spawning abundance, which can be difficult to predict due to high variability in the timing of peak abundance, both between years for a specific stream and between streams within a given year. For example, over a period of three years, the dates on which the peak numbers of coho salmon were seen in Black and French Creeks on Vancouver Island, British Columbia had ranges of 26 and 16 days, respectively (Irvine et al. 1992). In the Rakaia River system in New Zealand, the date of peak abundance for seven different chinook spawning aggregations located within the system extended over a range of 22 days within a single year (West and Goode, 1987).

Trapezoidal AUC (T-AUC)

The trapezoidal AUC method (hereafter referred to as "T-AUC") differs from the peak-count method in that information from all surveys conducted in a year are used in escapement estimation, not just the survey that produces the highest count. In the T-AUC method, an absolute estimate of escapement is obtained by (i) calculating the area under the observed run timing curve (also referred to as the escapement curve) using a trapezoidal approximation to the shape of the distribution, (ii) dividing the area-under-the-curve by an estimate of survey life, and (iii) scaling the product from (ii) by an estimate of observer efficiency (see Chapter 3 for equations). T-AUC escapement estimates have three major sources of uncertainty, one related to each of the above three components. The first is associated with the ability of survey counts to accurately represent the shape of the run timing curve, while the second and third are the estimates of survey life and observer efficiency. The application of this method requires the analyst to assume that survey life and observer efficiency parameters are known without error.

Ames and Phinney (1977) were the first to document the application of the T-AUC method to salmon escapement estimation, and it has been commonly used since then by salmon management agencies along the Pacific Coast (Bue et al. 1998). Most studies aimed at assessing the performance of the T-AUC method have been empirical, in which an T-AUC estimate of escapement for a given year is compared with that from a more reliable method such as mark-recapture or fence counts (although see Hill 1997, Manske and Schwarz 2000, and Korman et al. 2002 for simulation approaches). The T-AUC method has been shown to produce accurate escapement estimates when compared to fence counts and mark-recapture escapement estimates; however, the accuracy of T-AUC estimates is highly dependent on the use of year- and stream-specific estimates of survey life and observer efficiency (English et al. 1992, Bue et al. 1998, Parken et al. 2003). The extrapolation of more general survey life or observer efficiency estimates among years or streams can lead to positively or negatively biased escapement estimates (Bue et al. 1998, Parken et al. 2003). The precision and accuracy of T-AUC estimates generally improves with increasing survey frequency because a larger portion of the run timing curve is observed (Hill 1997, Bue et al. 1998).

An obvious advantage of the T-AUC method over the peak-count method is that it provides an estimate of absolute escapement. The disadvantages include greater survey effort (if less than three surveys per year are conducted, the analyst must make educated guesses to fill in missing count values) and the need for accurate year- and stream-specific estimates of survey life and observer efficiency (English et al. 1992, Irvine et al.

1992, Bue et al. 1998). The sampling methods required to estimate survey life and observer efficiency are usually intensive, requiring large financial and time commitments. Methods for estimating survey life include tagging programs and enumeration fences (English et al. 1992, Bue et al. 1998), capture-recapture studies (Manske and Schwarz 2000, Parken et al. 2003) and daily observations (Van den Berghe and Gross 1986, Trouton et al. 2004). The use of time-lapse-video technology has recently been shown to provide inexpensive estimates of redd residence time for sockeye salmon (Shardlow 2004); however, this method may not be suitable for small populations with more dispersed redd sites such as coho salmon. Methods for estimating observer efficiency include electrofishing, tagging programs, enumeration fences, and radio-telemetry (Irvine et al. 1992, Bue et al. 1998, Jones et al. 1998, Korman et al. 2002). English et al. (1992) and Irvine et al. (1992) both suggest that the requirement for year- and stream-specific estimates of survey life and observer efficiency should limit the application of the T-AUC method to high priority streams where absolute estimates of escapement are required.

Maximum Likelihood AUC (L-AUC)

The maximum likelihood approach to AUC escapement estimation (hereafter referred to as "L-AUC") was developed in response to concerns that the T-AUC method did not allow for the calculation of uncertainty in escapement estimates (Hilborn et al. 1999; but see Parken et al. 2003 for an example using bootstrap sampling with the T-AUC method). The L-AUC method involves (i) a run timing model to predict daily fish abundance in the survey area, (ii) an observation model to predict count values based on predicted abundance and observer efficiency, and (iii) a statistical model with a goodness of fit criterion to estimate total escapement using the predicted and observed count values. The run timing model requires the specification of a cumulative distribution to model fish arrival and death, as well as a survey life parameter that specifies the offset (in days) between the location of the two distributions (see Chapter 3 for equations).

The L-AUC method of escapement estimation is based on a framework for describing run timing dynamics initially developed by Quinn and Gates (1997). Hilborn et al. (1999) expanded on the work of Quinn and Gates (1997) by using cumulative normal and beta distributions to describe fish arrival, developing methods for incorporating uncertainty in survey life and observer efficiency parameter values in escapement estimation, and showing how the likelihood method could be used to estimate confidence bounds on escapement estimates. In addition, Hilborn et al. (1999) used pink salmon visual survey data from 18 streams with counting fences to compare the performance of the L-AUC method with that of the T-AUC method. Their results showed that when stream- and year-specific estimates of survey life and observer efficiency were assumed to be known without error and were applied to both methods, the T-AUC method produced slightly less biased escapement estimates than the L-AUC method. However, their results also showed that when stream life and observer efficiency were treated as uncertain parameters, the level of uncertainty in escapement estimates increased substantially compared to when they were treated as known using stream- and year-specific estimates, particularly in the case of observer efficiency. The authors concluded that when an estimate of uncertainty is required, the likelihood method should be applied with year-specific estimates of survey life and observer efficiency. With the exception of bootstrap sampling of replicate survey data (Parken et al. 2003), which can be highly effort-intensive, there have been no alternative methods developed for estimating uncertainty in visual survey escapement indices.

Su et al. (2001) and Adkison and Su (2001) expanded on the methodology of Hilborn et al. (1999) by incorporating a decline in survey life over the course of the spawning season into the run timing model, and using a hierarchical Bayesian approach that allowed the estimation procedure to “borrow information” from years with ample survey data to apply to years with sparse survey data or no post-peak counts. While the application of a hierarchical Bayesian approach reduces data requirements, this approach is more computationally intensive than other escapement monitoring options and, in the form used by Su et al. (2001), is only suitable for stocks with high annual consistency in run timing dynamics, which does not appear to be the case for coho salmon. Korman et al. (2002) further expanded on the above studies by exploring alternative run timing models, including the beta distribution and a pulsed arrival model, demonstrating the high sensitivity of escapement estimates to assumptions about run timing dynamics, and developing a likelihood model that incorporated mark-recapture data.

Two major uncertainties in the L-AUC method are the run timing model assumed (e.g. beta, normal, or pulsed) and the error structure assumed when fitting the run timing model to observed visual survey data. Uncertainty arising from the selection of a run timing model would likely be particularly important in cases where the assumed model is unimodal, but the true run timing curve is bimodal. In addition, L-AUC escapement estimates include uncertainties associated with the treatment of survey life and observer efficiency.

There are significant advantages to employing the L-AUC method. It allows stock assessment scientists to produce annual estimates of escapement with measures of uncertainty, and it can produce escapement estimates for years with sparse count data when used in a hierarchical framework. Disadvantages of the method include high computational requirements, especially in the case of the hierarchical Bayesian model, the required assumption of a statistical distribution to describe run timing dynamics, and the large number of survey counts required to estimate several model parameters. For example, the application of a normally distributed run timing model with uncertainty in survey life and observer efficiency parameters requires a minimum of six surveys within a year to ensure that sample size is greater than the number of parameters being estimated. Hilborn et al. (1999) recommended the development of run timing models using mixture distributions to describe arrival and death to deal with populations displaying highly skewed and bimodal run timing curves. However, these models would further increase the number of parameters to be estimated, and thus, the minimum number of surveys required.

Conclusions

There are two areas where further research is needed with regard to extensive survey programs for coho salmon. The first is the need to develop of an alternative estimation method that can accurately and precisely detect temporal trends in relative escapement with minimal survey effort and the second is a quantitative evaluation of the ability of several visual estimation methods to detect trends in escapement. For the first issue, existing methods of estimating escapement have several drawbacks. The peak-count method is not able to deal with high interannual variability in the timing and shape of run timing curves, which is commonly observed for coho salmon populations. While the AUC methods may be useful for intensive monitoring programs in which stream- and year-specific values of survey life and observer-efficiency are estimated, they are likely not suitable for extensive survey programs that aim to monitor escapement trends across broad geographic ranges with minimum levels of effort. In the absence of stream- and year-specific estimates of survey life and observer efficiency, which would be expected for extensive survey programs, the extrapolation of parameter values between years and streams has the potential to introduce an additional source of uncertainty into estimated escapement trends.

Several factors should be considered when selecting an estimation method for application to extensive visual survey programs. First and foremost, the level of detail extracted from visual survey programs should match the level of detail required for management. For extensive survey programs aimed primarily at detecting escapement trends, absolute estimates of escapement are not necessary, unless the population is at an extremely low abundance. Second, given the high levels of uncertainty in survey life and observer efficiency, and the high costs of estimating these parameters, a simple data-based method that requires no assumptions about these values is desirable. Finally, the application of an estimation method that utilizes information from all surveys conducted in a year on a given stream would be expected to reduce the level of uncertainty in trend detection caused by interannual variability in run timing.

One potential estimation method that has rarely been used, yet may meet several of the above criteria, is a simple mean-count method that uses the mean survey count observed for a given stream in a given year to index escapement. The application of an adjusted mean-count method, in which the average survey count was multiplied by the ratio of total days of fish presence in a survey area to survey life, has been previously documented (Gangmark and Fulton 1952 *in* Cousens et al. 1982); however, it is not commonly used. To the best of my knowledge, there have been no published studies that utilize a simple mean-count method.

The second area for further research is to quantify the ability of visual survey monitoring programs to detect long-term trends in escapement. Evaluations of visual survey methods have tended to focus on the utility of these methods for estimating absolute escapement within a single stream in a given year. These studies have been primarily empirical, in which the visual survey escapement estimate for a given stream is compared with that obtained from a more reliable enumeration method.

In Chapter 3, I use simulation modelling to address both areas for future research on extensive monitoring programs that I identified in the above literature review. I develop and implement a Monte Carlo simulation procedure to evaluate the ability of three commonly used visual survey estimation methods (peak-count, T-AUC, and L-AUC), as well as a mean-count method, to detect long-term escapement trends given realistic levels of process variation in coho run timing dynamics and observation error in count data and survey life estimates.

CHAPTER 3
EVALUATION OF VISUAL SURVEY PROGRAMS FOR MONITORING COHO SALMON ESCAPEMENTS: A SIMULATION MODELLING APPROACH

Introduction

As shown in Chapter 1, the ability of visual survey estimation methods to detect escapement trends depends upon annual consistency in (i) observer efficiency (Bevan 1961, Shardlow et al. 1987, Bue et al. 1998), and (ii) run timing, which is itself dependent on arrival timing into the survey area and survey life (English et al. 1992, Irvine et al. 1992, Bue et al. 1992). Visual surveys for coho salmon are known to be especially problematic due to high inconsistency in the above factors. In this Chapter, I present a Monte Carlo simulation procedure to evaluate the power of four estimation methods (peak-count, mean-count, trapezoidal AUC [T-AUC], and maximum likelihood AUC with a beta-distributed run timing model [L-AUC]) to detect long-term declines in coho salmon escapement given realistic levels of variation in observer efficiency and run timing. The specific objectives of this study are three-fold: (i) to compare the power of alternative estimation methods to detect declines in coho salmon escapement within a single stream that would warrant listing a population as "threatened" under the Canadian Species at Risk Act (SARA), (ii) to examine the effect of survey frequency on power, and (iii) to test the sensitivity of my results to a wide range of scenarios about survey designs and true population dynamics. A simulation modelling approach is well suited to these objectives because it allows for the evaluation of several escapement monitoring methods using a generated data set for which the true escapement is known. By creating a declining time series of escapement data, which represents the true state of nature, I can compare the predicted escapement time series for all four estimation methods with the true escapement time series. An additional advantage of a simulation modelling approach is the ease of incorporating hypothetical components into the analysis, such as the frequency and timing of visual surveys throughout a season, the rate at which the population is declining, and the level of interannual variability in run timing dynamics.

To measure the performance of trend detection, I chose power, which is defined as the probability of correctly rejecting the null hypothesis (in this case, that the rate of decline in escapement is less than or equal to the critical rate of decline specified in SARA). One minus power represents the probability (β) of the worst-case scenario occurring from a conservation point-of-view; that is, making a Type II error where the population is threatened due to a decrease in abundance, but we fail to realize it and take the appropriate conservation actions (Peterman 1990).

The results of this study will provide stock assessment scientists and fisheries managers with insight into the utility of alternative estimation methods for monitoring coho salmon escapement within a single stream. I provide specific recommendations for the design of visual survey monitoring programs, and discuss suitable performance measures that could be used to assess stock status in relation to escapement benchmarks that are yet to be developed under the WSP.

Methods

Methods are described in two sections: (i) model parameterization through analyses of existing data sets, and (ii) simulation of alternative monitoring options over a 10-year period and evaluation of performance of different monitoring methods.

Model Parameterization

Data Sources

Thirty-three existing coho salmon visual foot survey data sets were used to select an appropriate run timing model for simulating visual survey data, and to estimate interannual variability in run timing parameter values. Data sets were collected from 11 streams over a period of three years. Nine of the streams were tributaries to the North Thompson River in the Fraser River system, British Columbia, and the other two were located on the east coast of Vancouver Island, British Columbia. All data sets had a minimum of five surveys per year. Data from a subset of streams are presented in Figure 1, while Appendix A contains all data sets used to parameterize run timing dynamics.

Data on observer efficiency for adult coho salmon foot surveys were obtained from previous studies conducted on Black Creek, Vancouver Island, BC (Figure 2; unpublished data provided by J. Irvine, Fisheries and Oceans Canada and in Irvine et al. 1992). The study design, in which observers visually inspected fenced-off segments of stream prior to electrofishing, was repeated for 50 surveys occurring over seven years. A more detailed description of the study design and an analysis of the first three years of data are available in Irvine et al. (1992).

Analysis of Run Timing Data

Analysis of run timing data consisted of two basic steps. First, I used the 33 coho salmon visual survey data sets to select the best of five candidate models for describing coho salmon run timing dynamics (a normal distribution, a beta distribution, and three mixture models comprised of two normal distributions). In the second step, I used the "best" model from the first step to estimate interannual variability in run timing dynamics.

Model Selection

Equations describing the five run timing models considered are shown in Table 1 with model notation defined in Table 2. Model parameters are denoted using italicized lower-case letters (e.g. m), state variables are denoted using italicized upper-case letters (e.g. A), and functions are denoted using bold upper-case letters (e.g. **F**). Two probability density functions are used in the run timing models in Table 1. The first is the cumulative normal distribution function,

$$\mathbf{F_N}(d,m,\sigma)=\int_{i=0}^{d}\frac{1}{\sigma\sqrt{2\pi}}\,e^{\left(-\frac{(i-m)^2}{2\sigma^2}\right)}di \qquad (1)$$

and the second is the cumulative beta distribution function,

$$\mathbf{F_B}(x,\alpha,\beta)=\int_{i=0}^{x}x^{\alpha-1}(1-x)^{\beta-1}\,di \qquad (2)$$

Vectors of parameters estimated using maximum likelihood estimation for each of the run timing models are denoted $\hat{\Phi}_{model}$.

Run timing for the normal model (equations 3-6) is characterized using a cumulative normal distribution of fish arrival (Hilborn et al. 1999) that is scaled by the total number of fish observed over all surveys conducted on a single stream in a given year, C_T (equation 4). The cumulative number of counted fish that arrive by survey day, d, $(\hat{A}_d)$ is a function of the mean date of arrival into the survey area, m, and the standard deviation of arrival date, σ. The model used to describe fish death, $\hat{D}_d$, is the model used to describe fish arrival, offset by the average survey life, s (equation 5). The predicted count of fish during a given survey is the difference between $\hat{A}_d$ and $\hat{D}_d$ (equation 6). Run timing for the beta model (equations 6-9) is characterized using cumulative beta distributions for fish arrival and death (Hilborn et al. 1999, Korman et al. 2002). Two shape parameters, α and β, describe the shape of the beta distribution. Application of the beta model requires that start and end dates for stream arrival are explicitly specified so that day can be scaled between 0 and 1. Start and end dates are assumed to be the date of the first and last surveys each year, and n is the number of days between these two dates. As with the normal model, continuous arrivals and departures are scaled by the total number of fish observed over all surveys.

Run timing for the mixture models (equations 10-15) is characterized as two separate normal run timing models. When describing the mixture models, I use a superscript to denote which of the two component models a given parameter or state variable refers to (e.g. $\sigma^{'1}$ is the standard deviation of arrival for the first component and $\sigma^{'2}$ is the standard deviation of arrival for the second component). The standard deviation for both component models was constrained to be equal to or greater than two days during estimation. The mean date of arrival for the second model is assumed to occur k days after that of the first model. A parameter, z, is used to assign relative weight between 0 and 1 to the two component models in the mixture. The number of data points for most streams limited the number of parameters that could be estimated for the mixture models to four (equation 10), making it necessary to assume constant values of s and z to ensure parameter stability. The value of s was held constant at the mean of coho survey life estimates obtained from empirical studies (12.8 days; Table 3), while three different values of z were considered for each of the three mixture models ($z = 0.3, 0.5, 0.7$). The mixture models differed from the normal and beta models in that they could display both unimodal and bimodal shapes through adjustment of the k parameter. Small k values tended to create unimodal distributions while large values tended to create bimodal distributions.

Candidate models were fit to the observed count data sets using a maximum likelihood estimation (MLE) procedure assuming a Poisson error distribution for the counts, which is a commonly used distribution for dealing with discrete count data arising from continuous random sampling (Hilborn and Mangel 1997). The likelihood function for the Poisson distribution is,

$$L(C_d \mid \hat{C}_d) = \frac{1}{C_d!} \hat{C}_d^{\,C_d} e^{-\hat{C}_d} . \tag{16}$$

For each data set, the five candidate models were ranked using a small-sample version of Akaike's information criterion (AIC_c; Burnham and Anderson 2002). For each candidate model, an AIC_c value was calculated based on the maximum log-likelihood value (ℓ), as well as the number of parameters:

$$\mathrm{AIC}_c = -2\ell + 2K + \frac{2K(K+1)}{f_{max} - K - 1}, \tag{17}$$

where K is the number of parameters to be estimated and f_{max} is the number of surveys. The "best-model" for each of the 33 data sets, given the candidate models considered, was the one that produced the smallest $\mathrm{AIC_c}$ value.

The results of the $\mathrm{AIC_c}$ test showed that the mixture model with $z = 0.3$ was the best of the five candidate models for the largest number of data sets (Table 4). Based on these results, I selected the mixture model ($z = 0.3$) to estimate interannual variability in run timing parameters for the simulation procedure (below). While the use of a mixture model may seem unwarranted given the small number of data points and the possibility of over-fitting, its use is justified by the high interannual variability and bimodality in coho run timing dynamics and that arise from infrequent precipitation events (Figure 1; Appendix A). It is apparent that coho salmon are more likely to enter streams during, or immediately after, several days of high intensity precipitation, and that bimodal run timing curves are more common in years when dry periods are interspersed with periods of high precipitation. This level of variability cannot be produced using normal or beta run timing models that restrict run timing curves to a single mode.

