

A DISCUSSION AND ANALYSIS OF FOUR CONSTANT FRACTIONAL MARKING
ALTERNATIVES FOR CALIFORNIA'S CENTRAL VALLEY SALMON HATCHERIES

A Thesis

Presented in Partial Fulfillment of the Requirements for the

Degree of Master of Science

with a

Major in Statistics

in the

College of Graduate Studies

University of Idaho

by

Allan C. Hicks

June 2003

Major Professor: Ken B. Newman, Ph.D.

AUTHORIZATION TO SUBMIT THESIS

This thesis of Allan C. Hicks, submitted for the degree of Master of Science with a major in Statistics and titled “A discussion and analysis of four constant fractional marking alternatives for California’s central valley salmon hatcheries,” has been reviewed in final form. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor _____ Date _____
Ken B. Newman

Committee _____ Date _____
Members _____
David G. Hankin

_____ Date _____
Brian C. Dennis

Department _____ Date _____
Administrator _____
Christopher J. Williams

Science _____ Date _____
College Dean _____
Earl H. Bennett

Final Approval and Acceptance by the College of Graduate Studies

_____ Date _____
Margrit von Braun

Abstract

Hatchery chinook salmon are typically marked by clipping the adipose fin and tagged with a coded-wire tag for later identification. However, all juveniles released from a hatchery are not always marked and tagged, and tagging levels may vary each year. A constant fractional marking program is described where a certain percentage of juvenile hatchery fish are marked and tagged every year allowing for the estimation of natural production. Four marking alternatives where all of the hatchery fish were either marked or some were left unmarked, applied in combination to selective and non-selective fisheries, were considered. Estimation routines for these alternatives were derived using sampling theory and method of moments, and bootstrapping methods to estimate the variance of these estimators is discussed. An analysis using a factorial design was done to determine the effects of marking levels and sampling rates on the precision of the natural production estimates for each alternative. An increase in escapement sampling rate showed the most significant increased in precision for all alternatives, but when some hatchery fish were left unmarked, the constant fractional marking rate was also an important factor. The natural production estimates were less precise in alternatives that did not mark all of the hatchery fish.

Acknowledgments

From the time I applied to the graduate program, to the few days after my defense, there are many people that I would like to acknowledge. First and foremost, I would like to sincerely thank my adviser, Dr. Ken Newman. His patience, belief, and understanding were crucial to the completion of my thesis. My undergraduate adviser and committee member, Dr. Dave Hankin, presented me the opportunity to work with Dr. Newman, has always provided insight into any problem I have worked on, and has inspired me to become a fisheries scientist. I am also very grateful to my committee member and professor Dr. Brian Dennis. He amazes me with his talent to put anything ecological into a probabilistic framework, provided me with very important references important to this thesis, and taught me the importance of the scientific method.

The funding for this work was provided by Randy Bailey of Bailey Environmental. I am very thankful to him for answering all of my questions, providing me with financial support, and being a wonderful host when I presented the final simulation program to him.

I would have never been able to finish my graduate studies without the belief and support of my family. The pride that my father has in me pushes me to excel beyond where I believe I can go and he was always there financially and emotionally. The support of my two sisters is unsurpassed and they were always there when I needed them. My aunt and uncle always provided me with inspiration at those critical moments. Prince Lewis gave me an exceptional insight into life which keeps me thinking. And, the assurance and wisdom of my mother will never be forgotten.

I am also grateful to the many friends and professors that were supportive during my time as a graduate student. These people know who they are and I will mention a few. Ben Allen provided computer support and necessary “vacations” from my studies and research. Al Manson was a big influence with his incredible knowledge and amazing capabilities as head of the Division. He will be missed. The secretaries of the Division always knew what to do at those critical moments. And, while I was in New Zealand, Rosie Hurst, Di Tracey, Peter McMillan, Richard O’Driscoll, Garth Bloomfield, and Sandra Diesveld helped me through some very difficult times.

Finally, the love, support, belief, and understanding that Renée Ugarteche has given me will always be cherished. She has guided me through the toughest of times, and been there for the best. I am very grateful to her.

This thesis is dedicated to my mother.
She showed me how to make a path in the garden of life.

Contents

| | |
|---|------------|
| Authorization to submit thesis | ii |
| Abstract | iii |
| Acknowledgments | iv |
| 1 Introduction | 1 |
| 2 Central Valley hatcheries | 3 |
| 3 Marking, tagging, and data collection | 3 |
| 4 Model | 5 |
| 4.1 Notation | 5 |
| 4.2 Sequence of processes | 5 |
| 4.3 Underlying probability distributions | 7 |
| 4.3.1 Process error | 7 |
| 4.3.2 Observation error | 8 |
| 5 Production estimates | 9 |
| 5.1 Alternative NS1 | 10 |
| 5.1.1 NS1: Estimating hatchery specific production | 10 |
| 5.1.2 NS1: Estimating watershed specific natural production | 11 |
| 5.2 Alternative NS2 | 12 |
| 5.2.1 NS2: Estimating hatchery specific production | 13 |
| 5.2.2 NS2: Estimating watershed specific natural production | 13 |
| 5.3 Alternative AS1 | 14 |
| 5.3.1 AS1: Hatchery specific production | 14 |
| 5.3.2 AS1: Watershed specific natural production | 14 |
| 5.3.3 AS1: Effects of incidental fishing mortality | 15 |
| 5.4 Alternative AS2 | 16 |
| 5.4.1 AS2: Hatchery specific production | 16 |
| 5.4.2 AS2: Watershed specific natural production | 17 |

| | | |
|-----------|--|-----------|
| 5.4.3 | AS2: Effects of incidental fishing mortality | 18 |
| 6 | Variance estimation | 18 |
| 6.1 | Variances of catch and escapement estimates | 19 |
| 6.1.1 | Analytical variance calculation | 19 |
| 6.1.2 | Approximate variance calculation | 20 |
| 6.2 | Variance of Δ_{it} | 22 |
| 6.2.1 | Analytical variance calculation | 23 |
| 6.2.2 | Approximate variance calculation | 23 |
| 7 | Assumptions | 23 |
| 8 | Simulation and estimation program, <i>CFM Sim</i> | 24 |
| 8.1 | User input | 25 |
| 8.2 | Simulation details | 26 |
| 8.3 | Program output | 27 |
| 9 | Analysis of alternatives | 28 |
| 9.1 | Fixed parameters | 28 |
| 9.2 | No selective fishery alternatives | 30 |
| 9.2.1 | Design | 30 |
| 9.2.2 | Production estimates using alternative NS1 | 31 |
| 9.2.3 | Production estimates using alternative NS2 | 33 |
| 9.2.4 | Comparison of NS1 and NS2 | 36 |
| 9.3 | All selective fishery alternatives | 36 |
| 9.3.1 | Design | 37 |
| 9.3.2 | Production estimates using alternative AS1 | 37 |
| 9.3.3 | Δ_{nt} estimates under alternative AS1 | 39 |
| 9.3.4 | Production estimates using alternative AS2 | 40 |
| 9.3.5 | Δ_{nt} estimates under alternative AS2 | 42 |
| 9.3.6 | Comparison between alternatives AS1 and AS2 | 43 |
| 10 | Discussion | 44 |

| | | |
|-----------|---|-----------|
| 11 | References | 50 |
| A | Estimating Δ_{it} in selective fisheries | 83 |
| A.1 | Estimating life history parameters using the <i>Surrogate</i> group | 83 |
| A.2 | Estimating Δ_{et} for the <i>Stealth</i> group | 84 |
| A.3 | Estimating Δ_{nt} for the natural stock | 84 |
| B | Estimating Δ_{nt} when <i>Stealth</i> fish stray | 85 |

List of Tables

| | | |
|----|---|----|
| 1 | The 14 watersheds and the 5 hatcheries in the Central Valley and the total production of chinook salmon for each hatchery. | 52 |
| 2 | Different hatchery fish adipose fin clip and tagging schemes. NS and AS denote no selective fisheries and all selective fisheries, and f is the CFM rate. | 52 |
| 3 | The notation for abundance and life history parameters. | 53 |
| 4 | Notation for samples and sample recoveries. | 54 |
| 5 | The fixed minimum, mode, and maximum of the survival, ocean harvest, and freshwater harvest parameters (in percent) for all stocks. | 55 |
| 6 | Maturity parameters, in percent, for the two hatchery stocks. The natural stocks have the same parameters as its surrogate. | 55 |
| 7 | Straying rates for the two hatchery stocks. The natural stocks do not stray from their native watershed. | 55 |
| 8 | The experimental design for the simulation studies of alternatives NS1 and NS2. <i>CFM</i> is the constant fractional marking rate, <i>ESR</i> is the escapement sampling rate, <i>FSR</i> is the freshwater sampling rate, <i>Sur</i> is the number of surrogates released, and <i>TSR</i> is the terminal area sampling rate. The letters “l” and “h” denote low and high levels for the factors <i>CFM</i> , <i>ESR</i> , <i>FSR</i> , <i>Sur</i> , and <i>TSR</i> , respectively. | 56 |
| 9 | The number of simulations where the estimated production was less than the true production out of 1000 simulations for each natural stock and alternatives NS1 and NS2. “WS” refers to the watershed that the natural stock is natal to. The levels of the variables are given for reference. Refer to Table 8 for an explanation of the levels. | 57 |
| 10 | Number of negative estimates for each scenario out of 1000 simulations using alternative NS2. The levels of the variables are given for reference. Refer to Table 8 for an explanation of the levels. | 58 |
| 11 | Means and medians of the simulated natural production <i>MAE</i> 's (in %) for the high (H) and low (L) levels of the variables in the factorial design for the three natural stocks and alternatives NS1 and NS2. Section 9.2.1 defines the variables and their acronyms. | 59 |

| | | |
|----|--|----|
| 12 | All possible treatment combinations in the 2^5 factorial design used to analyze alternatives AS1 and AS2. CFM is the constant fractional marking rate, SR is the escapement and freshwater fisheries sampling rates, SM is the shaker mortality rate, Sur is the number of surrogates released, and Stl is the number of stealth fish released. The letters “l” and “h” denote low and high levels for the factors CFM , SR , SM , Sur , and Stl , respectively. | 60 |
| 13 | The number of simulations where the estimated production was less than the true production out of 1000 simulations for alternatives AS1 and AS2. The levels of the variables are given for reference. Refer to Table 12 for an explanation of the levels. | 61 |
| 14 | Means and medians of the simulated natural production MAE 's (in %) for the high (H) and low (L) levels of the variables in the factorial design for the three natural stocks and alternatives AS1 and AS2. Section 9.3.1 defines the variables and their acronyms. | 62 |
| 15 | Means and medians of the simulated Δ_{nt} MAE 's (in %) for the high (H) and low (L) levels of the variables in the factorial design for the three natural stocks and alternatives AS1 and AS2. Section 9.3.1 defines the variables and their acronyms. | 63 |
| 16 | Number of negative estimates for each scenario out of 1000 simulations using alternative AS2. The levels of the variables are given for reference. Refer to Table 12 for an explanation of the levels. | 64 |

List of Figures

| | | |
|----|--|----|
| 1 | Approximate locations of the five hatcheries on the Sacramento–San Joaquin River system in California. | 65 |
| 2 | The sequence of fates that a fish can have. See Table 3 for an explanation of the variables. Probabilities at each step are indicated by Greek letters. | 66 |
| 3 | Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative NS1. | 67 |
| 4 | Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the <i>MAE</i> of estimated production for each treatment level and natural stock simulated under alternative NS1. | 68 |
| 5 | Significant <i>ESR–FSR</i> and <i>TSR–Sur</i> interactions on the transformed <i>MAE</i> for Watershed 3 simulated under alternative NS1. | 69 |
| 6 | Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative NS2. | 70 |
| 7 | Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the <i>MAE</i> of estimated production for each treatment level and natural stock simulated under alternative NS2. | 71 |
| 8 | Significant <i>CFM–ESR</i> , <i>CFM–TSR</i> , and <i>ESR–TSR</i> interactions on the transformed <i>MAE</i> for Watershed 1 simulated under alternative NS2. | 72 |
| 9 | Significant <i>CFM–ESR</i> and <i>TSR–Sur</i> interactions on the transformed <i>MAE</i> for Watershed 3 simulated under alternative NS2. | 73 |
| 10 | Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative AS1. | 74 |
| 11 | Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the <i>MAE</i> of estimated production for each treatment level and natural stock simulated under alternative AS1. | 75 |
| 12 | Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the <i>MAE</i> of $\hat{\Delta}_{nt}$ for each treatment level and natural stock simulated under alternative AS1. | 76 |
| 13 | Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative AS2. | 77 |
| 14 | Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the <i>MAE</i> of estimated production for each treatment level and natural stock simulated under alternative AS2. | 78 |

| | | |
|----|---|----|
| 15 | Significant <i>CFM-SR</i> , <i>CFM-Sur</i> , <i>SR-Sur</i> , and <i>Sur-Stl</i> interactions on the transformed <i>MAE</i> for Watershed 1 simulated under alternative AS2. | 79 |
| 16 | Significant <i>CFM-SR</i> , <i>CFM-Sur</i> , <i>SR-Sur</i> , and <i>CFM-SM</i> interactions on the transformed <i>MAE</i> for Watershed 2 simulated under alternative AS2. | 80 |
| 17 | Significant <i>CFM-SR</i> , <i>CFM-Sur</i> , <i>SR-Stl</i> , and <i>Sur-Stl</i> interactions on the transformed <i>MAE</i> for Watershed 2 simulated under alternative AS2. | 81 |
| 18 | Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the <i>MAE</i> of $\hat{\Delta}_{nt}$ for each treatment level and natural stock simulated under alternative AS2. | 82 |