Estimation of Interannual Variability

Interannual variation for each of the four mixture model parameters was determined by fitting the mixture model ($z = 0.3$) to each of the 33 visual survey data sets. I assumed that each stream i had a set of stream-specific mean parameter values (m'^1, k, σ'^1, and σ'^2 in Table 2) from which run timing dynamics deviated each year. For a given mixture model parameter θ, the actual parameter value observed for stream i in year t was a function of the stream-specific mean as follows,

$$\theta_{i,t} = \overline{\theta}_i + \varepsilon_{i,t}, \tag{18}$$

where $\overline{\theta}_i$ is the stream-specific mean and $\varepsilon_{i,t}$ is the deviation in year t. The assumption of mean parameter values was necessary to allow for the presence of stream-specific factors influencing arrival timing, such as local adaptation and migration distance (Smoker et al. 1998; Hodgson and Quinn 2002; Keefer et al. 2004). For each of the 33 data sets, an estimate of $\varepsilon_{i,t}$ was calculated for each year by subtracting $\overline{\theta}_i$ from $\theta_{i,t}$. The standard deviation of $\varepsilon_{i,t}$ within each stream, τ_i, was then calculated using the three estimates of $\varepsilon_{i,t}$ ($t = 1, 2, 3$). The 11 values of τ_i were used to develop alternative scenarios for the simulation of interannual variability in run timing dynamics (see below).

Analysis of Observer Efficiency Data

Analysis of the Black Creek observer efficiency data showed that the relationship between "true" abundance estimated from electrofishing, N, and survey counts, C, was linear, and that the intercept of the linear best-fit line did not differ significantly from zero ($C = a + bN$; H_0: $a = 0$; $p = 0.412$, $r^2 = 0.92$). There was no significant difference in the slope of the regression between years (analysis of covariance; $F = 2.09$, $p = 0.155$). Based on these results, I estimated the average observer efficiency over all surveys as the slope of the zero-intercept linear regression between "true" abundance and counts (Figure 2; $r^2 = 0.95$). The estimated slope was 0.865, indicating that on average, 86.5% of fish alive in the stream during a survey were detected by observers.

Simulation Procedure

The simulation procedure consisted of four major components: (i) a model of "true" population dynamics, including annual escapement and daily run timing into spawning streams, (ii) generation of survey count data with observation error, (iii) calculation of alternative visual survey indices from those simulated survey count data using four different estimation methods, and (iv) evaluation of monitoring performance for each visual survey estimation method (Figure 3). When describing the simulation procedure, I use the same notation for parameters, state variables, and functions described in the previous section. In addition, I use a capital letter in normal font to denote a variable that is assumed to be random in the simulations. A summary of parameter values used in the baseline scenario is provided in Table 5. The basic steps in the simulation procedure are as follows:

i) Generate a "true" 10-year escapement time series with a rate of decline less than or equal to the critical rate of decline (r = -0.04) that would result in a spawning aggregation being assessed as "threatened" under SARA listing criteria.

ii) For each year t in the escapement time series, generate "true" daily run timing dynamics with random variation.

iii) For each survey frequency considered (f_{max} = 1, 2, … 8 surveys per year), select survey dates.

iv) For each selected survey date, generate count data with observation error.

v) For each of the four escapement estimation methods (peak-count, mean-count, T-AUC, and L-AUC), calculate an index of escapement for year t based on the count data from f_{max} surveys.

vi) Using the 10 years of observed index values, calculate probability associated with null hypothesis H_0: stock is "not threatened" ($r > -0.04$)

vii) If the probability of H_0 is less than a pre-specified confidence level (e.g. 20%), reject H_0 and designate stock as "threatened".

viii) Repeat steps (i) to (vii) for 1000 simulation trials.

viii) For each combination of estimation method and survey frequency considered, calculate power as the proportion of simulation trials that correctly assess stock status as "threatened".

Population Dynamics

I used an exponential growth model to generate a declining time series of "true" escapement values,

$$E_t = E_0 e^{rt} , \tag{19}$$

where E is true escapement, t is time in years, and r is the intrinsic rate of population growth (i.e., for a declining population, $r < 0$). For the baseline scenario, in which the true percent change in escapement, p, was a 40% decline over 10 years (p = -40%), the value of r was -0.057 (Appendix B). In comparison, the critical SARA rate of decline used to designate a population as "threatened" was a 30% decline in escapement over 10 years (p = -30%; r = -0.04).

To simulate daily run timing dynamics for year t, the mixture model (z = 0.3) was modified slightly so that the cumulative normal distributions presented in equations 11-14 were scaled by total escapement, E_t, instead of total counts. This allowed for the prediction of the total number of arrivals and deaths that actually occurred up to a given day d, as opposed to the total number of arrivals and deaths counted up to d. The difference between cumulative arrivals and cumulative deaths on day d thus determines the total abundance of fish that are alive and in the survey area on that day (N_d). The total duration of fish presence in the survey area ranged from 17 to 91 days (mean = 51.5 days, n = 1000) for the simulated "true" run timing dynamics, which was realistic compared to the range of 19 to 82 days (mean = 47.3 days, n = 33) observed for the 33 visual survey data sets used to parameterize run timing dynamics.

Interannual variability in run timing dynamics was incorporated into simulated "true" data using the following variation of equation (18) to describe a given run timing parameter θ,

$$\theta_{i,t} = \overline{\theta}_i + \Upsilon, \qquad (20)$$

where Y is a year-specific random effect. For each parameter θ, Y was assumed to be a normally distributed random variable with a mean of zero and standard deviation of τ,

$$\Upsilon \sim \mathbf{N}(0, \tau), \qquad (21)$$

where the value of τ was specific to each parameter (denoted as τ_m, τ_k, τ_{σ_1}, and τ_{σ_2} in Table 5). In the baseline scenario, values of τ were set equal to the mean τ_i value estimated from visual survey data sets, as described in the model parameterization section.

Visual Survey Model

For each of the four escapement estimation methods evaluated, I considered a range of survey frequencies extending from one to eight surveys per spawning season. The mean-count, T-AUC, and L-AUC methods required a minimum of two, three, and five surveys per year, respectively. Only the peak-count method allowed for the consideration of a single-survey scenario.

Survey dates were selected based on the assumption that observers would have some knowledge of historic run timing within a study stream. On average, survey dates were centred around the historic peak day of spawning abundance for the stream, d_{PK}, however, the start date for surveys each year was randomly selected from a seven-day window. In the baseline scenario, a simple algorithm was developed that clustered survey events near d_{PK} when survey frequency was low. The single survey for the one-survey case of the peak-count method was conducted on d_{PK}, while the surveys for the two-survey case were conducted one week before and one week after d_{PK}. For the three- to eight-survey cases, the total number of days over which counts were conducted (l) was dependent on survey frequency as follows:

$$l = \begin{cases} 42 \text{ days} & f_{max} = 3 \\ 56 \text{ days} & f_{max} = 4 \\ 70 \text{ days} & f_{max} = 5 \\ 84 \text{ days} & f_{max} \geq 6 \end{cases} \tag{22}$$

where f_{max} is the total number of surveys per year. While the number of days of simulated fish presence was often shorter than l, the advantage of a longer survey period derives from the increased probability of observing both the start and end of fish presence in years that the peak date of fish abundance occurs either several days (or weeks) before or after d_{PK}. The first survey date (d_{first}) was randomly selected from a seven-day period that occurred $0.5l$ days before the historic peak date,

$$d_{first} = U\left(d_{PK} - 0.5l - 3, d_{PK} - 0.5l + 3\right), \tag{23}$$

and the remaining surveys were evenly spaced over the next l days so that the last survey date (d_{last}) was always

$$d_{last} = d_{first} + l. \tag{24}$$

Observation error in survey counts was simulated using a Poisson distribution. The number of fish observed on day d was a random Poisson variable with a rate parameter λ, defined as

$$\lambda = vN_d, \tag{25}$$

where v is the average observer efficiency of 0.865 estimated from Black Creek observer efficiency studies (described above).

Alternative Escapement Estimation Methods

Peak-count

The highest count value generated for each year in the time series was used as an index of escapement for the peak-count method:

$$I_{PK} = \max(C). \tag{26}$$

Mean-count

The mean of count values generated for each year in the time series was used as an index of escapement for the mean-count method:

$$I_{MN} = \frac{1}{f_{\max}} \sum_{d=1}^{d=f_{\max}} C_d. \tag{27}$$

Trapezoidal-AUC (T-AUC)

In the T-AUC method, a simple trapezoidal approximation was used to calculate the area under the observed run timing curve for each year,

$$AUC = \sum_{x=2}^{j} (u_x - u_{x-1}) \frac{(C_x + C_{x-1})}{2}, \tag{28}$$

where j is 2 plus the number of surveys conducted in a season, x represents a survey event, and u_x is the day that the x^{th} survey was conducted. To calculate area under the curve prior to the first count (AUC_{first}), and after the last count (AUC_{last}), I used the same approximation as Bue et al. (1998):

$$AUC_{first} = \frac{C_1 \hat{s}_t}{2} \tag{29}$$

$$AUC_{last} = \frac{C_{last} \hat{s}_t}{2} \tag{30}$$

where $\hat{s}_t$ is an annual estimate of survey life. When year- and stream-specific estimates of survey life are not available, as would be expected for extensive survey programs that monitor a large number of streams with minimal budget and personnel, survey life estimates must be extrapolated between years and/or streams. In the baseline scenario, I model a monitoring program in which a year-specific survey life value is estimated for a single stream and then applied to multiple streams within that year. An annual escapement estimate is derived from equation 28 as,

$$\hat{E} = \frac{AUC}{\hat{s}_t \hat{v}}, \tag{31}$$

where $\hat{v}$ is an estimate of average observer efficiency. The selection of $\hat{v}$ and $\hat{s}_t$ values for each simulation trial was based on the assumption that stock assessment analysts had some knowledge of the "true" underlying parameter distributions. This assumption is reasonable given the numerous studies that have been conducted on these parameter values.

The value of $\hat{v}$ was held constant at 0.865, which was the average observer efficiency value used to generate "true" count data in equation 25. Observation error in $\hat{s}_t$ was incorporated into the simulation procedure by assuming that $\hat{s}_t$ was a random normal variable with a mean, μ_s, equal to s and a coefficient of variation, CV($\hat{s}$), of 0.2. While the selection of a CV=0.2 for survey life estimates in the baseline scenario was somewhat arbitrary, studies of the length of time coho salmon spend on their redds (redd residence time) in two coastal streams suggest that this value is reasonable for monitoring programs that estimate survey life for a single stream each year and then extrapolate that value to other streams. Over a period of four years, the redd residence time estimates for French Creek and Black Creek on eastern Vancouver Island, BC differed from each other by an average of 4.5 days (range of difference = 0.5 to 8.7 days; English et al. 1992, Irvine et al. 1992).

Likelihood AUC (L-AUC)

The L-AUC method (Hilborn et al. 1999) estimates escapement by treating it as a free parameter in a maximum likelihood estimation procedure. The application of the beta-distributed run timing model to the L-AUC method requires a slight modification from the version presented in equations 6-9. The scalar C_T in equations 7-8 is replaced with a total escapement parameter for year t, E_t. This modification allows equation 6 to predict the total abundance of fish alive in the survey area on day d, $\hat{N}_d$. A deterministic observation model is then used to predict the number of fish counted, $\hat{C}_d$, as a function of $\hat{N}_d$ and an estimate of observer efficiency $\hat{v}$,

$$\hat{C}_d = \hat{v}\hat{N}_d. \tag{32}$$

Using the modified run timing model and the observation model, maximum likelihood estimation was used to fit predicted counts to observed counts by estimating four parameters: total escapement $(\hat{E})$, survey life $(\hat{s})$, and two shape parameters of the beta distribution ($\hat{\alpha}$ and $\hat{\beta}$). The L-AUC method was only applicable to survey frequencies of 5 per year or greater to prevent model over-parameterization.

Parameters were estimated using a penalized likelihood function that assumed a Poisson error distribution for the count data (equation 16) and a normal prior distribution on the parameter s, with a mean of μ_s and a standard deviation of σ_s. The total likelihood was thus,

$$L\left(\mathrm{C} \middle| \hat{C}, \alpha, \beta, s\right) = \frac{1}{\sigma_s\sqrt{2\pi}} \exp^{\left(-\frac{(s-\mu_s)^2}{2\sigma_s^2}\right)} \prod_{d=1}^{d=f_{\max}} L\left(\mathrm{C}_d \mid \hat{C}_d\right). \tag{33}$$

where C_d is the actual count value observed on day d (equation 25) and $\hat{C}_d$ is the predicted count value for day d (equation 32). As with the T-AUC method, the value of μ_s was set equal to the true value of s used to generate data (above) and the value of $\hat{v}$ was held constant at 0.865 (equation 25). The inclusion of a prior distribution on s was necessary to ensure parameter stability in the estimation procedure. In some cases, v can be treated as a free parameter in the maximum likelihood estimation procedure (Hilborn et al. 1999); however, parameter estimates of α, β, and s were highly correlated at all survey frequencies examined when v was estimated rather than assumed.

Evaluation of Monitoring Performance

Trend Detection

For each Monte Carlo trial, an observed escapement index, I ($I = I_{t=1}, I_{t=2}, \ldots I_{t=t\max}$), was generated for each combination of estimation method and survey frequency considered (henceforth referred to as "monitoring designs"). Performance of each design was evaluated based on its power, which was the probability that a simple linear regression of the logarithm of the observed index on t would correctly conclude that the value of r was less than or equal to the critical r value, -0.04, (reject H_0: $r > -0.04$) that would lead to a population being assessed as "threatened" under the SARA assessment criteria (i.e., a 30% decline in abundance over 10 years; Appendix B). To calculate a predicted rate of population growth from each simulated escapement index I that could be compared with both the true r and the critical r value of –0.04, I used a simple regression on log-transformed index data,

$$\log_e(I_t) = \log_e(\hat{I}_0) + \hat{r}t, \tag{34}$$

where $\hat{r}$ is the predicted rate of population growth, calculated as the slope of the best-fit line to simulated index values.

Uncertainty in $\hat{r}$ was incorporated into trend detection using a Bayesian approach to regression analysis. I used a sampling-importance-resampling algorithm (SIR) to construct a joint probability density function (pdf) for the linear regression, $P(r, I_0 \mid I)$, from which I obtained the marginal posterior pdf, $P(r \mid I)$ (Rubin 1988, Gelman et al. 2004; Appendix C). The probability that r was within a specified range was thus the proportion of $P(r \mid I)$ that fell within that range (Figure 4). To evaluate the bias of each monitoring design, the expected r value from each Monte Carlo trial, $\mathbf{E}[r \mid I]$, was used over 1000 simulation trials.

Power Analysis

I used the Monte Carlo simulation procedure to approximate the statistical power of alternative monitoring designs to detect an r value ≤ -0.04 using a Bayesian approach to hypothesis testing (Gelman et al. 2004). In each trial, the posterior probability distribution $P(r \mid I)$ was used to calculate the probabilities for two competing hypotheses about the true value of r (H_0: $r > -0.04$ [or $p > -30\%$]; H_1: $r \leq -0.04$ [or $p \leq -30\%$]). Two examples of the calculation of probabilities for H_0 and H_1 are shown in Figure 4. In panel A of Figure 4, only 3% of $P(r \mid I)$ falls above -0.04, indicating a 3% probability that the rate of decline is less than that associated with $r = -0.04$ (i.e., that the time trend is less steep than $r = -0.04$). In this case, the probability assigned to H_0 is 3% while that assigned to H_1 is 97%. In panel B, 86% of $P(r \mid I)$ falls above -0.04, so the probabilities assigned to H_0 and H_1 are 86% and 14%, respectively.

I placed equal weight on the probabilities of type I and type II errors for detecting trends in escapement ($\alpha = \beta = 0.20$; power = $1-\beta = 0.80$). Because the r value used to generate escapement dynamics in the baseline scenario ($r = -0.057$) followed H_1 ($-0.057 \leq -0.04$), I considered trend detection successful when the probability that the slope declined less steeply than $r = -0.04$ was less than 0.2 (i.e., the probability for H_0, $P(r >$

-0.04), was < 0.2). The power of each monitoring design was the percentage of Monte Carlo trials in which trend detection was successful.

While classical hypothesis testing is sometimes considered incompatible with Bayesian analysis, posterior probabilities can correspond to conventional *p*-values for simple, one-sided hypothesis tests such as this (Gelman et al. 2004). I have chosen hypothesis testing to measure performance because I am primarily interested in whether the correct management decision will be made in my hypothetical monitoring program (i.e., whether the population is assessed as "threatened" and appropriate management action taken). I chose to employ a Bayesian approach to trend detection analysis, as opposed to a simple linear regression, to demonstrate how Bayesian methods can be used to produce estimates of r that communicate measures of uncertainty.

In addition to using the power of a monitoring program to detect $r \leq -0.04$ as a performance measure, I used power to examine two alternative performance measures for the baseline scenario that would likely be of interest to fisheries managers. The first was the minimum number of years required to detect $r \leq -0.04$ with 80% power and the second was the minimum "true" percent decline over 10 years that would allow for the detection of $r \leq -0.04$ with 80% power.

Sensitivity Analysis

I used sensitivity analyses to examine how deviations from three key assumptions about the "true" state of nature affected the power of each monitoring design to detect $r \leq -0.04$: (i) interannual variability in run timing, (ii) among-survey variability in observer efficiency, and (iii) the true rate of population decline (Table 6). For the first analysis, three additional levels of variability in run timing were examined (none, low, and high) by changing the values of τ in equation 21. To cover the range of between-year variability observed among the 11 streams analyzed, the value of τ in the low and high variability scenarios was set equal to the highest and lowest τ_i values estimated from the 11 streams. In the second sensitivity analysis, three alternative levels of variability in observer efficiency were examined (none, low, and high). Because the variance of a Poisson distribution is equal to the mean, variability in observer efficiency was altered by changing v in equation 25. For a given level of N, the variance of C was increased when v was increased. In the no-variability scenario, the number of fish observed was not treated as a random Poisson variable. Instead, C_d was set equal to N_d for all surveys. In the final sensitivity analysis, a range of alternative rates of population decline (p = 0, -2, -4, … -60% change over 10 years) were examined by changing the value of r in equation 19. For this analysis, I was interested in the ability of each monitoring design to correctly respond to changes in the "true" percent decline by adjusting the level of power achieved when detecting the SARA critical rate of decline of $r \leq -0.04$. I use the term "responsive" in my Results section to describe a monitoring program in which the power to detect $r \leq -0.04$ changes considerably per unit change in the true percent decline.

Survey Design Scenarios

To determine how the design of visual survey monitoring programs could affect the power of trend detection, I examined a range of alternative scenarios on (i) the level of error in estimates of survey life [CV($\hat{s}$)], (ii) the method used to select survey dates within a season (hereafter referred to as survey spacing), (iii) variation in survey frequency across years, and (iv) the number of years over which monitoring was conducted (Table 7). In the first of these analyses, I examined the sensitivity of the T-

AUC method's performance to three alternative scenarios for CV($\hat{s}$) (none, low, high). Only the T-AUC method was considered in the analysis because it is the only method that requires an annual estimate of s. In comparison, the L-AUC method treats s as an unknown parameter in maximum likelihood estimation. I considered CV($\hat{s}$) a component of the survey design through the assumption that stock assessment personnel have some control over this value based on the level of funding they choose to allocate towards obtaining year- and stream-specific estimates of $\hat{s}$. To examine how the spacing of survey events within a year affected the power of trend detection, I considered two alternatives to the baseline scenario shown in equations 22-24. The first was an even spacing design, in which surveys were evenly spaced over $l = 77$ days. The second was a random spacing design in which the survey period of $l = 77$ days was stratified into f_{max} intervals of even size and one survey date was randomly sampled from each interval using a random uniform distribution. To examine how across-year variation in survey frequency affected the power of trend detection, I considered four alternative scenarios, in addition to the constant frequency scenarios used in the baseline case. For each scenario, the survey frequency for a given year was selected from a random uniform distribution (Table 7). Due to the requirement for a minimum of five surveys per year for the L-AUC method, only one of the four alternative scenarios (low variability / high frequency) was applicable.