1 Introduction

Marked animals have been used to estimate survival and harvest rates in fisheries and wildlife studies for many decades (Youngs & Robson 1975, Brownie et. al. 1985, Lebreton et. al. 1992). More recent advances in fisheries have allowed the estimation of tag reporting rates using these methods in conjunction with creel surveys or port sampling (Pollock et. al. 1991) or by using multiple tagging models (Hearn et. al. 1998). Estimates of exploitation rates for multiple fisheries were derived by Brooks et. al. (1998). These models use a sequence of binary events to create multinomial distributions from which parameters are estimated.

Chinook salmon (*Oncorhynchus tshawytscha*) released from hatcheries are typically marked by clipping the adipose fin (also called an ad-clip) and tagged by inserting a coded-wire tag (CWT), engraved with a unique identifier for that release group, into the snout. The harvest and escapement of these release groups can be used to estimate various life-history parameters, and some release groups have even been used as indicators for natural populations (salmon produced from naturally spawning adults). Given the decline of salmon stocks in the Pacific Northwest (Committee on Protection and Management of Pacific Northwest Anadromous Salmonids, et. al., 1996), there is a great importance to adequately and consistently mark some or all hatchery released salmon to continue the research on and the monitoring of Pacific salmon stocks.

Healey (1991) presents a detailed look at the life-history of chinook salmon, which is summarized here for fall-run or “ocean-type” races. Young salmon emerge in the spring and migrate downstream to the ocean during their first year of life. They typically do not migrate far from the coastal waters and return to their natal river in the fall just days or weeks before spawning. The age at which they mature is variable, typically ranging from 2 through 5, resulting in catches and escapement consisting of multiple cohorts.

To help protect natural salmon stocks there has been an interest in selective fisheries, where only marked fish can be retained and unmarked fish must be returned to the water. Questions such as how many hatchery fish should be marked to allow a respectable sized fishery, and what is the impact on unmarked fish that are caught and released in a fishery have been of concern. Lawson and Sampson (1996) reported that the mortality rate suffered by unmarked fish in a selective fishery is dependent upon many factors such as harvest rate for the marked fish, the accuracy of mark recognition, and the proportion of marked and unmarked fish at the beginning of fishing. This mortality rate, or latent hooking mortality, can be broken down into two types. One is shaker mortality, which occurs when a fish dies after it has been caught and subsequently released back into the water. Its name is derived from the fishermen typically shaking the fish off the hook. The second type is called drop off mortality and occurs when a fish suffers mortality related to the escape from the hook prior to being landed, or when a predator, such as a shark or seal, removes the fish from the hook. Estimating either of these mortalities is difficult, especially when multiple fisheries are present. However, the effects of

these mortalities on escapement can be assessed and is discussed later.

The question of how many hatchery salmon should be marked has recently been answered by legislation. Language inserted into the Interior Appropriations bill states that all federally operated or federally funded hatcheries that release salmonids shall implement a mass marking program. This means that all fish released from a hatchery would be at least externally marked, resulting in some groups of fish being released with only an ad-clip and no CWT. Such mass marking may be beneficial to selective fisheries, but the effects on research and analysis are unknown.

Estimating the population size of each stock with precision is extremely important to adequately monitor the status of the stocks. However, it is difficult to accurately estimate the catches and escapement of natural stocks, especially when hatchery fish are present in the population. Hankin (1982) proposed constant fractional marking (CFM) as a means of estimating the percentage of hatchery fish in the escapement to a watershed. A fixed percentage, f , of all groups of juvenile salmon released from a hatchery are marked prior to release, where the percentage f remains constant over time. The mark is preferably one that can be detected without sacrificing the fish, such as clipping the adipose fin. When the marked and unmarked hatchery fish return from the ocean to the natal watershed, the percentage of hatchery fish in the escapement can be estimated from a sample of the escapement. And, because the CFM rate remains constant over time, aging of the samples is not necessary to account for the multiple cohorts in the sample.

More work was done to extend the CFM method after the passage of the Central Valley Project Improvement Act (CVPIA) in 1992. This required the U.S. Fish and Wildlife Service to develop a Comprehensive Assessment and Monitoring Program (CAMP) to monitor the trends and effects on anadromous fish populations subsequent to actions implemented under CVPIA. Newman (2000a) proposed eight different marking scenarios or alternatives to estimate hatchery and natural production of chinook salmon from the Sacramento-San Joaquin river system in California (Figure 1) using constant fractional marking, but only four were determined to be viable.

In this thesis, I extend further on the work of Newman (2000a) by adding a terminal area fishery, deriving equations to estimate the effects of selective fisheries on natural production, elucidating the natural and observation processes in the model, discussing possible methods for variance estimation, and finally, performing a formal analysis on the effects that some parameters have on the natural production estimates under all four alternatives. I begin by describing the five chinook salmon hatcheries in the Sacramento-San Joaquin Valley. I then review the work of Newman (2000a), discussing the notation, the model, and add in an explanation of the natural and observation probability distributions. Afterwards, I give an up to date listing of the estimation equations, followed by a discussion of the variability in the estimates. Before

explaining the simulation program and analysis, I review some of the important assumptions in the work, then conclude with a discussion of this work as well as future possibilities for improvement.

2 Central Valley hatcheries

There are five hatcheries in the Central Valley of California that release chinook salmon into the fourteen watersheds of the Sacramento-San Joaquin River Basin (Table 1 and Figure 1). Coleman National Fish Hatchery on Battle Creek near Red Bluff is the largest salmon hatchery in the contiguous 48 states. It was built in 1942 after the construction of Shasta and Keswick dams and is federally funded and managed. Feather River Hatchery on the Feather River near Oroville is operated by the California Department of Fish and Game (CDFG) with funding from the Department of Water Resources and the State Water Contractors. It was opened in 1967 after Oroville Dam blocked the upstream migration of adult salmon. Nimbus Hatchery on the American River near Sacramento was built after the construction of Nimbus and Folsom dams. The CDFG is funded by the U.S. Bureau of Reclamation to operate and maintain this hatchery. Mokelumne Hatchery on the Mokelumne River south of Sacramento is operated by the CDFG and funded by the East Bay Municipal Utility District. Merced Hatchery on the Merced River near Merced is operated and funded by the CDFG. In total, these hatcheries release over 30 million fall-run chinook salmon smolts each year. The approximate production levels of each hatchery are displayed in Table 1.