Results

Baseline Scenario

In the baseline scenario, which provided the closest representation to "true" population and survey dynamics based on my analysis of coho visual survey data, all survey designs considered produced relatively unbiased estimates of r when averaged across 1000 simulations. Percent bias remained within ± 3% for all monitoring designs evaluated.

The power of all four estimation methods to detect $r \leq -0.04$ increased with increasing survey frequency; however, the mean-count method was able to achieve greater power than the peak-count, T-AUC, and L-AUC methods (Figure 5). In the baseline scenario, the mean-count method achieved over 75% power at five or more survey counts per year and over 95% power at seven or more survey counts per year when the true rate of decline was 40% over 10 years ($r = -0.057$). The L-AUC method achieved only slightly lower power than the mean-count method at all survey frequencies examined, but it required at least five surveys. In contrast, the peak-count and T-AUC methods only achieved a maximum of 56 and 60% power, respectively, at eight survey counts per year. The single-survey case of the peak-count method had the lowest power of all monitoring programs considered, with "successful" trend detection occurring in only 39% of the simulation trials.

At survey frequencies of six or greater, the mean-count and L-AUC methods were able to achieve 80% power to detect $r \leq -0.04$ within six to eight years (Table 8). When only four surveys were conducted in a year, the mean-count method achieved 80% power to detect $r \leq -0.04$ in 13 years, whereas the L-AUC method could not be applied with so few surveys per year. The peak and T-AUC methods required as few as 16 years to achieve 80% power, but seven survey counts per year were required to achieve this.

The mean-count and L-AUC methods were able to achieve 80% power to detect $r \leq -0.04$ at seven or more survey counts per year when detecting declines in escapement of

36% or greater over 10 years (Table 8). Using the same power level, the peak-count and T-AUC methods were only able to detect 50% declines in escapement over 10 years at seven or eight survey counts per year.

Sensitivity Analysis

The power of all four monitoring methods depended on the level of interannual variability in run timing (τ in equation 21 and Table 6); however, the mean-count method maintained the highest power to detect $r \leq -0.04$ at all levels of variability examined (Figure 6). At five of more survey counts per year, the L-AUC method achieved levels of power only slightly less than the mean-count method. The power of the mean-count and L-AUC methods to detect $r \leq -0.04$ in the baseline, low variability, and no variability cases converged towards 100% at survey frequencies greater than or equal to five survey counts per year. In the high variability case, the mean-count method was able to achieve up to 65% power to detect $r \leq -0.04$, whereas the L-AUC never got above 50%. The peak-count method required perfectly consistent run timing over all years ($\tau = 0$) to achieve the level of power achieved by the mean-count method in the baseline scenario. The T-AUC method was unable to achieve greater than 60% power to detect $r \leq -0.04$, even when run timing was held constant.

The range of among-survey variability in observer efficiency examined, from v known with perfect information ($C_d = N_d$) to $v = 0.96$, had a relatively small effect on the power for all estimation methods to detect $r \leq -0.04$ (Figure 7). While all four methods experienced some decline in power when variability in observer efficiency was increased above the baseline level of 0.865 to 0.965, differences in power between the baseline scenario and the high variability scenario did not exceed 5.5% for any monitoring program examined. The gains in power achieved by eliminating among-survey variability in observer efficiency in the "no variability" scenario were also relatively small (< 6.6%).

The power of all four estimation methods to detect $r \leq -0.04$ was highly dependent on the "true" p value used to generate escapement dynamics (Figures 8-9). The effect of survey frequency on power depended on the value of p relative to that associated with the null hypothesis used in trend detection analysis (H_0: $r > -0.04$ [$p > -30\%$]). When the null hypothesis was false ($p = -40\%$ and $p = -60\%$ in Figure 8), the power of all four estimation methods to detect $r \leq -0.04$ (i.e., correctly reject H_0) increased with increasing survey frequency. As in the baseline scenario, the mean-count and L-AUC methods were able to achieve greater gains in power with increased survey frequency than the peak-count and T-AUC methods. Conversely, when the null hypothesis was true ($p = -20\%$ in Figure 8), the power of all four methods to detect $r \leq -0.04$ (i.e., incorrectly reject H_0) decreased with increasing survey frequency.

When the responsiveness of monitoring performance to changes in the "true" percent decline was examined over smaller increments of p, it was evident that the power of the mean-count and L-AUC methods to detect $r \leq -0.04$ became increasingly responsive to changes in the "true" percent decline as survey frequency increased, as shown by the narrowing of the contour lines in Figure 9. As noted in the methods section, I refer to a monitoring design as responsive when its power to detect $r \leq -0.04$ changes considerably in response to slight changes in the "true" percent decline. In comparison, the peak-count and T-AUC methods were less responsive, displaying more gradual changes in power per unit change in the "true" percent decline.

The high responsiveness of the mean-count and L-AUC methods to the "true" percent decline increased the probability of correctly assessing a population as "threatened" when it really was threatened and reduced the probability of incorrectly assessing a population as "threatened" when it was not. As an example of the latter case, when the true percent decline was only 20% over 10 years (i.e., not threatened), the mean-count method had a less than 10% probability of incorrectly assessing the population as "threatened" at five or more survey counts per year, while the peak-count and T-AUC method had greater than 10% probability (Figure 9). The single survey case of the peak-count method was highly unresponsive to the "true" rate of decline. This method had only a 60% probability of correctly assessing the population as "threatened" when the true percent decline was 60% over 10 years, and had a greater than 10% probability of incorrectly assessing the population as threatened when escapement remained constant (0% decline over 10 years).

Survey Design Scenarios

The power of the T-AUC method to detect $r \leq -0.04$ was highly sensitive to the level of error in estimates of survey life, [CV($\hat{s}$)] (Figure 10). Perfect information about survey life (CV($\hat{s}$)=0) was needed to obtain the high level of power achieved by the mean-count method in the baseline scenario (CV($\hat{s}$)=0.20). In the low error case (CV($\hat{s}$) = 0.10), the T-AUC method was unable to achieve more than 78% power at eight survey counts per year. In the high error case (CV($\hat{s}$) = 0.30), it achieved a maximum of 48% power at eight survey counts per year.

The power of all four monitoring methods to detect $r \leq -0.04$ was dependent on the spacing of survey events within a year, although this was particularly true for the mean-count and L-AUC methods (Figure 11). For most combinations of survey frequency and estimation method, power to detect $r \leq -0.04$ was greatest when surveys were spaced according to the baseline algorithm that increasingly clustered surveys around the historic peak date of spawning abundance when survey frequency was low (equations 22-24). Between three and five survey counts per year, gains in power achieved by using the baseline algorithm over even spacing became smaller with each additional increment in survey counts per year due to increased similarity of survey spacing between the two designs. At six or more survey counts per year, survey spacing was the same under the baseline and even spacing design, which resulted in comparable levels of power to detect $r \leq -0.04$. The random spacing design resulted in the lowest performance for all methods; however, the loss of power under the random spacing design was especially large for the mean-count and L-AUC methods. In the baseline and even spacing designs, the mean-count and L-AUC methods achieved over 95% power to detect $r \leq -0.04$ at six or more survey counts per year, while in the random spacing design, these methods were unable to achieve greater than 85% power at eight survey counts per year. The reduced power of all four methods to detect $r \leq -0.04$ under the random spacing design, compared to the baseline and even spacing designs, shows that monitoring performance is greatest when the number of days between surveys remains constant between years.

Examination of alternative scenarios regarding among-year variability in survey frequency revealed that when survey frequency varied among years, power to detect $r \leq -0.04$ tended to be limited by the lowest survey frequency in the time series. While this effect was observed for all estimation methods, it was particularly notable for the mean-count and L-AUC methods, which achieved the greatest gains in power with increased survey frequency. For example, when survey frequency varied between three and five counts per year, the power of each method to detect $r \leq -0.04$ was similar to that achieved with a constant survey frequency of three, and when survey frequency varied between three and eight, power was similar to that achieved with a constant survey frequency of four.

The power of all four estimation methods to detect $r \leq -0.04$ increased as the number of years of escapement monitoring increased (Figure 12). For a given survey frequency, larger gains in power associated with increased duration of monitoring would be accured to the mean-count and L-AUC methods, as shown by the narrower spacing of contour lines for these two methods in Figure 12.

Discussion

Performance of Alternative Estimation Methods

The results of these power analyses suggest that a simple mean-count method is more suitable than commonly used visual survey estimation methods for monitoring coho salmon escapement for a variety of reasons (Table 9). While all four methods that I examined estimated the "true" rate of population decline with relatively low bias, the mean-count method consistently achieved higher levels of precision (as reflected by greater power for trend detection) than peak-count, T-AUC, and L-AUC methods over a wide range of scenarios about true population parameters and survey design. High precision is a desirable property for monitoring programs because it increases the probability of correctly estimating the true status of a population. In the baseline scenario, in which the true rate of decline was 40% over 10 years, the mean-count method was able to achieve 75% or greater power when detecting declines in coho escapement of 30% over 10 years with five or more survey counts per year. In order for the power of the T-AUC method to be comparable to that of the mean-count method, survey life estimates must be constant among years. Similarly, in order for the monitoring performance of the L-AUC method with a beta-distributed run timing model to be comparable to the mean-count method, observer efficiency estimates must remain constant, as was done in the baseline scenario. In both of these cases, escapement estimates should be viewed as relative indices instead of absolute escapement estimates because survey life and observer efficiency vary annually. While the two AUC methods may be more appealing for management purposes because they produce index values on a scale comparable to actual escapements, the application of a constant scalar to annual mean-count values would have the same effect. The poor performance of the peak-count method when interannual variability in run timing dynamics was set at the baseline level suggests that this method is not suitable for monitoring coho salmon populations because, as I have shown, they generally display high interannual variability in run timing dynamics.

The level of power achieved by the L-AUC method with a beta-distributed run timing model was only slightly less than that of the mean-count method; however, the L-AUC method is limited to survey designs with five or more survey counts per year, which is not always possible for extensive survey programs (D. Peacock, Fisheries and Oceans Canada, Prince Rupert, BC, pers comm.). The mean-count method compared to the L-AUC method, the mean-count method is better suited for extensive monitoring of coho salmon escapements given its higher levels of power, the simpler estimation procedure, and the smaller minimum survey frequency (as few as two surveys per year).

This study differed from previous studies evaluating the utility of visual survey programs for monitoring Pacific salmon escapement in two important ways. First, because I was interested in the utility of visual survey programs for monitoring time trends in escapement as opposed to estimating absolute escapement, I measured performance as the power of trend detection over a period of several years. Previous studies have tended to focus on the ability of visual surveys to estimate absolute escapement within a single year (e.g. English et al. 1992, Hilborn et al. 1999). However, when the objective of a monitoring program is to detect trends in escapement, which is the case for extensive survey programs described in the WSP, the accuracy of annual escapement estimates is not of direct interest. Of greater interest is the ability of visual survey indices to detect biologically important trends in escapement in a time frame that is relevant to management. This is not to say that absolute estimates of escapement are not important for other assessment purposes. However, because I have chosen to focus on the extensive survey component of WSP monitoring plans, I am not interested in absolute estimates of escapement because they are not essential for detecting time trends in escapement.

The second way in which this study differs from others is that I used a simulation modelling approach that allowed me to explicitly incorporate observation error in visual survey methods and interannual variability in coho run timing dynamics into the generation of count data, and to identify an estimation method that is reliable at detecting time trends under a wide range of scenarios about these factors. Korman and Higgins

(1997) used Monte Carlo simulations to examine the utility of escapement data based on visual surveys to monitor changes in chinook salmon escapements caused by habitat alterations. However, they did not explicitly model the individual components of a visual survey program. Specifically, while their results showed that the power of visual survey methods to detect trends increased substantially with increased precision in escapement estimates, they were unable to provide insight about the specific components of the visual survey process that limited power (e.g. estimation method, survey frequency, spacing of surveys, variable observer efficiency).

The usefulness of employing a simulation modelling approach to evaluate visual survey designs and estimation methods has been demonstrated for monitoring programs for migratory birds (Thomas 1996) and marine mammals (Adkison et al. 2003). Hill (1997) used simulation modelling to determine the effect of survey frequency on the precision of AUC escapement estimates for Nechacko River chinook salmon; however, he focused on the escapement estimate derived from a single year and did not consider variable run timing dynamics or alternative estimation methods. Korman et al. (2002) used a similar approach to mine to examine how survey design affected the reliability of an annual escapement estimate for a stock of winter-run steelhead (*O. mykiss*) returning to spawn in a single year. To the best of my knowledge however, my evaluation of visual survey programs for monitoring coho salmon populations is the first to use simulation modelling to examine how interannual variability in salmon run timing dynamics and survey design affect the power of long-term trend detection for alternative estimation methods.

Sensitivity Analysis of Power

Power is a function of effect size, sample size, sample variance, and the specified significance level (Peterman 1990). Thus, the levels of power to detect $r \leq -0.04$ achieved by the alternative estimation methods are contingent on model assumptions made about these four factors. My sensitivity analyses provided insight about how the first three factors affect the power of trend detection for coho salmon visual survey programs. Power to detect $r \leq -0.04$ was positively related to effect size and sample size, as evident from the increases in power associated with increasing rates of decrease in escapement and number of years of monitoring, respectively. Positive relationships between effect size and power, and sample size and power, are well known (Zar 1999), and have been demonstrated in previous power analyses of fisheries monitoring programs (Peterman and Bradford 1987, Maxwell and Jennings 2005). Changes in power associated with different levels of variability in run timing, survey frequency, survey spacing, precision in survey life estimates, and variability in observer efficiency can be explained in terms of the effect these variables have on sample variance, which is negatively related to power (Peterman 1990).

The varied response of the peak-count, mean-count, T-AUC, and L-AUC methods to the level of variability in run timing is a function of the assumptions required by each estimation method with regard to run timing. The peak-count method utilizes information from only one survey event each year (i.e., the survey in which the highest count value was obtained) and therefore does not incorporate any information about the shape of the run timing curve. Instead, it relies on the assumption that the ratio of the peak-count to total escapement is constant across years. My results show that this assumption is weakly supported for coho salmon stocks. Visual survey data used to parameterise run timing dynamics showed high interannual variability in both the timing of fish presence and the shape of the observed run timing curve. The high sensitivity of monitoring performance of the peak count method to the level of variability in simulated run timing dynamics further supports this conclusion. While increases in survey frequency produced slight improvements in the power of the peak-count method by increasing the probability of observing the "true" peak abundance, variability in the width of the run timing curve and the number of modes limited performance of peak-counts. In contrast, the mean-count method incorporates all available information on the shape of run timing curves into the escapement index by using data from all surveys conducted in a year and is thus able to achieve large gains in power with increased survey frequency.

The T-AUC method also uses information from all survey events conducted in a year; however, error in survey life estimates limited its ability to achieve the large gains in power associated with increased survey frequency that were observed for the mean-count method. Even when variability in run timing was low, the T-AUC method was unable to achieve greater than 65% power with the level of error in survey life estimates assumed in the baseline scenario [CV($\hat{s}$) = 0.20]. When error in survey life estimates was low [CV($\hat{s}$) = 0.10] or absent [CV($\hat{s}$) = 0], increased survey frequency had a positive effect on power. Hill (1997) also found that when errors in survey life were relatively low (i.e., survey life treated as a random normal variable with mean =10 days, standard error = 0.5 days, and CV[$\hat{s}$] = 0.13), the reliability of annual escapement estimates derived using the T-AUC method increased with increasing survey frequency. The lower level of error assumed by Hill (1997), compared to my baseline scenario, was because he was simulating a monitoring program that obtained year- and stream-specific estimates of survey life.

The low sensitivity of monitoring performance for all estimation methods to among-survey variability in observer efficiency is likely a result of the small differences in variability among the four scenarios. For each scenario, I assumed that observation error followed a Poisson distribution with a constant rate parameter set at the mean observer efficiency value estimated from Black Creek observer efficiency studies (Irvine et al. 1992). While varying the mean observer efficiency between different scenarios allowed me to affect the variance of the Poisson distribution, the overall change in the coefficient of variation decreased as daily abundances increased. The assumption of a Poisson error distribution for generating count data is highly uncertain. Unfortunately, the limited data available on between-survey variability for coho salmon and the challenge of separating observation errors from process errors when calculating observer efficiency made it necessary for me to assume an error structure. A Poisson distribution was used because it is commonly associated with count data (Hilborn and Mangel 1997). Korman et al. (2002) also assumed observation error followed a Poisson distribution for visual surveys of spawning steelhead, while Hill (1997) randomly sampled from empirically-derived estimates of observation error for chinook salmon visual surveys. In my study, the equal sensitivity of all estimation methods to variability in observer efficiency suggests that, even if a Poisson distribution is not suitable, the approach I have taken is sufficient for comparing the relative performance of the four methods.

While escapement monitoring programs should be able to reliably detect situations of concern, such as a rate of population decline that would result in a stock being assessed as "threatened" under SARA, they should also minimize the probability of falsely detecting these situations. My sensitivity analysis of the response of power to changes in the "true" percent decline showed that the mean-count and L-AUC methods performed better than the other two methods using both criteria. While the power of the mean-count and L-AUC methods to detect the critical SARA rate of decline increased quickly as the "true" rate of percent decline increased, power also decreased quickly as the "true" rate of decline decreased, resulting in a lower probability of false positives (incorrectly rejecting H_0: $r > -0.04$) for these two methods than for the peak-count and T-AUC methods. The performance of the single survey case of the peak-count method was especially poor using the above criteria. Even when escapement was constant, this monitoring design incorrectly "detected" the critical SARA rate of decline in greater than 10% of simulations.

Survey Design Scenarios

The requirement for highly precise estimates of survey life in order for the T-AUC method to achieve high power in trend detection analysis is consistent with previous studies showing that the accuracy of annual AUC escapement estimates is highly dependent on both year- and stream-specific estimates of survey life (English et al. 1992, Irvine et al. 1992, Bue et al. 1998). Unfortunately, the cost and effort requirements of survey methods used to estimate this parameter can be high. Previously applied methods for estimating survey life for Pacific salmon include tagging programs and enumeration fences (English et al. 1992, Bue et al. 1998), capture-recapture studies (Manske and Schwarz 2000) and daily observations (Van den Berghe and Gross 1986).

If year- and stream-specific estimates of survey life are not used for the T-AUC method, the costs of all four estimation methods are expected to be comparable. In addition to fixed costs associated with operating visual survey field programs that are independent of survey frequency (e.g. administration, data analysis and equipment), there are variable costs linearly related to survey frequency (e.g., field personnel hours, travel expenses). The requirement for an additional field program aimed at estimating survey life for the T-AUC method would greatly increase both variable and fixed costs for this method, making it less attractive for extensive monitoring from a budgetary point of view. While the increase in cost incurred in this case would be dependent on the method used to estimate survey life, most methods are highly effort-intensive, requiring field personnel to be present in the survey area on a daily basis while fish are present. Irvine et al. (1992) found that, while each visual survey count required only two person-days (i.e,, 10 person-days for a survey frequency of five counts per year), the estimation of observer efficiency and survey life required an additional 22 person-days. Equipment costs, such as radio telemetry gear or enumeration fences, would further increase budgets for programs estimating survey life annually.

When year- and stream-specific estimates of survey life are not available, there are two ways in which survey life estimates can be extrapolated between years and streams (Perrin and Irvine 1990). The first approach is the one assumed in the baseline scenario, in which a year-specific survey life value is estimated for a single stream and then applied to multiple streams within that year. The second approach is to apply a constant survey life estimate to all years in a time series. While the second approach of a constant survey life would reduce the accuracy of annual escapement estimates by ignoring interannual variability in that parameter, the high levels of power to detect $r \leq -0.04$ achieved by the T-AUC method when survey life was held constant shows that this approach provides a more reliable index of escapement than the first approach, which was used in the baseline case. In this case however, the T-AUC method would serve as only a relative index of escapement, providing the same level of information as the mean-count method.

The evaluation of alternative survey design scenarios regarding the spacing of surveys within a year and variability in survey frequency across years demonstrates the importance of establishing a standardized sampling protocol that can be applied consistently over multiple years. In order for the mean-count and L-AUC methods to achieve the high levels of power for trend detection seen in the baseline scenario, both the spacing of surveys and the frequency of surveys should remain constant between years. The simple algorithm for spacing surveys used in the baseline scenario (equations 22 - 24), in which survey dates were increasingly clustered around the historical peak date of spawning abundance at low survey frequencies (3 – 5 surveys per year), tended to produce higher levels of power at these frequencies than when surveys were spaced evenly over the historic spawning period. While the particular spacing design that maximizes power would likely differ between streams due to varying lengths of freshwater residence, the advantage of generally clustering surveys near the historic peak date of spawning abundance is expected to apply to all streams. The decreased levels of power for trend detection achieved by the mean-count and L-AUC methods when survey dates were spaced randomly shows that, in order for these methods to achieve maximum levels of power, the number of days between surveys should be held constant among years. These results have implications not only for the design of coho visual survey programs, but also for the application of these methods to historic time series of coho count data that have not followed a consistent sampling protocol. When inconsistencies in survey spacing and frequency exist within a time series, the gains in power achieved by the mean-count and L-AUC methods compared to the other two methods will be reduced. However, the mean-count and L-AUC methods still achieved higher power than the other two estimation methods in all survey design scenarios considered.

While my results are specific to coho spawning aggregations from the interior Fraser River and the east coast of Vancouver Island, British Columbia, the sensitivity analyses and the examination of alternative survey design scenarios can be used to make inferences about the suitability of alternative estimation methods for other species of Pacific salmon. For example, pink and sockeye salmon tend to have relatively low interannual variability in run timing dynamics (Burgner 1991, Heard 1991), and thus, the peak-count method may be able to achieve higher levels of power in trend detection than was observered for coho salmon. Conversely, chum salmon (*O. keta*) are similar to coho salmon in their strong dependence on environmental cues for initiating upstream migration (Salo 1991), potentially making the peak-count method less suitable. While the extrapolation of survey life estimates between multiple streams within a single year had a strong negative effect on the power of the T-AUC method when CV($\hat{s}$) was assumed to be 0.2, which I based on coho salmon populations on the east coast of Vancouver Island, this may not be the case for streams that display high regional consistency in survey life within a year, resulting in a lower CV($\hat{s}$) value.

Establishing Benchmarks for Extensive Surveys

My results suggest that a goal of obtaining accurate estimates of absolute coho salmon escapement is unrealistic given the limited information available from extensive survey programs. Escapement estimates produced by the T-AUC method were unreliable and the L-AUC method was not able to estimate absolute escapement for fewer than five surveys per year, or to estimate observer efficiency as part of the maximum likelihood procedure. While the application of benchmarks based on absolute escapement are commonly employed for salmon management, the development of rate-based benchmarks, such as the rate of change in escapement over time, or relative benchmarks, such as the mean-count over all surveys in a year, may be more suitable for the extensive survey component of WSP monitoring plans. Through the application of a Bayesian approach to trend detection analysis, I have shown how uncertainty in escapement monitoring can be incorporated into extensive monitoring programs using rate-based benchmarks. The calculation of a probability distribution to summarize information about estimated rates of change in escapement provides decision-makers with a simple, visual representation of uncertainty in estimated rates of change and encourages discussions about the biological relevance of the results (Wade 2000). For example, a person presented with the two plots in Figure 4 can easily infer that the "best-estimate" of r (the r value with the highest probability given the data) in plot A is more certain than that in plot B, as shown by the narrower distribution for plot A. An alternative approach to using rate-based benchmarks would be to establish benchmarks in the units of a relative index, such as the mean-count. Relative indices based on peak-counts are currently used for monitoring by some management agencies (e.g. Geiger and McPherson 2004); however, to my knowledge, the mean-count has not been applied for this purpose.

In this study, I focused on the ability of visual survey methods to monitor escapement within a single spawning stream. However, given that the primary purpose of extensive visual survey programs, as identified the WSP, is to examine distribution and consistency in escapement trends throughout a region by monitoring a subset of streams within a large geographic area, future research is needed to develop a means of combining data from several streams into a single performance measure. There are two ways in which the approach taken here could be adapted for application to extensive survey programs that monitor multiple streams within a CU. The first is to develop a composite index that represents all streams monitored (e.g. the sum of coho salmon peak-counts over all streams in a year). For example, composite indices of several vulnerable species have been shown to have a greater power to detect trends than individual species indices for the English bottom trawl survey when all composite species displayed similar trends (Maxwell and Jennings 2005). Composite indices of peak-counts from several streams are currently used to monitor escapement relative to escapement goals in southeast Alaska (Geiger and McPhearson 2004). The second approach is to apply monitoring criteria consisting of three components (Peterman 2004): (i) a biologically significant rate of decline, (ii) a statement about the proportion of spawning aggregations within a CU that have rates of decline greater than the biologically significant rate of decline (e.g. 3 out of 10 spawning aggregations), and (iii) a statement about the probability of component (ii). Using this approach, an example benchmark for assessing status could be "less than a 40% probability that more than 3 out of 10 spawning aggregations within a CU are declining at a rate greater than 30% over 10 years".

The number of years of escapement monitoring and the significance level used for hypothesis testing should be selected based on the specific objectives of individual escapement monitoring programs. The two additional performance measures I examined, the minimum detectable "true" rate of decline with 80% power and the number of years of monitoring required to achieve 80% power, demonstrate alternative options for communicating the results of power analyses that explicitly highlight trade-offs between power and the minimum detectable effect size or sample size. For example, if the goal of a visual survey monitoring program is to detect the critical SARA rate of decline in 10 years with 80% power, my results indicate that the mean-count and L-AUC methods are able to do so for rates of decline of 36 % or higher, while the other two methods require rates of at least 50%. The latter two methods would be unacceptable because they require at least a 50% decline in escapement before achieving a probability of 80% or greater of detecting the critical SARA rate of decline. If the goal of a monitoring program is to detect the critical SARA rate of decline with 80% power, my results indicate that the mean-count and L-AUC methods require as little as six or seven years of escapement monitoring to do so, while the other two methods require between 15 and 17 years. Once again, the mean-count and L-AUC methods are obviously preferable.

Limitations

My use of existing coho visual survey data sets to parameterize run timing dynamics for this study was somewhat problematic due to the large portion of the run timing curve not observed during surveys. The data sets I used had between five and ten survey counts per year, making it necessary to assume that count values between surveys could be described by the estimated mixture model. Given the available data however, there was no better alternative. An additional complication arises from variable survey frequency among data sets. Presumably, data sets with only five count values would be more likely to miss bimodal run timing distributions than data sets with ten count values. Fortunately, the data sets used to parameterize run timing tended to have high survey frequencies. Only five out of 33 data sets (15%) were limited to five count values per year, while 18 data sets had eight or more count values per year (55%). Finally, the time series of visual survey data sets I used to parameterize interannual variability in run timing was restricted to three years for each stream. The use of a time series of daily coho migration past an enumeration fence may have addressed some of these concerns. Daily fence counts would have allowed for the entire run to be observed, thus standardizing the proportion of the run observed among years. Furthermore, some coho enumeration fences have been operated over several consecutive years, which would have increased the length of the time series. However, the tendency of fish to hold in pools below fences and wait for environmental cues before migrating past the fence could potentially increase the number of modes in run timing distributions above what would occur in an unfenced stream. If this was the case, the use of fence data to parameterise run timing dynamics could overestimate interannual variability because visual surveys are usually conducted on streams without enumeration fences. Ideally, a time series of visual survey data sets from an unfenced stream with almost daily surveys should be used to parameterize run timing dynamics. However, such data sets are rare.

The small number of counts within each data set restricted the biological information on run timing dynamics that could be extracted from the maximum likelihood models. While the assumptions of constant survey life and proportion parameter values in the mixture run timing model are not expected to be biologically realistic, these assumptions were necessary given the small number of counts in visual survey data sets. Because the primary purpose of fitting a run timing model to the count data was to generate plausible levels of interannual variability in coho salmon run timing dynamics, a biologically realistic model was not required, as long as the selected model adequately represented interannual variability in coho salmon run timing dynamics, which the restricted mixture model appeared to do.

While my simulation modelling approach allowed for a detailed examination of how uncertainty in coho run timing dynamics and visual survey count data affected the relative performance of alternative estimation methods, all models are approximations of the true state of nature, and the results of this study should be interpreted accordingly. Simulation models are a useful tool for selecting survey designs that help achieve specific monitoring objectives; however, empirical studies are necessary to ensure that the results produced by the model are consistent with reality. In order to address this discrepancy for visual surveys of salmon escapements, long-term monitoring programs should be established in which visual survey data are collected from streams, in addition to using an alternative means of estimating annual escapement, such as mark recapture methods or resistivity counters. By comparing escapement trends observed for the absolute escapement time series with those observed for the visual survey time series, the accuracy and power of trend detection for alternative estimation methods could be better assessed.

Extensions

The application of a hierarchical Bayesian model (HBM) to run timing parameter estimation (Adkison and Su 2001, Su et al. 2001) would likely increase the number of parameters that can be estimated for the mixture run timing model. Adkison and Su (2001) found that escapement estimates of Alaskan pink salmon populations obtained using a HBM formulation for the L-AUC method were more reliable than separate L-AUC estimates for each year. The HBM allowed the estimation procedure to borrow information for years with sparse survey data from years with substantial data, which was well suited for Alaskan pink salmon populations displaying high interannual consistency in run timing (Su et al. 2001). While high interannual variability in coho run timing dynamics as a result of precipitation patterns may limit the utility of a HBM that assumes similar run timing parameters among years, there may be regional similarities in run timing dynamics that would allow for a spatially explicit HBM. For example, run timing parameter estimates for neighbouring streams with similar rainfall patterns could potentially improve escapement estimates for each stream within a single year.

The apparent relationship between the timing of precipitation and coho arrival into spawning streams lends itself to the development of coho run timing models that relate fish arrival to environmental variables. Incorporating such models into the simulation framework I have developed would allow for the evaluation of systematic sampling designs that are based on observed environmental conditions. For example, it may be found that surveys conducted immediately after a high intensity rainfall event are more useful than surveys conducted during an extended period of no rain. This type of information could be used to increase the efficiency of extensive visual survey programs by maximizing the amount of information available on interannual run timing dynamics for a given survey frequency.

Another useful extension of the simulation framework would be the expansion of the current single-stock model to a multi-stock model, in which individual populations with co-varying escapement and run timing dynamics are randomly selected for monitoring. The development of a multi-stock model would allow for the evaluation of alternative multi-stock performance measures, as discussed above, as well as explicit consideration of the effects of different spatial and temporal allocation of survey effort on the power of trend detection within a CU. For example, the question of "what is the optimal allocation of effort between the number of streams monitored with a given method and the frequency applied to each stream?" could be examined for individual CUs. This type of information would be useful for the development of WSP monitoring plans aimed at maximizing the power of trend detection while minimizing survey costs.

Another priority for future research is an evaluation of visual survey methods with regard to the fourth requirement of WSP monitoring plans described in Chapter 1, which asserts that all data collected should be relevant to the provision of management advice. In order to meet this requirement, clear management actions should be established to ensure that appropriate management or conservation actions are taken when the status of a CU falls below pre-specified benchmarks. Given that the level of information available for management decisions regarding a CU will be partially dependent on the level of effort allocated to visual surveys, it follows that proposed alternative designs for extensive monitoring programs should be evaluated as part of a larger management system that includes data collection, stock assessment, management actions based on stock assessment results, and the responses of the fishery and stock dynamics to management decisions (de la Mare 1998, Cooke 1999). The evaluation approach presented in the current chapter could be used as one component of such a larger closed-loop policy simulation framework (Walters 1986) that evaluates the ability of alternative visual survey methods to achieve management objectives. The development of a closed-loop simulation procedure would allow for the comparison of management systems relying on relative indices of escapement, as provided by the mean-count method, with those that derive absolute estimates of escapement using AUC methods.

Conclusions

Simulation modelling provides a useful tool for evaluating the ability of alternative visual survey estimation methods and survey designs to detect changes in salmon escapement. The results of my Monte Carlo simulations show that when trend detection is the primary goal of escapement monitoring, as is true for extensive survey programs in the WSP, a simple mean-count method could be applied broadly to monitor rates of change for coho salmon escapements. Despite high interannual variability in coho run timing dynamics and variable observer efficiency, the mean-count method provides a higher level of confidence in trend detection than peak-count, T-AUC and L-AUC methods when survey dates are constant between years. The success of the mean-count method can be attributed to its simple, data-based estimation procedure that requires no assumptions about the shape of the run timing curve or the length of time fish remain in the survey area.

The power of trend detection for the mean-count method increased with increased survey frequency within a year. While the level of effort afforded to visual survey programs will depend on trade-offs made by fisheries managers between program costs and the level of detail required to assess status of fish populations relative to management and conservation goals, applying a mean-count method and maintaining constant survey dates among years will maximize the power of trend detection for a given level of effort.

Tables

Table 1 Alternative run timing models used in model selection analysis. The functions F_N and F_B are defined in equations (1-2). Symbols defined in Table 2.

Normal run timing model

Parameters

(3) $$\hat{\Phi}_{Normal} = \{m, \sigma, s\}$$

Predicted states

(4) $$\hat{A}_d = C_T \mathbf{F_N}(d, m, \sigma) \qquad \text{where, } C_T = \sum_{f=1}^{f_{max}} C_f$$

(5) $$\hat{D}_d = C_T \mathbf{F_N}(d - s, m, \sigma)$$

(6) $$\hat{C}_d = \hat{A}_d - \hat{D}_d$$

Beta run timing model

Parameters

(7) $$\hat{\Phi}_{Beta} = \{\alpha, \beta, s\}$$

Predicted states

(8) $$\hat{A}_d = C_T \mathbf{F_B}\left(\frac{d}{n}, \alpha, \beta\right)$$

(9) $$\hat{D}_d = \begin{cases} C_T \mathbf{F_B}\left(\frac{d-s}{n}, \alpha, \beta\right) & \text{for } d \geq s \\ 0 & \text{for } d < s \end{cases}$$

(6) $$\hat{C}_d = \hat{A}_d - \hat{D}_d$$

Mixture run timing model

Parameters

(10) $$\hat{\Phi}_{Mixture} = \{m'^1, \sigma'^1, k, \sigma'^2\}$$

Predicted states

(11) $$\hat{A}'^1{}_d = z\, C_T \mathbf{F_N}(d, m'^1, \sigma'^1) \qquad \sigma'^1 \geq 2$$

(12) $$\hat{D}'^1{}_d = z\, C_T \mathbf{F_N}(d - s, m'^1, \sigma'^1)$$

(13) $$\hat{A}'^2{}_d = (1 - z) C_T \mathbf{F_N}(d, m'^1 + k, \sigma'^2) \qquad k \geq 0, \sigma'^2 \geq 2$$

(14) $$\hat{D}'^2{}_d = (1 - z) C_T \mathbf{F_N}(d - s, m'^1 + k, \sigma'^2)$$

(15) $$\hat{C}_d = (\hat{A}'^1{}_d + \hat{A}'^2{}_d) - (\hat{D}'^1{}_d + \hat{D}'^2{}_d)$$

Table 2 Definition of symbols used to describe run timing models in Table 1.

Symbol	Definition
Index variables	
f	Survey index ($f = 1, 2, \ldots f_{max}$ surveys)
f_{max}	Total number of surveys conducted on a given stream within a year
d	Day (from annual calendar; January 1 = day 1)
n	Total number of days of fish presence in survey area
Observed data	
C_d	Number of fish counted within a single stream on day d
C_T	Total number of fish counted within a single stream over all surveys in a year
Predicted states	
Normal and beta models	
$\hat{A}_d$	Cumulative number of fish arriving by survey day d
$\hat{D}_d$	Cumulative number of fish departing by survey day d
$\hat{C}_d$	Predicted number of fish counted on survey day d
Mixture models	
$\hat{A}'^{1}{}_{d}$	Cumulative number of fish from first component model arriving by d
$\hat{A}'^{2}{}_{d}$	Cumulative number of fish from second component model arriving by d
$\hat{D}'^{1}{}_{d}$	Cumulative number of fish from first component model departing by d
$\hat{D}'^{2}{}_{d}$	Cumulative number of fish from second component model departing by d
Run timing model parameters	
Normal model	
m	Mean date of arrival (from annual calendar)
σ	Standard deviation of arrival date
s	Stream life (days)
Beta model	
α	Beta shape parameter 1
β	Beta shape parameter 2
Mixture models	
m'^{1}	Mean date of arrival for first component model
k	Number of days that the mean date of arrival for second component model is offset from that of the first component model
σ'^{1}	Standard deviation of arrival for first component model
σ'^{2}	Standard deviation of arrival for second component model
z	Proportion of counted fish belonging to first component model

Table 3 Summary of survey life estimates collected from the literature (Location codes: AK = Alaska, BC = British Columbia, OR = Oregon, WA = Washington State).

Stream	Location	Year	Estimate	Source
Spring Creek	OR	1952	11.5	Perrin and Irvine 1990
Flynn Creek	OR	1966	13.1	Perrin and Irvine 1990
Oregon streams	OR	1980	11.0	Perrin and Irvine 1990
Harris Creek	WA	1980-3	10.0	Perrin and Irvine 1990
Deer Creek	WA	1981	9.2	Perrin and Irvine 1990
Eagle River	BC	1982	12.5	Perrin and Irvine 1990
Salmon River	BC	1982	15.0	Perrin and Irvine 1990
Adams River	BC	1982	10.0	Perrin and Irvine 1990
Coldwater River	BC	1982	12.5	Perrin and Irvine 1990
Keogh River	BC	1985	13.0	Perrin and Irvine 1990
Little Qualicum	BC	1986	13.3	Perrin and Irvine 1990
French Creek	BC	1987	13.3	Irvine et al. 1992
Black Creek	BC	1987	16.6	Irvine et al. 1992; English et al. 1992
Trent River	BC	1987	7.1	Perrin and Irvine 1990
French Creek	BC	1988	16.7	Irvine et al. 1992
Black Creek	BC	1988	8.0	Irvine et al. 1992; English et al. 1992
Trent River	BC	1988	9.6	Perrin and Irvine 1990
French Creek	BC	1989	15.5	Irvine et al. 1992; English et al. 1992
Black Creek	BC	1989	15.0	Irvine et al. 1992; English et al. 1992
Chase River	BC	1989	16.3	Manske and Schwarz 2000
French Creek	BC	1990	20.3	English et al. 1992
Black Creek	BC	1990	15.0	English et al. 1992
Chase River	BC	1990	10.4	Manske and Schwarz 2000
All streams, OR	OR	-	11.3	Perrin and Irvine 1990
Clear Creek	AK	1996	13.8	Hetrick and Nemeth 2003
Mean			12.8	

Table 4 Results of model selection using Akaike's Information Criterion (AIC) to select the "best" of five candidate run timing models for describing coho salmon visual foot survey data sets.