3 Marking, tagging, and data collection

To implement and monitor a constant fractional marking program, various release categories are required with different types dependent on whether or not a selective fishery is in place. Newman (2000a) developed two marking alternatives for no selective fisheries, and two alternatives when selective fisheries are present in all fisheries. These alternatives were labeled 5–6 in Newman (2000a) and have been changed to NS1, NS2, AS1, and AS2 here. “NS” refers to no selective fisheries, while “AS” refers to all selective fisheries. The numbers 1 and 2 respectively refer to whether the remaining fish after the required numbers have been tagged with a coded-wire tag (CWT) are marked by clipping the adipose fin (ad-clip) or left unmarked.

As specified by Newman (2000a), the hatchery releases can be partitioned into one of five different categories, subscripted here with a , b , c , d , or e .

Ad hoc releases, a: fish that are not assumed to represent natural fish, e.g., experimental releases, and are internally tagged with a CWT and ad-clipped. There are no restrictions on the numbers in this group and multiple CWT codes can be used.

Surrogate releases, *b*: fish that are assumed to have the same natural survival rates, migration paths, fisheries vulnerability (in non-selective fisheries), and maturation probabilities as a designated natural stock. These fish are marked and tagged and the CWT code must be unique for later identification.

CFM releases, *c*: fish that are CWT'd and ad-clipped, and are a constant fraction, f , of a larger group of releases (Hankin 1982). This larger group, (*CFM* and *Remainder*), may be a mixture of types of releases, fingerlings or yearlings, for example, and include so-called production releases. The CWT codes need to be unique, however, for each of the subgroups in the mixture if inferences about the subgroups are desired.

Remainder releases, *d*: fish that are not CWT'd and may or may not receive an ad-clip. They are the complement of the CFM releases, i.e., the remainder is a constant fraction, $1-f$, of a larger group of releases.

Stealth releases, *e*: fish that are CWT'd but not marked (ad-clipped), and assumed to be biologically equivalent to the *Surrogate* release (thus represent the same natural stock as the *Surrogates*). These fish are assumed to have the same straying rates (if any) as the natural stock. The CWT code needs to be unique. These releases are only required in the presence of selective fisheries.

The no selective fisheries alternatives, NS1 and NS2, require four release groups: *Ad hoc*, *Surrogate*, *CFM*, and *Remainder*. The only difference between NS1 and NS2 is that the *Remainder* group is ad-clipped in NS1 and left unmarked in NS2. The all selective fisheries alternatives, AS1 and AS2, are similar to alternatives NS1 and NS2, except for the inclusion of the required *Stealth* group. The *Remainder* group is ad-clipped in AS1 and left unmarked in AS2. Table 2 displays the release requirements for each alternative.

It is assumed that the sampled fish with a CWT are identified, which may require scanning the head of a fish when *Stealth* fish are present, or when ad-clipped fish without CWTs are present. Scanning can be done a number of ways, and it is assumed here that the CWTs are identified accurately. Alternative NS1 consists of fish that have an ad-clip, but may not have a CWT. Therefore, all fish with an ad-clip must be scanned for the presence of a CWT. Alternative NS2 requires no scanning because all fish with an ad-clip also contain a CWT. The scanning requirements for the sampled catch in the all selective fisheries alternatives, which consist of only ad-clipped fish, are the same as in the no selective fishery alternatives. The main difference is that the unmarked fish in the escapement sample are scanned to identify stealth fish. Once a CWT is detected, the head would be collected for CWT extraction and reading.

4 Model

The underlying model simplifies the life history of chinook salmon by assuming they experience a sequence of binary events such as survival, harvest, and maturation. First, the notation will be defined, and then the sequence of events will be described. Following the explanation of the model, the underlying probability distributions will be discussed.

4.1 Notation

The notation is similar to that of Hankin and Healey (1986) and Newman (2000a). The suffix i is used to denote a particular stock, and in later sections n and h will also be used to distinguish natural and hatchery stocks. The term stock, as used here, refers to releases from a particular hatchery, or the progeny of fish that spawned naturally in a particular watershed. The term “natural” is used to describe fish that were the result of natural spawning in a river, regardless of whether or not the parents were of hatchery origin.

Table 3 explains the notation for abundances and life-history parameters used to describe the model. Abundances are noted with a capital Roman letter while Greek letters refer to life-history parameters. Catch and escapement summed over stocks and/or ages are denoted by dropping subscripts; e.g., C_O is total ocean catch of all ages and stocks for a given year.

Sample values are notated using a small letter for the sample origin (i.e., ocean catch) followed by another small letter indicating the origin of the fish (i.e., release group or natural). These may also be subscripted with i for stock, a for age, and/or j for watershed, as well as other descriptors. For example, the sample of age a terminal caught fish from watershed j in the *Surrogate* release group from stock i is denoted tb_{iaj} .

4.2 Sequence of processes

The sequence of binary events experienced by a fish still alive at each point in time attempts to imitate the life-history of a fall-run chinook salmon from the Central Valley in California. Juvenile chinook salmon leave freshwater and enter the ocean sometime during their first year (age 1). They then experience overwinter survival, harvest by the ocean fishery, and possible maturation as early as age 2. The fish that mature pass through a freshwater mainstem fishery after which they will enter a terminal area (or watershed). Natural fish are assumed to enter the watershed from which they were born, while hatchery fish may stray to any of the watersheds in the system. In a watershed they may experience a terminal fishery if one is present. The fish that survive the terminal fishery are counted as escapement, which is divided between fish that stayed in the river to spawn (in-river escapement) and those that enter the hatchery in that watershed (hatchery escapement), if one is present. The fish that do not mature, stay in the

ocean to age one year, and again experience overwinter survival, ocean harvest, and possible maturation. All fish are assumed to mature by age 5. Mature fish in the freshwater system are assumed to have 100% survival, outside of the fisheries, thus natural mortality is only experienced during overwinter survival in the ocean. Figure 2 graphically shows the sequence of events in the salmon's life history.

The probability of each fate is conditional on the fish surviving up to the point where it would experience the fate. For example, denote the probability that a fish is caught in the ocean fishery at age 3, given it has survived up to age 3, as μ_{O3} . This probability is conditional on the fish surviving up to the time immediately preceding the ocean fishery. However, the probability that the fish would be caught in the ocean fishery at age 3 after release from the hatchery is the product of the probabilities of all previous conditional fates and μ_{O3} .

$$\Pr(C_{Oi3}) = \phi_{iI}(1 - \mu_{Oi2})(1 - \sigma_{i2})\phi_{i3}\mu_{Oi3} \quad (1)$$

The parameters refer to overwinter survival (ϕ), maturation (σ), and ocean harvest (μ).