Candidate run timing model	Number of data sets for which model was the "best-fit"
Normal	1 (3%)
Beta	7 (21%)
Mixture ($z = 0.3$)	10 (31%)
Mixture ($z = 0.5$)	7 (21%)
Mixture ($z = 0.7$)	8 (24%)

Table 5 Simulation parameters for the baseline scenario. The symbol "++" indicates that sensitivity analyses were conducted on the parameter value (see Tables 6 – 7).

Parameter	Value	Description
True population parameters		
t_{max}	10 $^{++}$	Number of years in time series
E_0	500	Initial escapement in year 0
p	- 40 % $^{++}$	Percent change in escapement over t_{max} years
r	- 0.057 $^{++}$	Annual rate of population growth associated with p
m	301.8	Stream-specific mean date of arrival (annual days)
k	12.4	Stream-specific distance parameter
σ'^1	4.7	Stream-specific standard deviation of arrival timing for first curve in mixture model
σ'^2	5.7	Stream-specific standard deviation of arrival timing for second curve in mixture model
τ_m	7.1 $^{++}$	Standard deviation of year-specific random effect for m
τ_k	8.1 $^{++}$	Standard deviation of year-specific random effect for k
τ_{σ_1}	2.3 $^{++}$	Standard deviation of year-specific random effect for σ_1
τ_{σ_2}	3.0 $^{++}$	Standard deviation of year-specific random effect for σ_2
s	12.8	Survey life for fish arriving on the median arrival date
z	0.3	Proportion parameter for the mixture run timing model
v	0.865 $^{++}$	Average "true" observer efficiency
Survey parameters		
f_{max}	(1, 2, 3, ... 8) $^{++}$	Number of surveys per year
$\hat{v}$	0.865	Estimated observer efficiency
CV($\hat{s}$)	0.2 $^{++}$	Co-efficient of variation of survey life estimate
μ_s	12.8	Mean of prior distribution on s
Trend detection parameters		
p^*	- 30%	Critical percent change in escapement over t_{max} years
r^*	- 0.04	Critical annual rate of population growth associated with p^*

Table 6 Alternative scenarios about "true" population dynamics tested in sensitivity analyses.

Variable	Scenarios	Values
Variability in run timing	None	$\tau = (0, 0, 0, 0)$
$\tau = \{\tau_m, \tau_k, \tau_{\sigma 1}, \tau_{\sigma 2}\}$	Low	$\tau = (2.5, 1.4, 2.6, 2.9)$
	Baseline	$\tau = (6.7, 5.6, 2.9, 2.4)$
	High	$\tau = (9.4, 8.9, 4.0, 1.0)$
Variability in observer efficiency	None	see text
	Low	$v = 0.76$
	Baseline	$v = 0.86$
	High	$v = 0.96$
Percent population change over 10 years		$p = (0, -2, -4, \ldots -60\ \%)$

Table 7 Alternative survey design scenarios tested in sensitivity analyses. The function U(*g*,*h*) denotes the random generation of a number from a uniform distribution with a lower bound of *g* and an upper bound of *h*.

Variable	Scenarios	Values
Error in survey life estimate	None	$CV(\hat{s}) = 0$
	Low	$CV(\hat{s}) = 0.1$
	Baseline	$CV(\hat{s}) = 0.2$
	High	$CV(\hat{s}) = 0.3$
Survey spacing	Even	
	Baseline	See text
	Random	
Survey frequency	Constant	$f_{max} = (1, 2, 3 \ldots 8)$
	High variability	$f_{max} = \mathbf{U}(3, 8)$
	Moderate variability	$f_{max} = \mathbf{U}(4, 7)$
	Low variability / low frequency	$f_{max} = \mathbf{U}(3, 5)$
	Low variability / high frequency	$f_{max} = \mathbf{U}(3, 6)$
Number of years monitored		$t_{max} = (5, 6, 7, \ldots 20)$

Table 8 Minimum number of years required to detect a 30% decline in escapement over 10 years with 80% power (top) and minimum detectable "true" rate of decline when 80% power is required for trend detection analysis (bottom) in the baseline scenario.

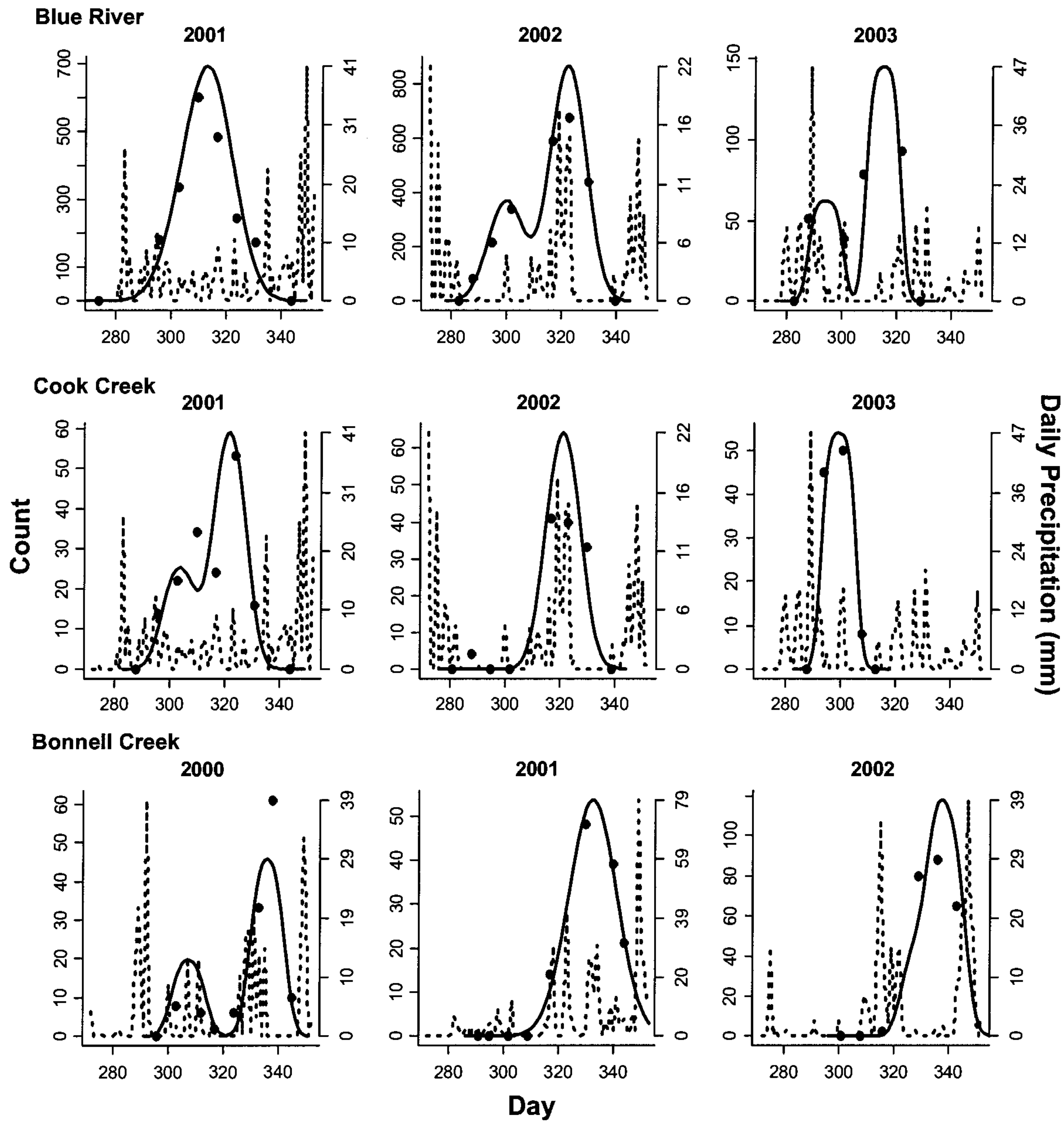
	Survey frequency							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	Minimum number of years							
Peak	> 20	20	19	19	19	18	16	16
Mean	-	20	16	13	11	8	7	6
Trapezoidal	-	-	19	18	17	17	16	16
Likelihood	-	-	-	-	12	8	7	7
	Minimum detectable "true" percent of decline (% over 10 years)							
Peak	> 60	> 60	58	54	54	52	50	50
Mean	> 60	> 60	50	44	42	38	36	36
Trapezoidal	-	-	56	52	50	50	50	50
Likelihood	-	-	-	-	44	40	36	36

Table 9 Summary of monitoring performance, data requirements, and recommended features for each estimation method.

Method	Max. power to detect $r \leq -0.04$[a]	Absolute or relative escapement estimates	Min. No. of Surveys per Year	Annual estimates of OE[b] and SL[c]	Measure of uncertainty in estimated escapement	Recommended features
Peak-count	56 %	Relative	1	Not required	No (unless replicate surveys conducted)	- Not suitable for stocks with high interannual variability in run timing - Maintain a constant number of days between surveys each year - Cluster surveys around historical mean peak date when ≤ 5 surveys per year conducted
Mean-count	99 %	Relative	2	Not required	No (unless replicate surveys conducted)	- Maintain a constant number of days between surveys each year - Cluster surveys around historical mean peak date when ≤ 5 surveys per year conducted
T-AUC	60 %	Absolute	3	Required	No (unless replicate surveys conducted)	- Use year- and stream-specific estimates of SL and OE - Maintain a constant number of days between surveys each year - Cluster surveys around historical mean peak date when ≤ 5 surveys per year conducted
L-AUC	96 %	Absolute	5 (or more)	Depends on number of surveys used	Yes (but dependent on prior information about OE and SL)	- Use estimates of SL for prior information - Use stream- and year-specific estimates of OE - Maintain a constant number of days between surveys each year - Cluster surveys around historical mean peak date when 5 surveys per year conducted

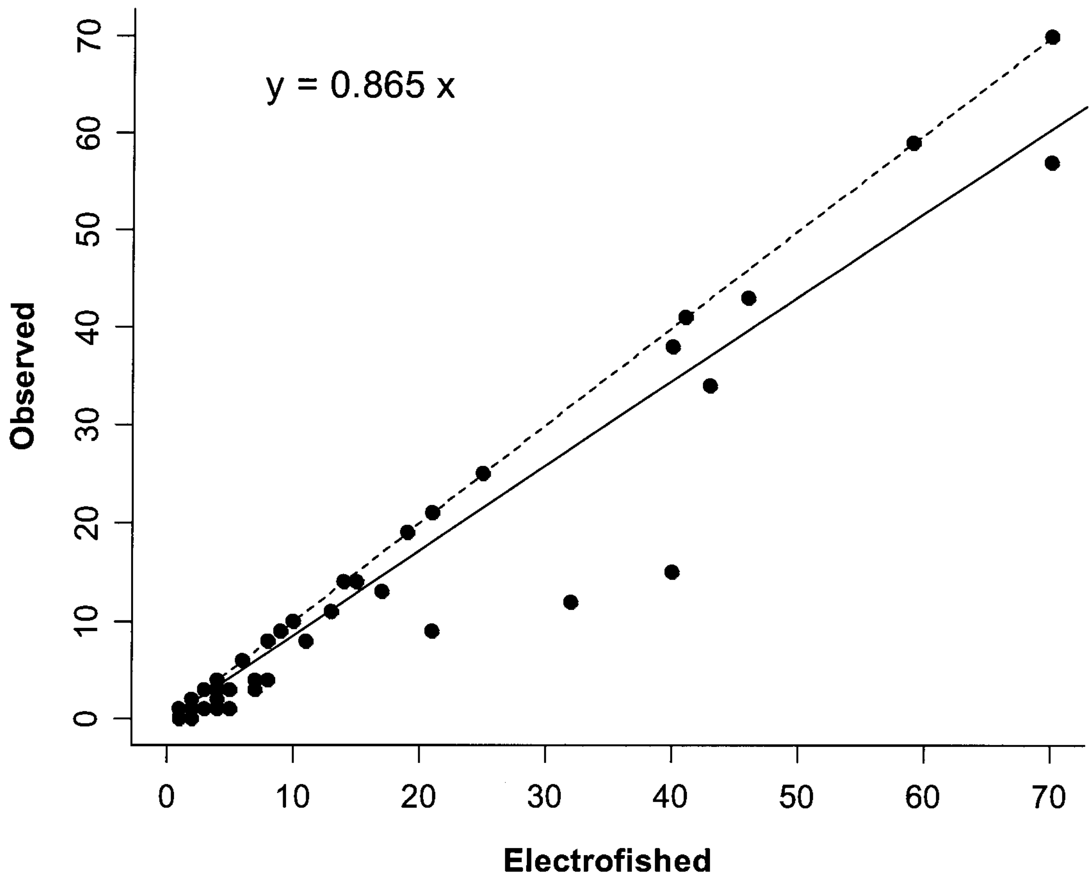
[a] At eight survey counts per year when the true r value is -0.057 over 10 years [b] OE = observer efficiency [c] SL = survey life

Figures



Data sources: Observed spawner abundance data provided by Fisheries and Oceans Canada, Interior Fraser River and South Coast stock assessment divisions. Daily precipitation data was taken from Environment Canada's online database (http://climate.weatheroffice.ec.gc.ca/climateData/canada_e.html)

Figure 1 Daily precipitation levels (dashed lines) plus observed (dots) and predicted (solid lines) spawner abundance from visual survey counts for coho salmon visual survey data sets from Blue River and Cook Creek, North Thompson Watershed, BC and Bonnell Creek, Vancouver Island, BC.



Data source: Jim Irvine, Fisheries and Oceans Canada (unpublished data) and Irvine et al. 1992

Figure 2 Predicted relationship (solid line) between abundance estimates from electrofishing ("true" abundance) and the number of coho estimated during visual foot surveys for Black Creek, Vancouver Island between 1987 and 1993. Multiple surveys (black dots) were done in each year. For comparison, the 1:1 line is also shown (dashed line).

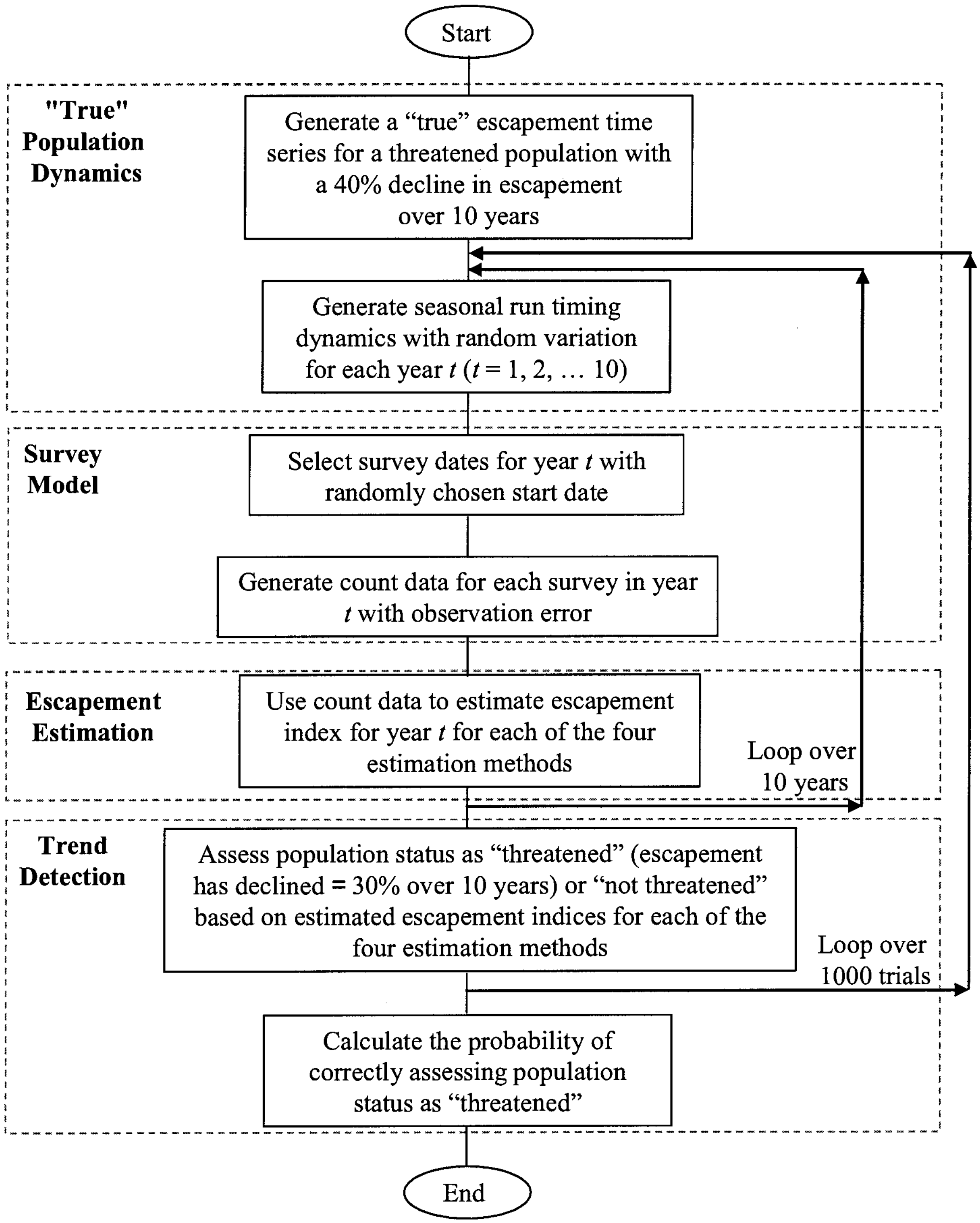


Figure 3 Flow diagram for Monte Carlo simulation procedure used to estimate the statistical power of alternative estimation methods to detect a 30% decline in escapement over 10 years ($r \leq -0.04$).

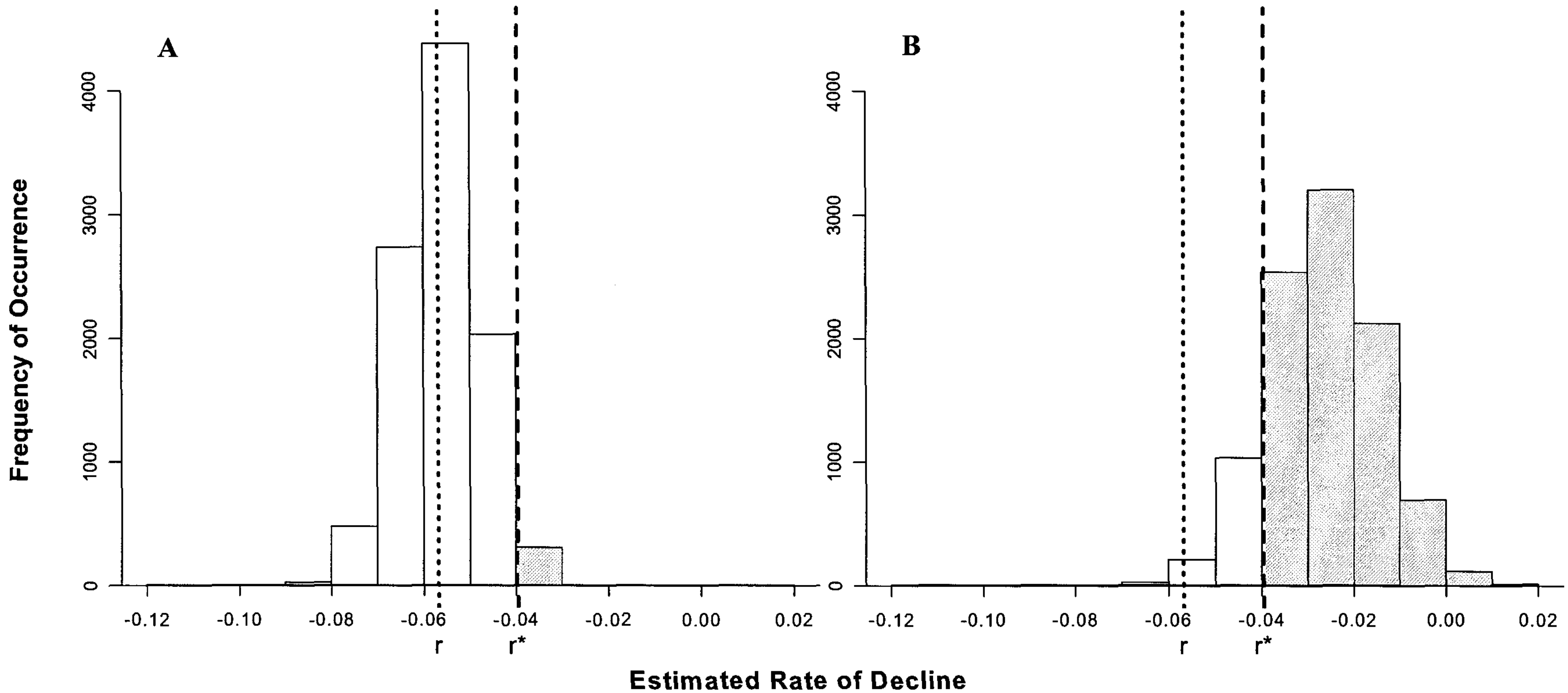


Figure 4 Two examples of approximate marginal posterior distributions for r constructed using a sampling-importance-resampling (SIR) algorithm (Appendix C). The dashed line shows the critical r value associated with a 30% decline in escapement over 10 years, r^*, and the dotted line shows the true r value used in the baseline scenario which corresponds with a 40% decline over 10 years. The probability assigned to the null hypothesis $r > r^*$ [$P(r > r^*)$] is the proportion of simulations for which the predicted rate of decline is less steep than r^* (shaded regions on plots A and B). Plot A represents a trial for which trend detection is "successful" [$P(r > r^*) = 0.03$], while plot B represents a trial for which trend detection is "not successful" [$P(r > r^*) = 0.86$].

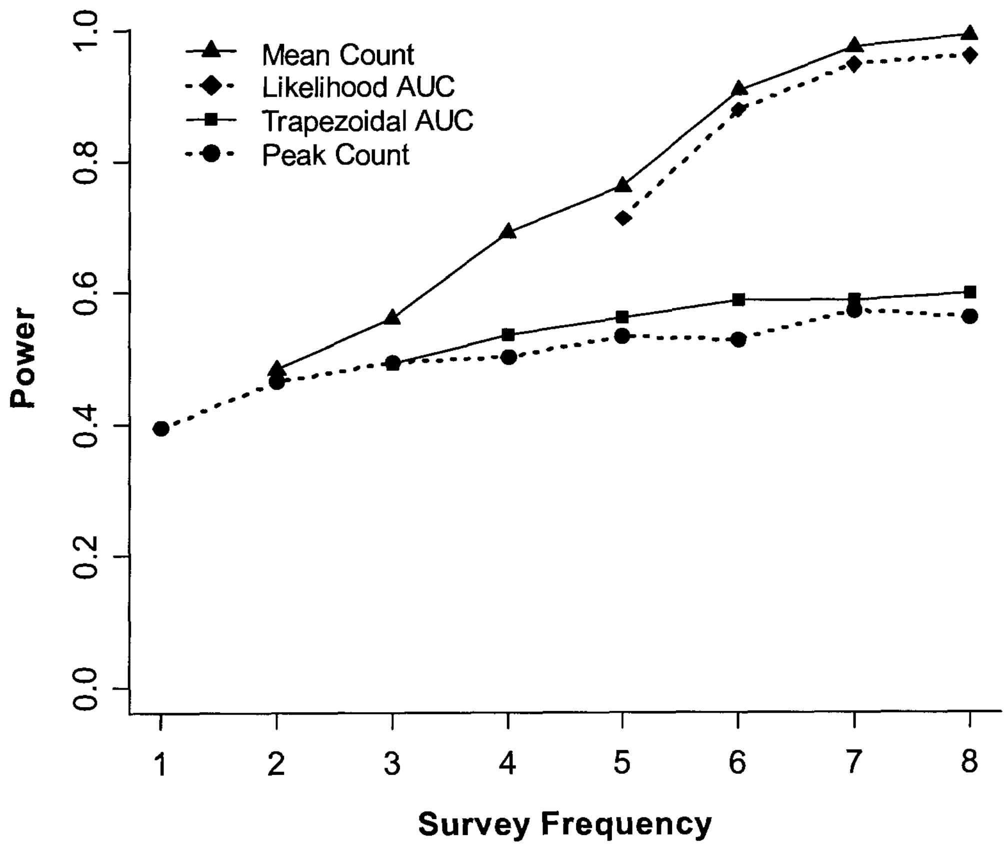


Figure 5 **Power to detect $r \leq -0.04$ as a function of survey frequency for each of the four estimation methods in the baseline scenario. The mean-count, trapezoidal AUC, and likelihood AUC methods require minimum survey frequencies of 2, 3, and 5 surveys per year, respectively.**

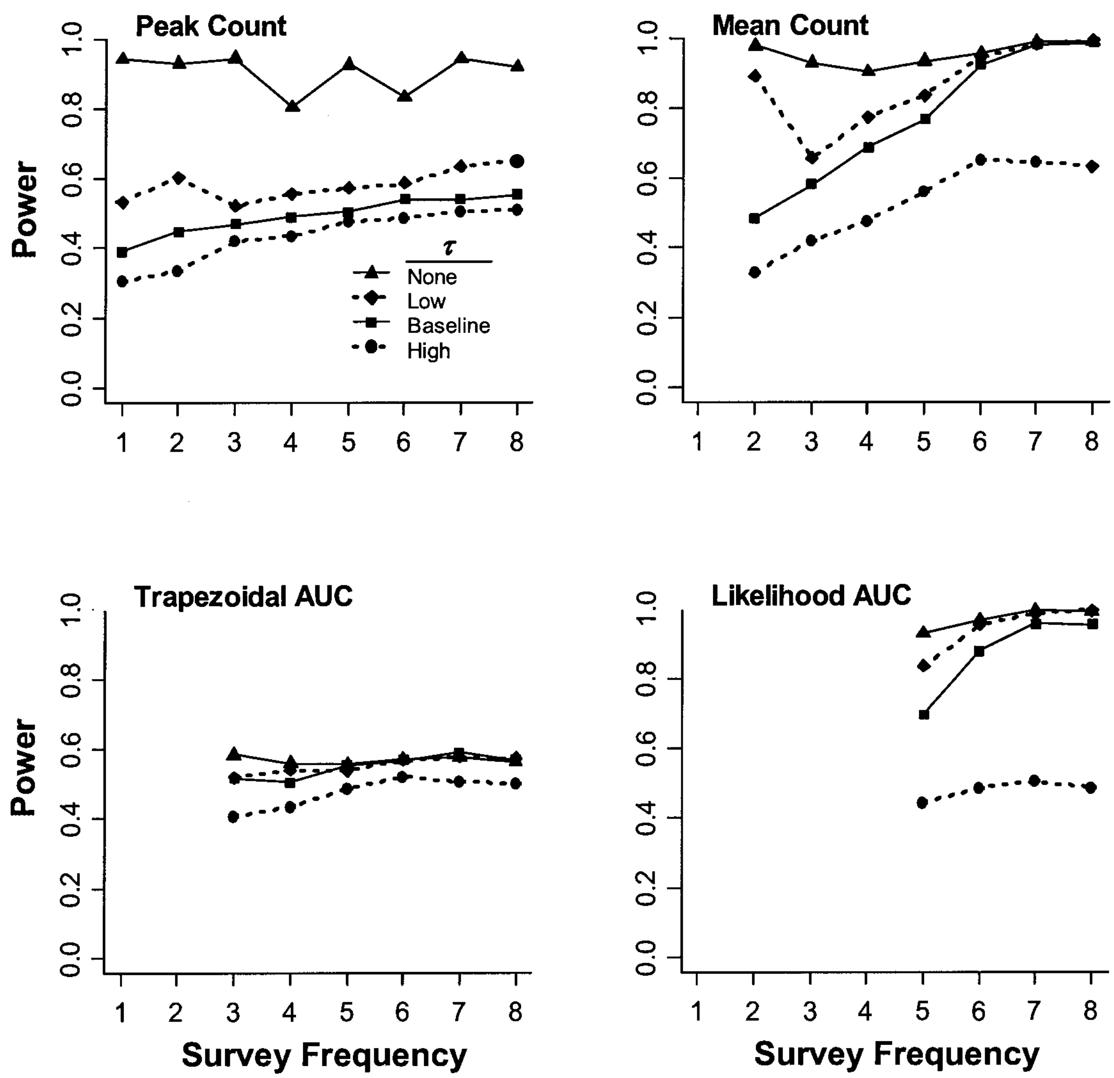


Figure 6 Sensitivity of power to detect $r \leq -0.04$ to the level of interannual variability in run timing dynamics, τ (Table 6), for each of the four estimation methods.

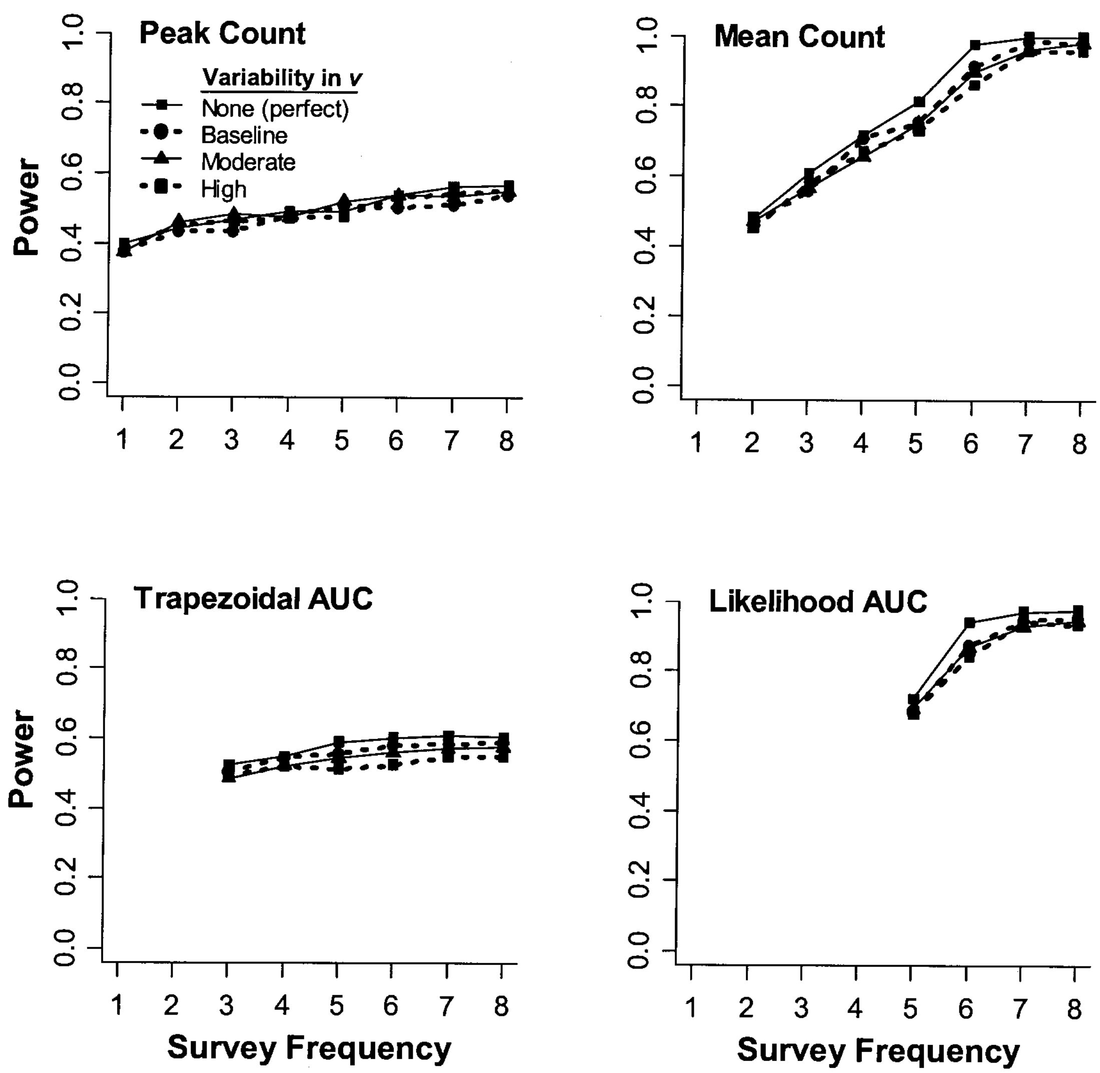


Figure 7 Sensitivity of power to detect $r \leq -0.04$ to the level of among-survey variability in observer efficiency, as determined by the value of v (Table 6), for each of the four estimation methods.

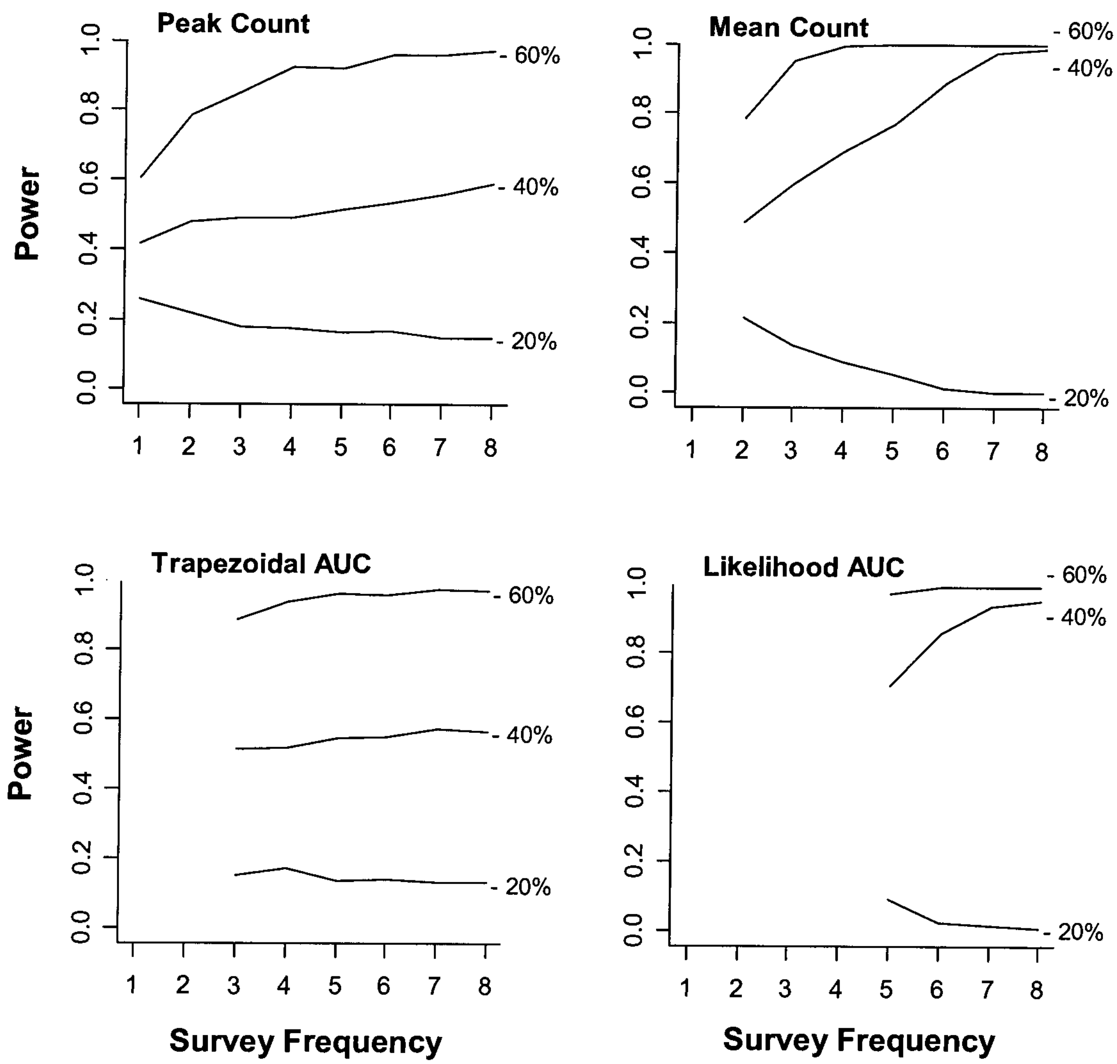


Figure 8 Sensitivity of power to detect $r \leq -0.04$ to the "true" percent change in escapement (-20%, -40%, or -60% decline over 10 years) for each of the four estimation methods.