Using the notation described in Table 3, the probabilities and expected values of all the fates in Figures 2 can be defined. Given this assumed sequence of events, the expected abundance in the ocean of age a fish from stock i , is a function of previous survival, harvest, and maturation rates. For example, assuming ocean harvest begins at age 2

$$E[N_{i4}] = R_i\phi_{iI}(1 - \mu_{Oi2})(1 - \sigma_{i2})\phi_{i3}(1 - \mu_{Oi3})(1 - \sigma_{i3})\phi_{i4}.$$

Conditional on the abundance in the ocean at age a , N_{ia} , the expected catches and escapements for age a fish are

$$E[C_{Oia}|N_{ia}] = N_{ia}\mu_{Oia} \quad (2)$$

$$E[C_{Fia}|N_{ia}] = N_{ia}(1 - \mu_{Oia})\sigma_{ia}\mu_{Fia} \quad (3)$$

$$E[C_{Tiaj}|N_{ia}] = N_{ia}(1 - \mu_{Oia})\sigma_{ia}(1 - \mu_{Fia})\rho_{iaj}\mu_{Tiaj} \quad (4)$$

$$E[E_{iaj,\alpha}|N_{ia}] = N_{ia}(1 - \mu_{Oia})\sigma_{ia}(1 - \mu_{Fia})\rho_{iaj}(1 - \mu_{Tiaj})\eta_{iaj} \quad (5)$$

$$E[E_{iaj,\beta}|N_{ia}] = N_{ia}(1 - \mu_{Oia})\sigma_{ia}(1 - \mu_{Fia})\rho_{iaj}(1 - \mu_{Tiaj})(1 - \eta_{iaj}) \quad (6)$$

$$a = 2, 3, 4, 5$$

The production for a given stock, natural or hatchery, is the harvest and escapement summed over ages. It is assumed that ages 2, 3, 4, and 5 are harvested, maturation begins at age 2 (and such "jacks" or "jills" are counted in production) and all uncaught age 5 fish mature.

Therefore, the production is

$$P_i = \sum_{a=2}^5 [C_{Oia} + C_{Fia} + C_{Tia} + E_{ia,\alpha} + E_{ia,\beta}], \quad (7)$$

where the terminal catch and escapements are summed over all watersheds.

4.3 Underlying probability distributions

There are a number of different states in the model stemming from the natural processes described in Section 4.2. For example, there are unobservable states such as immediately prior to ocean harvest and after overwinter survival, after maturation when a fish has entered the freshwater system and will go through the mainstem fishery, and the state prior to terminal harvest and escapement after moving to a specific watershed. There are also catches and escapement, which are also part of the natural process, and samples are usually taken from these states. These samples are subject to sampling error and can be considered an observable process.

The underlying probability distributions associated with these processes arise from the assumed binary sequence of events. First, the unobservable process error is discussed, then the observable, or sampling, error is evaluated. The stock subscript, i , is dropped for convenience.

4.3.1 Process error

The state of most interest is the stock abundance immediately prior to ocean harvest, after overwinter survival (N_a). By conditioning on this state, all of the other states for age a can be found using the sequence of binomial events (see equations 2-6). The sequence of binary events up to N_a results in a binomial distribution, which may be conditioned on N_{a-1} .

$$N_2 \sim \text{Binomial}(R, \phi_I) \quad (8)$$

$$N_a | N_{a-1} \sim \text{Binomial}[N_{a-1}, (1 - \mu_{O,a-1})(1 - \sigma_{a-1})\phi_a] \quad a = 3, 4, 5 \quad (9)$$

Without conditioning on the previous state, N_{a-1} , the distributions are

$$N_2 \sim \text{Binomial}(R, \phi_I) \quad (10)$$

$$N_a \sim \text{Binomial} \left[R, \phi_I \prod_{g=3}^a (1 - \mu_{O,g-1})(1 - \sigma_{g-1})\phi_g \right] \quad a = 3, 4, 5 \quad (11)$$

Conditional on N_a , there are 5 general states of interest: ocean catch (C_{Oa}), freshwater mainstem catch (C_{Fa}), terminal area catch (C_{Taj}), in-river escapement ($E_{aj,\alpha}$), and hatchery

escapement ($E_{aj,\beta}$) (again, see equations 2-6). Terminal catch and escapements are watershed specific thus have j realizations. Each of these states, conditional on it surviving to the point immediately prior to that state, has a binomial distribution with a probability equal to the life-history parameter associated with that state.

More interesting, however, is that the sequence of processes can be seen as a multinomial distribution, conditioned on the state immediately prior to ocean harvest (N_a). For example, the catches and escapements of age 5 fish, conditional on N_5 , make a multinomial distribution because all age 5 fish are assumed to mature ($\sigma_5 = 1$).

$$\begin{aligned}
[C_{O5}, C_{F5}, C_{T5j}, E_{5j,\alpha}, E_{5j,\beta}] &\sim \text{Multinomial} [N_5, \mu_{O5}, (1 - \mu_{O5})(1)\mu_{F5}, \\
&(1 - \mu_{O5})(1)(1 - \mu_{F5})\mu_{T5j}, \\
&(1 - \mu_{O5})(1)(1 - \mu_{F5})(1 - \mu_{T5j})\eta_{i5j}, \\
&(1 - \mu_{O5})(1)(1 - \mu_{F5})(1 - \mu_{T5j})(1 - \eta_{i5j})] \quad (12)
\end{aligned}$$

The catches and escapements for ages 2–4, however, have an additional two unobservable states to complete the multinomial distribution because some fish may not mature. These states are the number that survive in the ocean to the next age (N_{a+1}) and the number that die in the ocean (D_{a+1}).

$$\begin{aligned}
[C_{Oa}, C_{Fa}, C_{Taj}, E_{aj,\alpha}, E_{aj,\beta}, N_{a+1}, D_{a+1}] &\sim \text{Multinomial} [N_a, \mu_{Oa}, (1 - \mu_{Oa})\sigma_a\mu_{Fa}, \\
&(1 - \mu_{Oa})\sigma_a(1 - \mu_{Fa})\mu_{Taj}, \\
&(1 - \mu_{Oa})\sigma_a(1 - \mu_{Fa})(1 - \mu_{Taj})\eta_{aj}, \\
&(1 - \mu_{Oa})\sigma_a(1 - \mu_{Fa})(1 - \mu_{Taj})(1 - \eta_{aj}), \\
&(1 - \mu_{Oa})(1 - \sigma_a)\phi_{a+1}, \\
&(1 - \mu_{Oa})(1 - \sigma_a)(1 - \phi_{a+1})] \quad (13)
\end{aligned}$$

Furthermore, the entire process, including all ages can be viewed using a multinomial distribution. Each possible state has a probability equal to the product of all probabilities of survival, harvest and maturation up to that state, with the “n” parameter equal to the release size.

4.3.2 Observation error

The states above represent the unobservable processes in the model. In addition, there are observations coming from the catch and escapement states which can be summarized as follows. Samples of fish belonging to different release groups from the hatcheries or natural fish from various watersheds are taken from the catches and escapements. For example, given the total

ocean catch (C_O), a sample will contain fish of different ages from various *Ad hoc*, *Surrogate*, *CFM*, *Remaining*, and *Stealth* release groups, when present, as well as natural fish originating from various watersheds. Each specific group of sample recoveries (i.e., *Ad hoc* fish of a specific age) can be thought of as following a hypergeometric distribution, which can be approximated by a binomial distribution with a probability equal to the proportion of fish from that group and age in the ocean catch (Evans et. al. 1993). Furthermore, thinking of all the recovery groups in the sample as a series of hypergeometric distributions, one can approximate the sampling distribution of all the recoveries using a multinomial distribution.

More specifically, assume there are two hatcheries each with the following release groups: *Ad hoc*, *Surrogate*, *CFM*, and *Remaining*. Also, assume there are three watersheds which all contain viable natural spawning populations. In the ocean catch, the multinomial sampling distribution can be summarized as

$$[xa_{12}, xb_{12}, xc_{12}, xa_{22}, \dots, xc_{25}, xd, xn] \sim \text{Mn}(n_O, pa_{12}, pb_{12}, pc_{12}, pa_{22}, \dots, pc_{25}, pd, pn) \quad (14)$$

where p refers to the proportion of that group in the total catch, and Mn refers to the multinomial distribution. The remaining and natural fish cannot be separated to a specific origin or age, thus are separately lumped into single groups. Other catches and escapement samples can also be generalized as with the ocean catches.

5 Production estimates

Newman (2000a) explained how to estimate the production for alternatives NS1, NS2, AS1, and AS2, assuming no terminal fishery for the non-selective fishery alternatives and only an ocean fishery for the all selective fisheries alternatives. Estimates of catch from a terminal fishery, and estimates of the effects from selective fisheries have since been derived and are included here.

Sampling theory and a method of moments approach were used by Newman (2000a) to derive the original estimation routines for alternatives NS1, NS2, AS1, and AS2. The same methods were used to derive the estimates of terminal catch and selective fishery effects. Some states had direct observations or samples, thus basic simple random sampling theory was used to estimate the total size of that state (Thompson 1992). Other states were not as easily identified by the samples, thus a method of moments approach was used. Method of moments is the process where the lowest possible order moments are used to solve for unknowns (Rice 1995). For the estimates provided here, the expected value, or the first moment, was used to solve for unknown states, given some strict assumptions. For some estimates, specific examples are given to exemplify this process.

The total non-stock specific catches and escapement are used in the estimation equations and it is assumed that these estimates of the total catch and total escapement, such as \hat{C}_O , are

available. The details of their estimation is not discussed here, but it is not of great importance because the point estimates given here will be calculated using the same methods, regardless of how the total catch or escapement is estimated. What is of a greater concern is that simple random samples are assumed to be taken from the catches and escapements which are typically sampled using a stratified sampling plan. This simplification may result in an underestimation of the precision in the catches.

Many of the equations presented by Newman (2000a) are repeated here for completeness and convenience.

5.1 Alternative NS1

In alternative NS1 there is no selective fishery, there are four categories of fish released at each hatchery, and simple random samples are taken of ocean catch and freshwater mainstem catch, as well as the terminal catch and escapement of each watershed. All fish are marked, thus natural fish are distinguishable from hatchery fish.

5.1.1 NS1: Estimating hatchery specific production

Newman (2000a) derived the equations to estimate the ocean catch and freshwater mainstem catch assuming that simple random samples are taken. For hatchery i , let xa_i , xb_i , xc_i be the number of fish sampled from the ocean catch that are from the *Ad hoc*, *Surrogate*, and *CFM* release groups, respectively, and let y denote the freshwater mainstem samples. Furthermore, \hat{C}_O and \hat{C}_F are the estimates of the total ocean and mainstem catches, while n_O and n_F are the sample sizes from each catch.

$$\hat{C}_{Ohi} = \frac{\hat{C}_O}{n_O} \left(xa_i + xb_i + \frac{xc_i}{f} \right) \quad (15)$$

$$\hat{C}_{Fhi} = \frac{\hat{C}_F}{n_F} \left(ya_i + yb_i + \frac{yc_i}{f} \right) \quad (16)$$

The constant fractional marking rate is noted with f and may be hatchery specific, which would change it to f_i .

For terminal harvest and escapement estimation it is assumed that hatchery fish can return to any watershed (in addition to the one the natal hatchery is located in). Let ta_{ij} , tb_{ij} , and tc_{ij} be the number of sample recoveries from the *Ad hoc*, *Surrogate*, and *CFM* release groups, respectively, in watershed j . \hat{C}_{Tj} is the total estimated terminal catch and n_{Tj} is the sample size from that catch. The estimated terminal harvest of fish from hatchery i is then

$$\hat{C}_{Thi} = \sum_{j=1}^k \frac{\hat{C}_{Tj}}{n_{Tj}} \left(ta_{ij} + tb_{ij} + \frac{tc_{ij}}{f} \right). \quad (17)$$

Escapement is estimated as in Newman (2000a) except that in-river and hatchery escapements are separated and denoted by the subscripts α and β , respectively.

$$\widehat{E}_{hi,\alpha} = \sum_{j=1}^k \frac{\widehat{E}_{j,\alpha}}{n_{E_{j,\alpha}}} \left(za_{ij,\alpha} + zb_{ij,\alpha} + \frac{zc_{ij,\alpha}}{f} \right) \quad (18)$$

$$\widehat{E}_{hi,\beta} = \sum_{j=1}^k \frac{\widehat{E}_{j,\beta}}{n_{E_{j,\beta}}} \left(za_{ij,\beta} + zb_{ij,\beta} + \frac{zc_{ij,\beta}}{f} \right). \quad (19)$$

The hatchery escapement sample size, $n_{E_{j,\beta}}$, may equal $\widehat{E}_{j,\beta}$ because all fish returning to the hatchery are typically sampled. An expansion of the hatchery escapement numbers is thus not usually necessary.