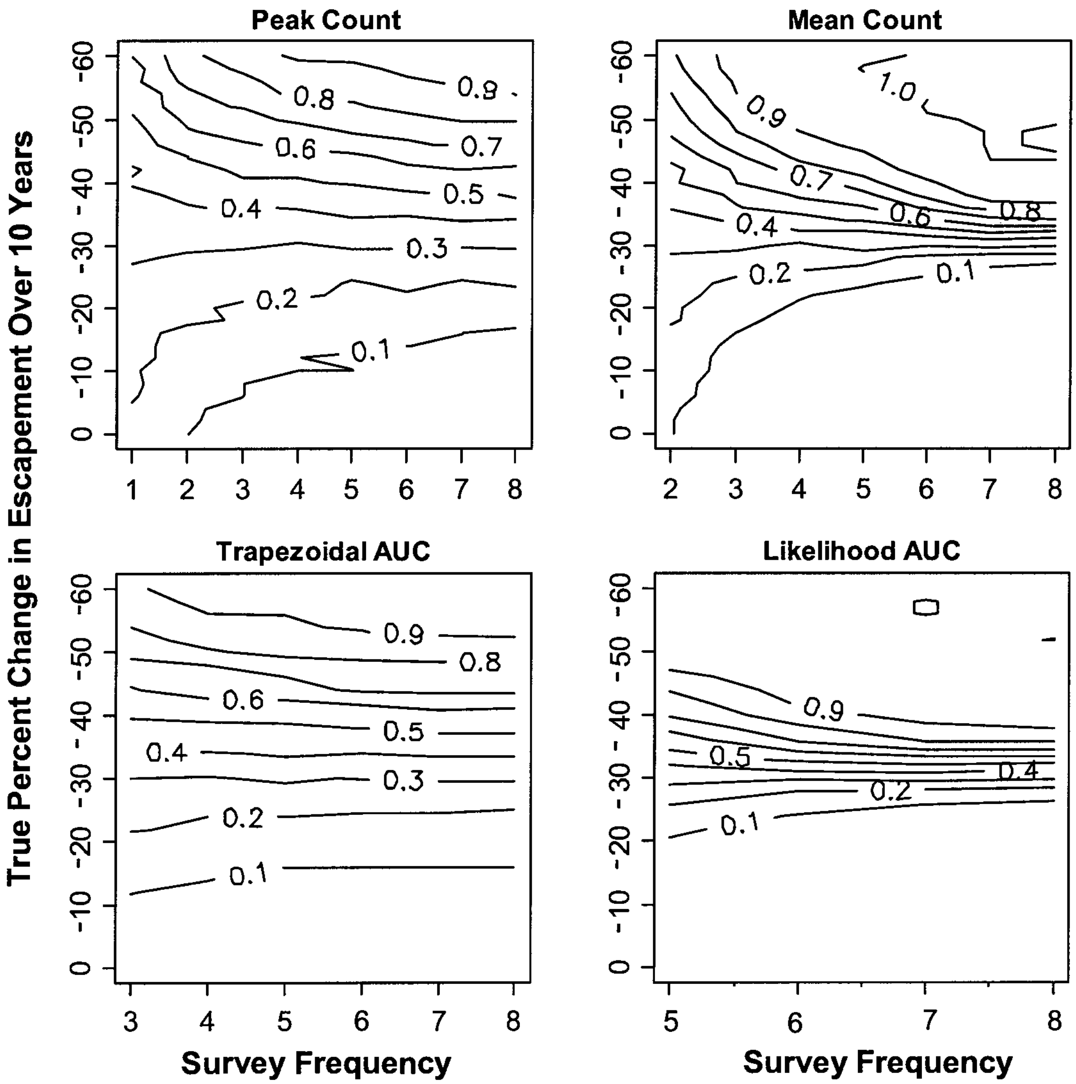


Figure 9 **Power (contour lines) to detect $r \leq -0.04$ (30% decline in escapement over 10 years; $p \leq -30\%$) as a function of the "true" percent change in escapement and survey frequency within each year for each estimation method. Note that x-axes have different scales for each estimation method.**

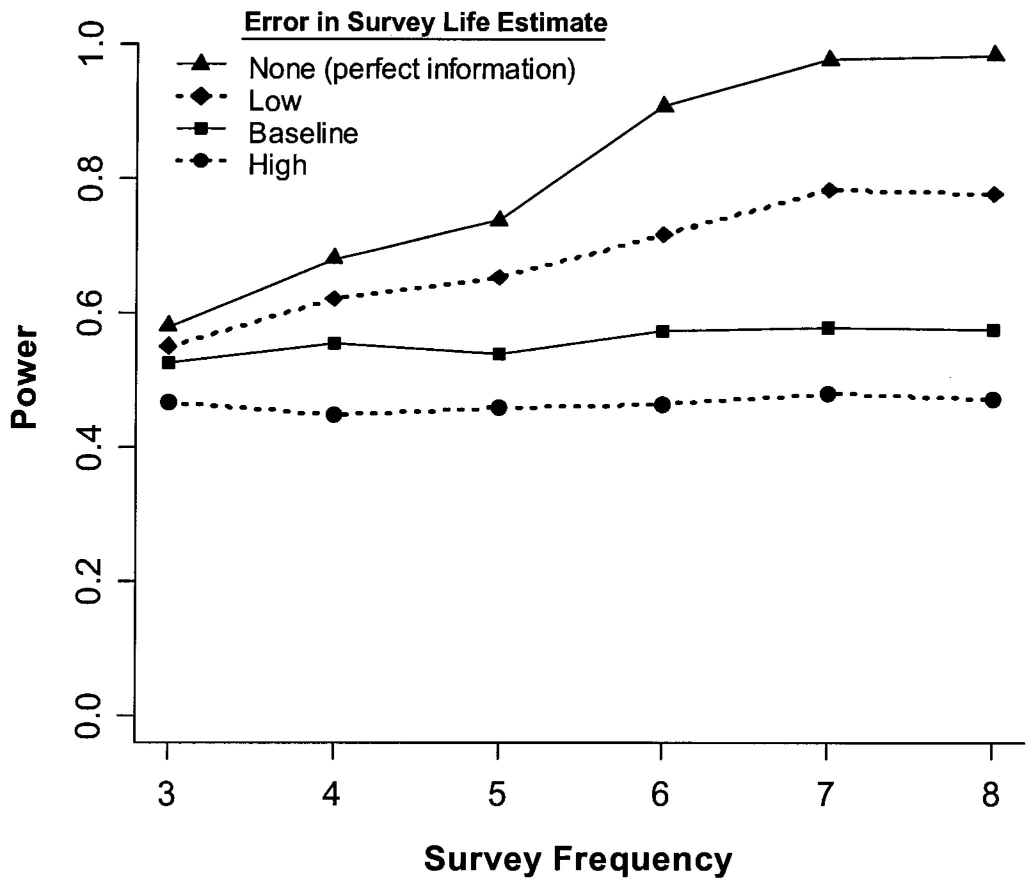


Figure 10 Sensitivity of power to detect $r \leq -0.04$ to the level of error in annual survey life estimates for the T-AUC method (none: CV($\hat{s}$) = 0 , low: CV($\hat{s}$) = 0.1 , baseline: CV($\hat{s}$) = 0.2 , high: CV($\hat{s}$) = 0.3).

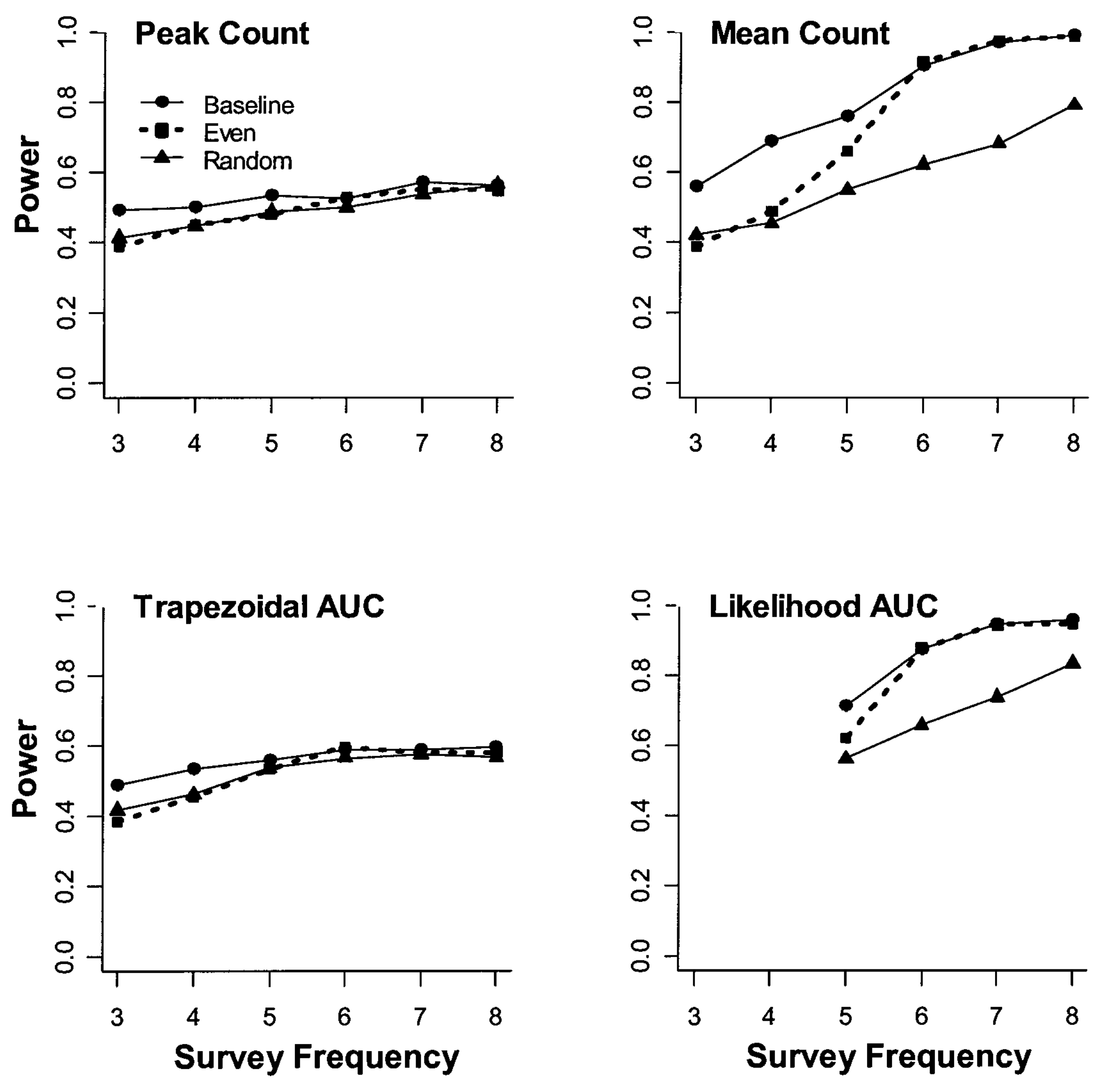


Figure 11 **Sensitivity of power to detect $r \leq -0.04$ to the spacing of survey dates within a year (baseline = dates selected using equations 22-24, even = dates evenly spaced within a year, random = dates randomly selected from one-week strata) for each of the four estimation methods.**

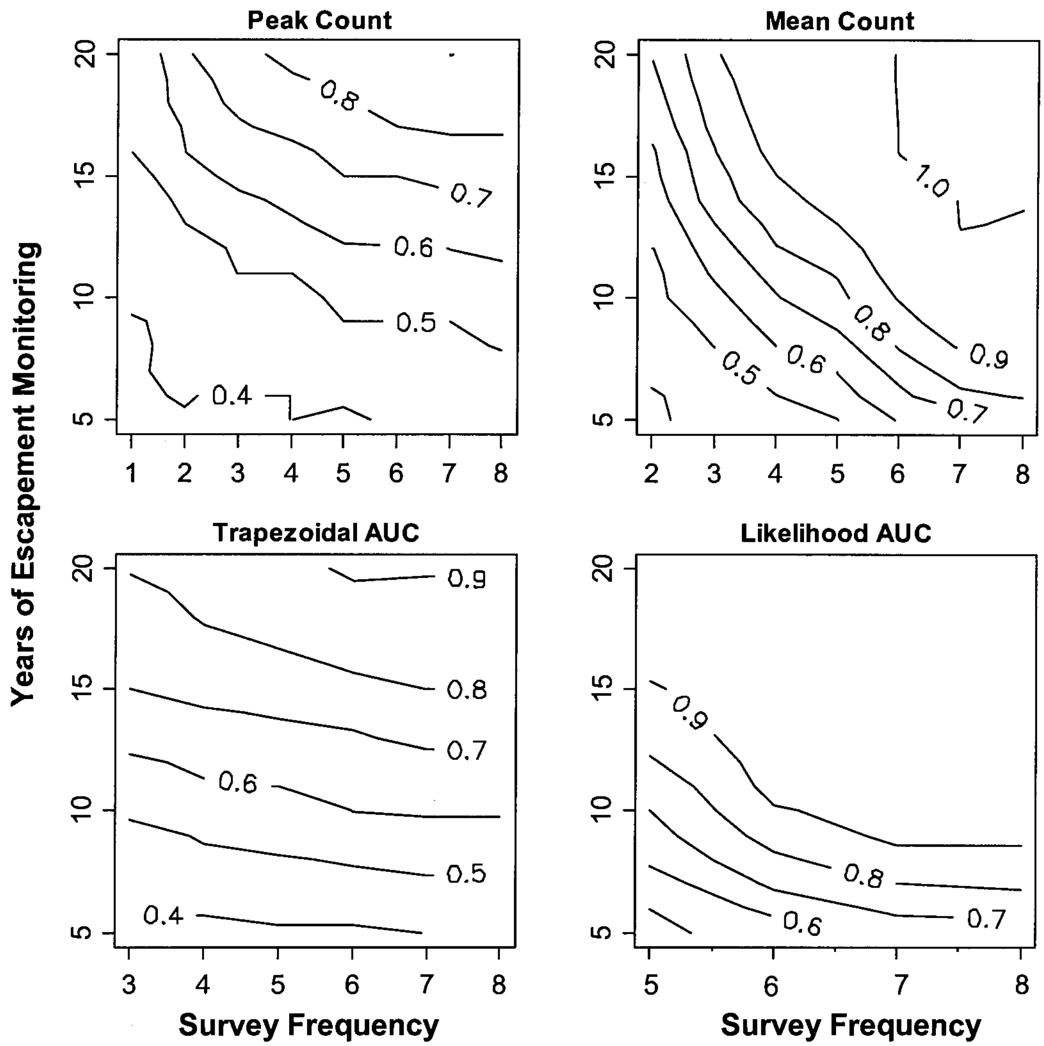
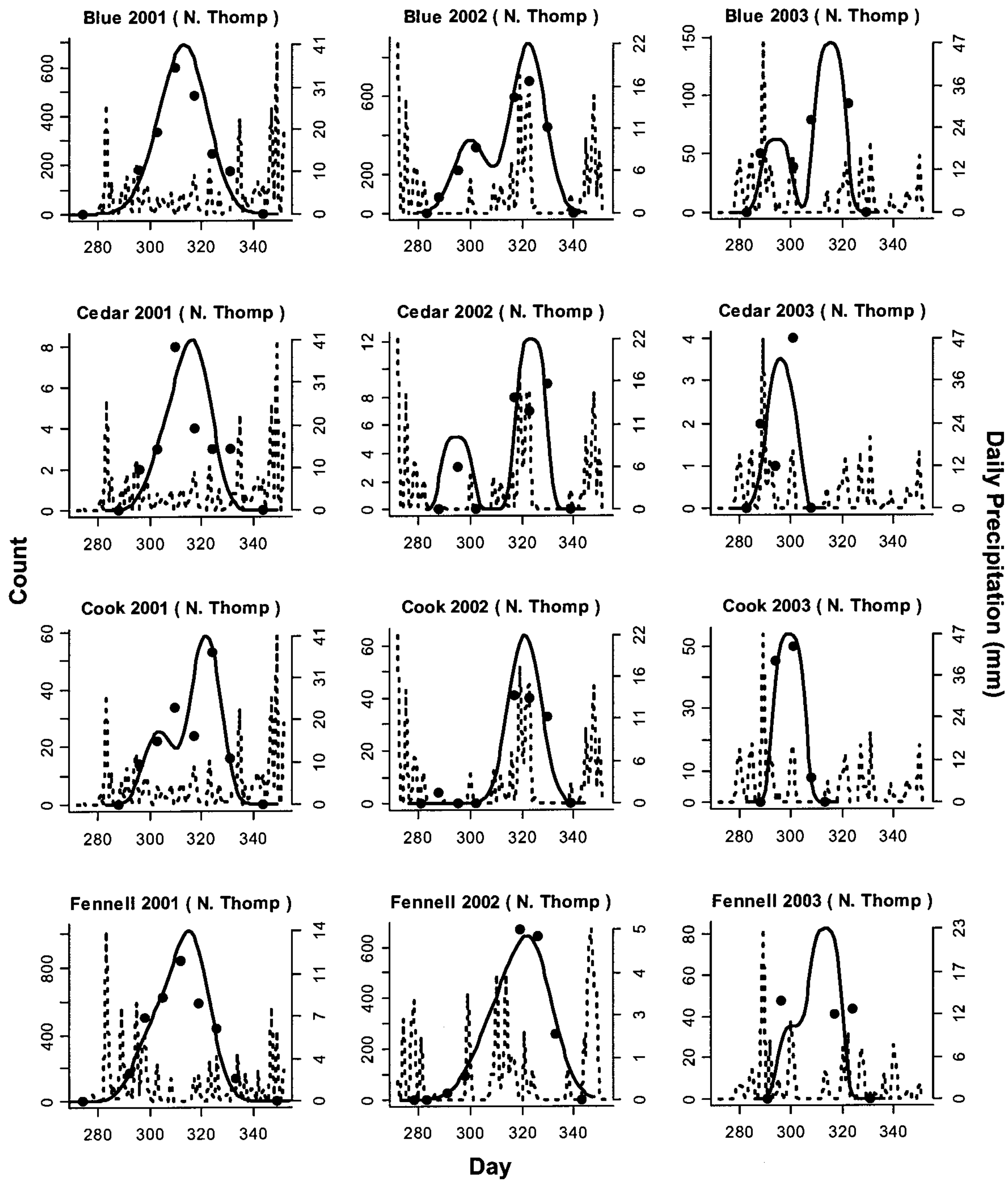
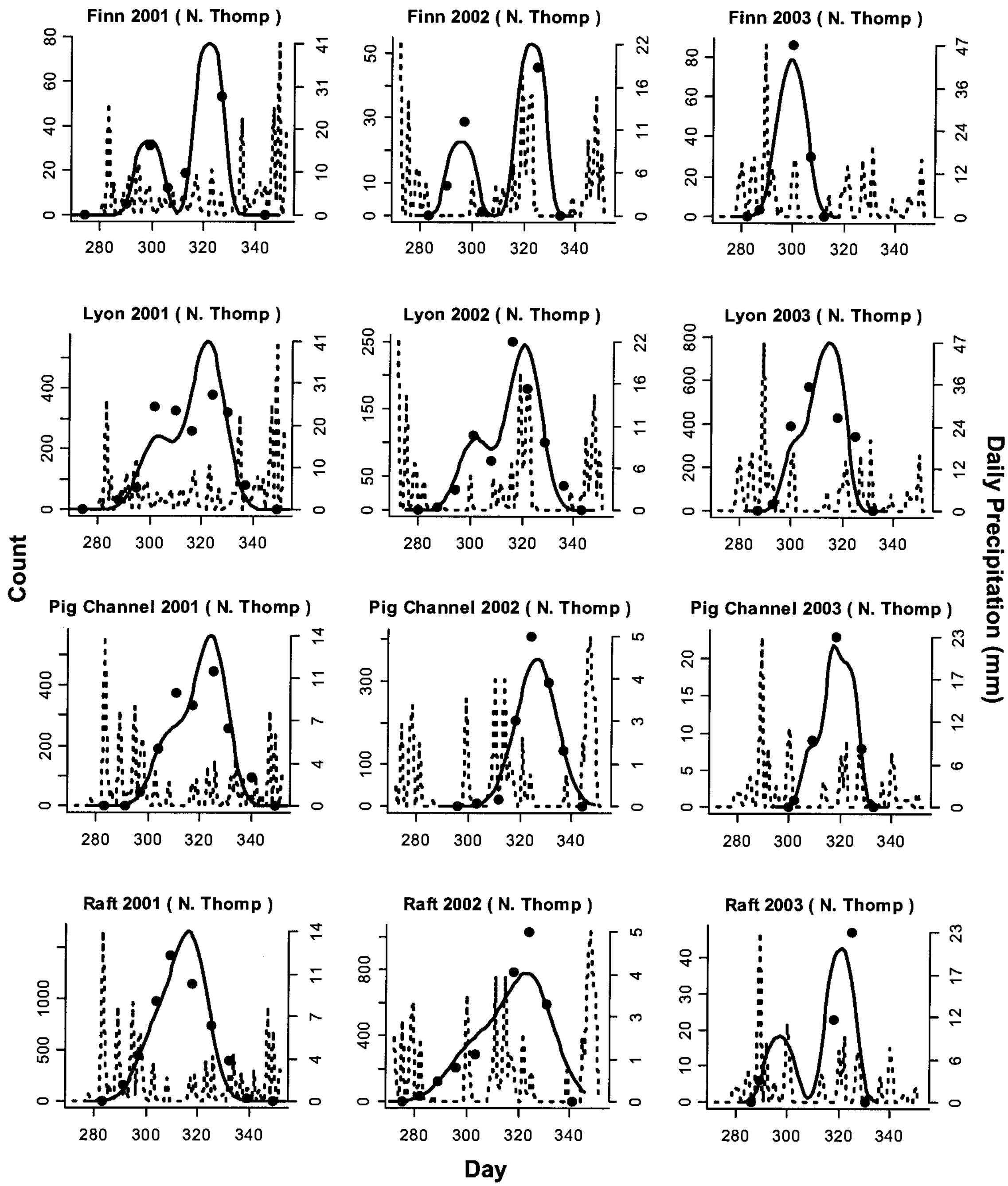