The total escapement for hatchery stock i is then the sum of the in-river and hatchery two escapement estimates.

$$\widehat{E}_{hi} = \widehat{E}_{hi,\alpha} + \widehat{E}_{hi,\beta} \quad (20)$$

5.1.2 NS1: Estimating watershed specific natural production

Because it is assumed that natural fish do not stray to other watersheds, it is quite simple to estimate the watershed specific natural terminal harvest and escapement. All hatchery fish have an ad-clip, thus any fish in the terminal catch or escapement without an ad-clip should be a natural fish. Let tn_j be the number of unclipped fish, or natural fish, in the terminal harvest sample from watershed j , and $zn_{j,\alpha}$ and $zn_{j,\beta}$ be the number of unclipped fish observed in the hatchery and in-river escapement samples. Then, the estimates of natural terminal harvest and natural escapement to watershed j are

$$\widehat{C}_{Tnj} = \frac{\widehat{C}_{Tj}}{n_{Tj}} tn_j \quad (21)$$

$$\widehat{E}_{nj} = \frac{\widehat{E}_{j,\alpha}}{n_{E_{j,\alpha}}} zn_{j,\alpha} + \frac{\widehat{E}_{j,\beta}}{n_{E_{j,\beta}}} zn_{j,\beta} \quad (22)$$

The ocean and freshwater mainstem catches of natural stock j are estimated with a method of moments style estimator using the hatchery surrogates sampled from the ocean catch, and the hatchery and natural escapement and terminal catch estimates.

$$\widehat{C}_{Onj} = \sum_{a=2}^5 \left(\frac{\widehat{C}_O}{n_O} xb_{ia} \right) \frac{\widehat{C}_{Tnja} + \widehat{E}_{nja}}{\sum_{j=1}^k \left(\frac{\widehat{C}_{Tj}}{n_{Tj}} tb_{ija} + \frac{\widehat{E}_{j,\alpha}}{n_{E_{j,\alpha}}} zb_{ija,\alpha} + \frac{\widehat{E}_{j,\beta}}{n_{E_{j,\beta}}} zb_{ija,\beta} \right)} \quad (23)$$

$$\widehat{C}_{Fnj} = \sum_{a=2}^5 \left(\frac{\widehat{C}_F}{n_F} y b_{ia} \right) \frac{\widehat{C}_{Tnja} + \widehat{E}_{nja}}{\sum_{j=1}^k \left(\frac{\widehat{C}_{Tj}}{n_{Tj}} t b_{ija} + \frac{\widehat{E}_{j,\alpha}}{n_{Ej,\alpha}} z b_{ija,\alpha} + \frac{\widehat{E}_{j,\beta}}{n_{Ej,\beta}} z b_{ija,\beta} \right)} \quad (24)$$

where

$$\widehat{C}_{Tnja} = \widehat{C}_{Tnj} \widehat{p}_{Tnja} \quad (25)$$

$$\widehat{E}_{nja} = \widehat{E}_{nj,\alpha} \widehat{p}_{nja,\alpha} + \widehat{E}_{nj,\beta} \widehat{p}_{nja,\beta}. \quad (26)$$

The estimated percentage of age a natural fish in the terminal catch, the in-river escapement, and the in-hatchery escapement is denoted \widehat{p}_{Tnja} , $\widehat{p}_{nja,\alpha}$ and $\widehat{p}_{nja,\beta}$. The specifics of finding the estimates of these proportions are not given here, but determining the ages of fish in a sub-sample using scales may be one method.

The intuition behind the estimate of C_{Onj} can be seen by substituting expected values (Section 4.2) for the estimated values in the right hand side of equation 23.

$$\begin{aligned} \widehat{C}_{Onj} &\approx \sum_{a=2}^5 N_{bia} \mu_{Oa} \frac{N_{nja} (1 - \mu_{Oa}) \sigma_a (1 - \mu_{Fa}) (\mu_{Tja} + (1 - \mu_{Tja}))}{\sum_{j=1}^k N_{bia} (1 - \mu_{Oa}) \sigma_a (1 - \mu_{Fa}) \rho_{ij} (\mu_{Tja} + (1 - \mu_{Tja}))} \\ &= \sum_{a=2}^5 N_{bia} \mu_{Oa} \frac{N_{nja} (1 - \mu_{Oa}) \sigma_a (1 - \mu_{Fa})}{N_{bia} (1 - \mu_{Oa}) \sigma_a (1 - \mu_{Fa}) \sum_{j=1}^k \rho_{ij}} \\ &= \sum_{a=2}^5 N_{bia} \mu_{Oa} \frac{N_{nja}}{N_{bia}} \\ &= \sum_{a=2}^5 N_{nja} \mu_{Oa} \end{aligned}$$

where N_{bia} and N_{nja} are the abundances of age a fish from hatchery surrogate i and natural stock j prior to ocean harvest. The life-history parameters are assumed to be the same for the natural stock and its surrogate, thus the stock subscripts were dropped. However, if they are not the same, the estimate of natural ocean catch will be biased. The natural production of stock j is the sum of the estimated catches and escapements.

$$\widehat{P}_{nj} = \widehat{C}_{Onj} + \widehat{C}_{Fnj} + \widehat{C}_{Tnj} + \widehat{E}_{nj,\alpha} + \widehat{E}_{nj,\beta}, \quad (27)$$

The assumption that the natural stock and its surrogate behave in the same way is crucial to estimate the production without bias. A similar argument applies to the estimation of C_{Fnj} .

5.2 Alternative NS2

Recall that there is no selective fishery and not all hatchery fish are marked.

5.2.1 NS2: Estimating hatchery specific production

The estimates of ocean catch, freshwater catch, terminal catch, and escapement of hatchery releases are identical to alternative NS1 (equations 15-20) because the *Remaining* release group is estimated by expanding the CFM group recoveries.