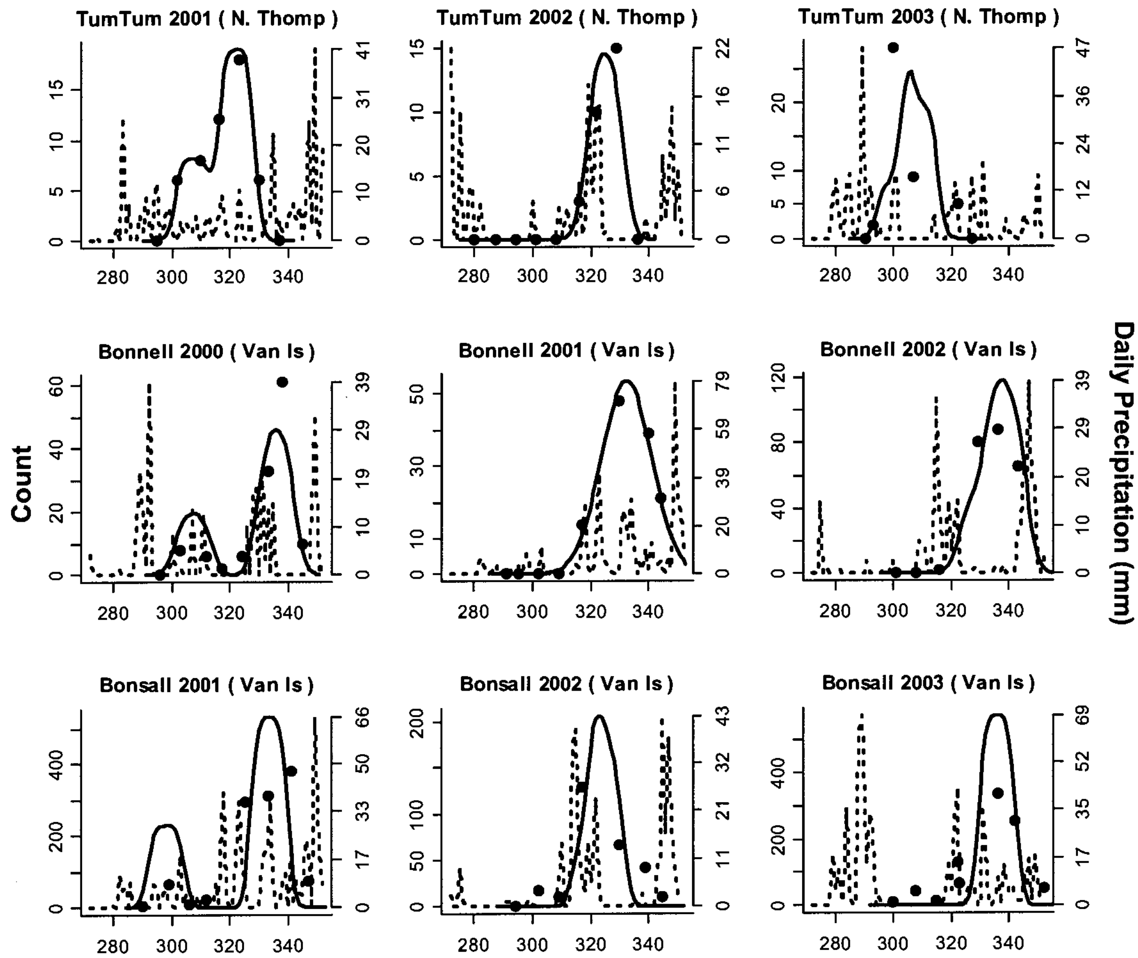
Figure 12 **Power (contour lines) to detect $r \leq -0.04$ as a function of the number of years of escapement monitoring and survey frequency within each year for each of the four estimation methods. Note that x-axes have different scales for each estimation method.**

Appendix A

Figure A-1. Thirty-three data sets of daily precipitation levels (dashed lines) plus observed (dots) and predicted (solid lines) spawner abundance from visual surveys of tributary streams to the North Thompson River in interior British Columbia (N. Thomp) and coastal streams on the east coast Vancouver Island, British Columbia (Van. Is).




Finn 2001 (N. Thomp)
Finn 2002 (N. Thomp)
Finn 2003 (N. Thomp)
Lyon 2001 (N. Thomp)
Lyon 2002 (N. Thomp)
Lyon 2003 (N. Thomp)
Pig Channel 2001 (N. Thomp)
Pig Channel 2002 (N. Thomp)
Pig Channel 2003 (N. Thomp)
Raft 2001 (N. Thomp)
Raft 2002 (N. Thomp)
Raft 2003 (N. Thomp)
Count
Daily Precipitation (mm)
Day


TumTum 2001 (N. Thomp)
TumTum 2002 (N. Thomp)
TumTum 2003 (N. Thomp)
Bonnell 2000 (Van Is)
Bonnell 2001 (Van Is)
Bonnell 2002 (Van Is)
Bonsall 2001 (Van Is)
Bonsall 2002 (Van Is)
Bonsall 2003 (Van Is)
Count
Daily Precipitation (mm)
280 300 320 340

Appendix B

For a specified percent change p in escapement E over t years,

$$p = \frac{E_t - E_0}{E_0} * 100, \tag{B1}$$

escapement in year t can be written as

$$E_t = E_0 (1 - \frac{-p}{100}). \tag{B2}$$

To calculate the intrinsic rate of population growth, r, required to produce an exponential escapement time series,

$$E_t = E_0 e^{rt}, \tag{19}$$

with p percent decline, equation 19 was rearranged to solved for r

$$r = \frac{\log_e(E_t) - \log_e(E_0)}{t} \tag{B3}$$

$$r = \log_e \left(\frac{E_t}{E_0} \right)^{-t} \tag{B4}$$

and B2 was substituted into B4 as follows:

$$r = \log_e \left(\frac{E_0 \left(1 - \frac{-p}{100} \right)}{E_0} \right)^{-t} \tag{B5}$$

$$r = \log_e \left(1 - \frac{-p}{100} \right)^{-t}. \tag{B6}$$

Using equation B6, the “true” value of r used in equation 19 to generate a “true” escapement time series with a 40% decline in escapement over 10 years (p = - 40%) was calculated as -0.057. Similarly, the critical SARA rate of decline associated with a 30% decline in escapement over 10 years (p = -30%) that was used in hypothesis testing for the trend detection analysis was calculated as -0.04.

Appendix C

I used the sampling-importance-resampling algorithm (SIR; Rubin 1988) to approximate the marginal posterior distribution of the population growth rate r in the growth model:

$$\log_e(I_t) = \log_e(I_0) + rt \tag{C1}$$

where I_t is the escapement index in year t, and I_0 is the y-intercept, which represents the predicted index value in year 0. The algorithm outlined below is based on the SIR algorithm presented in Rubin (1988) and McAllister et al. (1994). For computational efficiency, I fit a simple linear model to $I\,(I = I_{t=1}, I_{t=2}, \ldots I_{t=t\max})$ to construct a joint prior distribution for (r, I_0). The prior distributions for each variable were assumed normally distributed with means and standard errors equal to those predicted by the linear model.

1) For simulation k ($k = 1,2, \ldots m$; $m = 10{,}000$), randomly draw $(r, I_0)_k$ from the joint prior distribution $p(r, I_0)$

2) Calculate a predicted escapement index, $\hat{X}_t$, by using $(r, I_0)_k$ in the regression equation

$$\ln(\hat{X}_t) = \ln(I_{0k}) + r_k t \tag{C2}$$

3) Calculate the normal log-likelihood of $(r, I_0)_k$, $[\mathrm{L}(r_k, I_{0k} \mid I)]$ proportional to the sums-of-squares between observed and predicted escapement indices over t years

4) Calculate the importance ratio $w(r_k, I_{0k})$ as,

$$w(r_k, I_{0k}) = \frac{L(r_k, I_{0k} | I_t)\, p(r_k, I_{0k})}{h(r_k, I_{0k})} \tag{C3}$$

where, $h(r_k, I_{0k})$ is the importance function, which is a common probability density function that can be used to generate samples for all parameters in the set

(r_k, I_{0k}). For simplicity, I set $h(r_k, I_{0k})$ equal to the prior $p(r_k, I_{0k})$ (*e.g.* McAllister *et al.* 1994), simplifying the importance ratio to,

$$w(r_k, I_{0k}) = L(r_k, I_{0k} | I) \tag{C4}$$

5) Repeat steps 1 to 4 over *m* simulations

6) Randomly draw *m* samples, with replacement, from r_k using probabilities proportional to $w(r_k, I_{0k})$, where $k = (1, 2, \ldots m)$. Because $w(r_k, I_{0k})$ can be used to approximate the joint posterior distribution of (r, I_0) using

$$\hat{P}(r_k, I_{0k} | I) \cong \frac{w(r_k, I_{0k})}{\sum_{k=1}^{m} w(r_k, I_{0k})}, \tag{C5}$$

the generated sample is an approximation of the marginal $p(r \mid I)$.

REFERENCE LIST

Adkison, M.D. and Z. 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.

Adkison, M.D., T.J. Quinn II, and R.J. Small. 2003. Evaluation of the Alaskan harbour seal (Phoca Vitulina) population survey: a simulation study. *Marine Mammal Science* 19: 764 – 790.

Ames, J. and D.E. Phinney. 1977. Puget Sound summer-fall Chinook methodology: escapement goals, run size forecasts, and in-season run size updates. Washington Department of Fisheries Technical Report No. 29.

Bevan, D.E. 1961. Variability in aerial counts of spawning salmon. *Journal of the Fisheries Research Board of Canada* 18: 337-348.

Bue, B.G., M. Fried, S. Sharr, D.G. Sharp, J.A. Wilcock, and H.G. Geiger. 1998. Estimating salmon escapement using area-under-the-curve, aerial observer efficiency, and stream-life estimates: the Prince William Sound pink salmon example. *North Pacific Anadromous Fish Commission Bulletin* 1: 240-250.

Burgner, R.L. 1991. Life history of sockeye salmon (*Oncorhynchus nerka*). *In* Pacific Salmon Life Histories. *Edited by* C. Groot and L. Margolis. pp. 1-118. UBC Press, Vancouver.

Burnham, K.P. and D.R. Anderson. 2002. *Model Selection and Multimodal Inference: A Practical Information-Theoretical Approach*, 2nd edition. Springer-Verlag, New York.

Cooke, J.G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithims. *ICES Journal of Marine Science* 56: 797-810.

Cousens, N.B.F., G.A. Thomas, C.G. Swann, and M.C. Healey. 1982. A review of salmon escapement estimation techniques. *Canadian Technical Report of Fisheries and Aquatic Sciences* 1108.

de la Mare, W.K. 1998. Tidier fisheries management requires a new MOP (management orientated paradigm). *Reviews in Fish Biology and Fisheries* 8: 349-356.

DFO (Fisheries and Oceans Canada). 2005. Canada's policy for conservation of wild Pacific salmon. Available at: http://www-comm.pac.dfo-mpo.gc.ca/ publications/wsp/default_e.htm (May 2006). Vancouver, British Columbia.

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.

Fraser, F.J., E.A. Perry, and D.T. Lightly. 1983. Big Qualicum River salmon development project. Volume 1: a biological assessment of 1959 - 1972. *Canadian Technical Report of Fisheries and Aquatic Sciences* 1189: 198 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.

Geiger, H.J. and S. McPherson, editors. 2004. Stock status and escapement goals for salmon stocks in Southeast Alaska. *Alaska Department of Fish and Game, Divisions of Sport Fishing and Commercial Fisheries, Special Publication 04-02.*

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2004. *Bayesian Data Analysis*, 2nd edition. Chapman and Hall, New York.

Hansen, L.P. and B. Jonsonn 1991. Evidence of a genetic component in the seasonal return pattern of Atlantic salmon, *Salmo salar*. *Journal of Fish Biology* 38: 25-1258.

Heard, W.R. 1991. Life history of pink salmon (*Oncorhynchus gorbuscha*). *In* Pacific Salmon Life Histories. *Edited by* C. Groot and L. Margolis. pp. 119-230. UBC Press, Vancouver.

Hetrick, N.J. and M.J. Nemeth. 2003. Survey of coho salmon runs on the Pacific coast of the Alaska Peninsula and Becharof National Wildlife Refuges, 1994 with estimates of escapement for two small streams in 1995 and 1996. *Alaska Fisheries Technical Report* No. 63.

Hilborn, R., B.G. Bue, and S. Sharr. 1999. Estimating spawning escapements from periodic counts: a comparison of methods. *Canadian Journal of Fisheries and Aquatic Sciences* 56: 888-896.

Hilborn, R. and M. Mangel. 1997. *The Ecological Detective: Confronting Models with Data*. Princeton University Press, New Jersey.

Hill, R.A. 1997. Optimizing aerial count frequency for the area-under-the-curve method of estimating escapement. *North American Journal of Fisheries Management* 14: 461 - 466.

Hodgson, S. and T.P. Quinn. 2002. The timing of adult sockeye salmon migration into freshwater: adaptations by populations to prevailing thermal regimes. *Canadian Journal of Zoology* 80: 542-555.

Holtby, L.B., G.F. Hartman, and J.C. Scrivener. 1984. Stream indexing from the perspective of the Carnation Creek experience, p. 87-111. *In:* P.E.K. Symons and M. Waldichuck (eds.) Proceedings of the Workshop on Stream Indexing for salmon escapement Estimation, West Vancouver, BC, 2-3 February, 1984. *Canadian Technical Report of Fisheries and Aquatic Sciences* 1326.

Irvine, J.R., R.C. Bocking, K.K. English, and M. Labelle. 1992. Estimating coho salmon (*Onchorhynchus kisutch*) spawning escapements by conducting visual surveys in areas selected using stratified index sampling designs. *Canadian Journal of Fisheries and Aquatic Sciences* 49: 1972-1981.

Jones, E.L., III, T.J. Quinn, and B.W. VanAlen. 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.

Keefer, M.L., C.A. Perry, M.A. Jepson, K.R. Tolotti, T.C. Bjornn, and L.C. Stuehrenberg. 2004. Stock-specific migration timing of adult spring-summer chinook salmon in the Columbia River Basin. *North American Journal of Fisheries Management* 24: 1145-1162.

Korman, J., R.N.M. Ahrens, P.S. Higgins, C.J. Walters. 2002. Effects of observer efficiency, arrival timing, and survey life on estimates of escapement for steelhead trout (*Oncorhynchus mykiss*) derived from repeat mark-recapture experiments. *Canadian Journal of Fisheries and Aquatic Science* 59: 1116-1131.

Korman, J. and P.S. Higgins. 1997. Utility of escapement time series data for monitoring the response of salmon populations to habitat alteration. *Canadian Journal of Fisheries and Aquatic Sciences* 54: 2058 - 2067.

Manske, M. and C.J. Schwarz. 2000. Estimates of stream residence time and escapement based on capture-recapture data. *Canadian Journal of Fisheries and Aquatic Science* 57: 241-246.

Maxwell, D. and S. Jennings. 2005. Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. *Journal of Applied Ecology* 42: 25 – 37.

McAllister, M.K., E.K. Pikitch, A.E. Punt, and R. Hilborn. 1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences* 51: 2673-2687.

Neave, F. 1943. Diurnal fluctuations in the upstream migration of coho and spring salmon. *Journal of the Fisheries Research Board of Canada* 6:158-163.

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: 554-556.

Neilson, J.D. and B.E. Banford. 1983. Chinook salmon (*Oncorhynchus tshawytscha*) spawner characteristics in relation to redd physical features. *Canadian Journal of Zoology* 61: 1524-1531.

Parken, C.K., R.E. Bailey, 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.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries management. *Canadian Journal of Fisheries and Aquatic Science* 47: 2 – 15.

Peterman, R.M. 2004. Possible solutions to some challenges facing fisheries scientists and managers. *ICES Journal of Marine Science*, 61: 1331-13343.

Peterman, R.M. and M.J. Bradford. 1987. Statistical power of trends in fish abundance. *Canadian Journal of Fisheries and Aquatic Science* 44: 1879 – 1889.

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.

Quinn, T.J. and R. Gates. 1997. Estimation of salmon escapement: models with entry, mortality, and stochasticity. *Natural Resource Modeling* 10: 217-250.

Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. *In* Bayesian Statistics 3. *Edited by* J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith. pp. 395-402. Oxford University Press.

Salo, E.O. 1991. Life history of chum salmon (*Oncorhynchus keta*). *In* Pacific Salmon Life Histories. *Edited by* C. Groot and L. Margolis. pp. 231-310. UBC Press, Vancouver.

Shardlow, T., R. Hilborn, and D. Lightly. 1987. Components analysis of instream escapement methods for Pacific salmon (*Onchorhynchus spp.*) Canadian Journal of Fisheries and Aquatic Sciences 44: 1031-1037.

Shardlow, T. 2004. Using time-lapsed video to estimate survey life for area-under-the-curve methods of escapement estimation. *North American Journal of Fisheries Management* 24: 1413-1420.

Shaul, L., S. McPhearson, E. Jones, and K. Crabtree. 2004. Stock status and escapement goals for coho salmon stocks in Southeast Alaska [*in*] Stock Status and Escapement Goals for Salmon Stocks in Southeast Alaska. H.J. Geiger and S. McPherson [*eds*]. Alaska Department of Fish and Game, Divisions of Sport Fish and Commercial Fisheries, Special Publication No. 04-0.

Smoker, W.W., A.J. Gharrett, and M.S. Stekoll. 1998. Genetic variation of return date in a population of pink salmon: a consequence of fluctuating environment and dispersive selection? *Alaskan Fishery Research Bulletin* 5: 46-51.

Su, Z., M.D. Adkison, and B.W. Van Alen. 2001. A hierarchical Bayesian model for estimating historical escapement and escapement timing. *Canadian Journal of Fisheries and Aquatic Sciences* 58: 1648-1662.

Stewart, D.C. G.W. Smith, and A.F. Youngson. 2002. Tributary-specific variation in timing of return of adult Atlantic salmon (*Salmo salar*) to freshwater has a genetic component. *Canadian Journal of Fisheries and Aquatic Science* 59: 276-281.

Thomas, L. 1996. Monitoring long-term population change: why are there so many analysis methods? *Ecology* 77: 49 - 58.

Trouton, N.D. 2004. *An investigation into the factors influencing escapement estimation for chinook salmon (Oncorhynchus tshawytscha) on the Lower Shuswap River, British Columbia.* M.Sc. Thesis. Simon Fraser University.

Van den Berghe, E.P. and M.R. Gross. 1986. Length of breeding life of coho salmon (*Oncorhynchus kisutch*). *Canadian Journal of Zoology* 64: 1482-1486.

Wade, P.R. 2000. Bayesian methods in conservation biology. *Conservation Biology* 14: 1308 – 1316.

Walters, C.J. 1986. *Adaptive Management of Renewable Resources.* Macmillan, New York.

West, J.F. and R.H. Goode. 1987. Aerial counts of spawning chinook salmon (*Oncorhynchus tshawytscha*) on the Rakaia River system, Canterbury, New Zealand, 1973-76. *New Zealand Journal of Marine and Freshwater Research* 21: 563-572.

Zar, J.H. 1999. *Biostatistical Analysis*, 4th edition. Prentice Hall Inc., New Jersey.