5.2.2 NS2: Estimating watershed specific natural production

Estimation of natural terminal catch and escapement to a given watershed differs from alternative NS1 in that estimates of hatchery terminal catch or escapement to the watershed are subtracted from the estimate of total terminal catch or escapement. Assuming that there are r hatchery stocks that contribute to the escapement in watershed j ,

$$\widehat{C}_{Tnj} = \widehat{C}_{Tj} - \sum_{i=1}^r \widehat{C}_{Thij} \quad (28)$$

$$\widehat{E}_{nj} = \widehat{E}_{j,\alpha} - \sum_{i=1}^r \widehat{E}_{hij,\alpha} + \widehat{E}_{j,\beta} - \sum_{i=1}^r \widehat{E}_{hij,\beta}, \quad (29)$$

where

$$\begin{aligned} \widehat{C}_{Thij} &= \frac{\widehat{C}_{Tj}}{n_{C_{Tj}}} \left(ta_{ij} + tb_{ij} + \frac{tc_{ij}}{f} \right), \\ \widehat{E}_{hij,\alpha} &= \frac{\widehat{E}_{j,\alpha}}{n_{E_{j,\alpha}}} \left(za_{ij,\alpha} + zb_{ij,\alpha} + \frac{zc_{ij,\alpha}}{f} \right), \text{ and} \\ \widehat{E}_{hij,\beta} &= \frac{\widehat{E}_{j,\beta}}{n_{E_{j,\beta}}} \left(za_{ij,\beta} + zb_{ij,\beta} + \frac{zc_{ij,\beta}}{f} \right). \end{aligned}$$

The catches of the natural stock in the ocean and freshwater mainstem fisheries are estimated as in alternative NS1 (equations 23 and 24). However, estimation of the age-specific natural stock numbers in the terminal catch and the escapement is more difficult because unmarked fish consists of natural and *Remaining* fish. A possible method to estimate the age proportions of natural fish is explained briefly here. First, find the proportion of unmarked fish in the terminal catch or escapement sample. Also, find the proportion of each age in a sample from only the unmarked fish. Then, multiply these two proportions and the total terminal catch or escapement estimate to get an estimate of the numbers of unmarked fish of a certain age in the terminal catch or escapement ($\widehat{C}_{Tuj a}$, $\widehat{E}_{uj a,\alpha}$, and $\widehat{E}_{uj a,\beta}$).

The contribution of the *Remainder* group to the terminal catch and escapement can be found using the *CFM* recoveries and the constant fractional marking rate, f .

$$\widehat{C}_{Tdi j a} = \frac{\widehat{C}_{Tj}}{n_{Tj}} \left[tc_{ija} \frac{1-f}{f} \right]$$

$$\begin{aligned}\widehat{E}_{dija,\alpha} &= \frac{\widehat{E}_{j,\alpha}}{n_{Ej,\alpha}} \left[zc_{ija,\alpha} \frac{1-f}{f} \right] \\ \widehat{E}_{dija,\beta} &= \frac{\widehat{E}_{j,\beta}}{n_{Ej,\beta}} \left[zc_{ija,\beta} \frac{1-f}{f} \right]\end{aligned}$$

Finally, to obtain estimates of age specific natural terminal catch and escapement, subtract the *Remainder* catches from the unmarked catches.

$$\widehat{C}_{Tnja} = \widehat{C}_{Tuja} - \sum_{i=1}^r \widehat{C}_{Tdijs} \quad (30)$$

$$\widehat{E}_{nja,\alpha} = \widehat{E}_{uja,\alpha} - \sum_{i=1}^r \widehat{E}_{dija,\alpha} \quad (31)$$

$$\widehat{E}_{nja,\beta} = \widehat{E}_{uja,\beta} - \sum_{i=1}^r \widehat{E}_{dija,\beta} \quad (32)$$

5.3 Alternative AS1

In alternative AS1 a selective fishery is assumed and there are five categories of hatchery releases including a stealth group.

It is assumed that all unmarked fish in the escapement samples are scanned for CWTs, thus recoveries from the stealth group are identified. It is further assumed that the selective fisheries do not retain any unmarked fish, but there will be latent hooking mortality for the unmarked fish.

5.3.1 AS1: Hatchery specific production

The estimates of hatchery specific ocean (\widehat{C}_{Ohi}), mainstem (\widehat{C}_{Fhi}), and terminal (\widehat{C}_{Thi}) harvests are the same as in equations 15-17 under alternative NS1. The “stealth” group can not be kept in any fishery, thus does not appear in catch samples. The expansion of the CFM recoveries remains appropriate because the harvest rate on the ad-clipped only fish should be the same as for the CFM group.

The estimate of hatchery specific escapement, \widehat{E}_{hi} , is also identical to alternative NS1 (equation 20). The data collection procedure differs, however, in that all unmarked fish in the escapement sample must be scanned, killing those with tags, or a subset of them, if they are not dead already.

5.3.2 AS1: Watershed specific natural production

No natural fish are kept in a selective fishery, thus natural production consists of only escapement. To estimate the natural escapement to watershed j , again assume no straying of natural

fish to other watersheds. Because of stealth groups some escapement sample recoveries will include hatchery fish without ad-clips, but assuming all unclipped fish are scanned for CWTs, the number of unclipped fish without CWTs are presumably natural fish and equation (22) can be used.

5.3.3 AS1: Effects of incidental fishing mortality

Incidental mortality, θ_{iat} is defined as the probability that a fish from stock i suffers mortality in any of the fisheries up to age a in year t . More specifically,

$$\theta_{iat} = 1 - \left[\prod_{g=2}^a (1 - \mu_{Oig,t+g-a}^*) \right] (1 - \mu_{Fiat}^*) (1 - \mu_{Tiajt}^*), \quad (33)$$

where μ_{Oiat}^* , μ_{Fiat}^* , and μ_{Tiajt}^* are the latent hooking mortality probabilities for an age a fish during year t in the ocean, freshwater mainstem, and terminal fisheries. The incidental mortality cannot be separately estimated for different fisheries and ages, because there are more unknown variables than known variables. However, a difference between the expected escapement if there was no incidental mortality (or no fishery intervention with the stock) and the actual escapement can be calculated for *Stealth* and natural fish. This is denoted by Δ_{it} , where t is the year of interest, and is defined as

$$\Delta_{it} = \sum_{a=2}^5 [\mathcal{E}(E_{iat} | \theta_{iat} = 0) - \mathcal{E}(E_{iat} | \theta_{iat} \neq 0)]. \quad (34)$$

\mathcal{E} indicates expected value and θ_{iat} is the probability that an unmarked fish from stock i suffers incidental mortality up to age a during year t . The subscripts e and n , for *Stealth* and natural, replace the i subscript to indicate those type of fish from a specific stock.

The notation $\theta_{iat} = 0$ indicates that no fishery intercepted the chosen stock, or that no fish died upon release from any fishery. Therefore, Δ_{it} is an indication of how many more fish would have escaped if there was no selective fishery or all unmarked fish survived the selective fishery. The ratio of $\mathcal{E}(E_{iat} | \theta_{iat} \neq 0)$ and $\mathcal{E}(E_{iat} | \theta_{iat} = 0)$ results in an estimate of θ_{iat} .

$$\theta_{iat} = 1 - \frac{\mathcal{E}(E_{iat} | \theta_{iat} \neq 0)}{\mathcal{E}(E_{iat} | \theta_{iat} = 0)} \quad (35)$$

To estimate Δ_{it} , it is *assumed* that age 3, 4, and 5 natural survival rates are known and constant. There may be ways to avoid assuming known survival rates, but some other parameters will likely have to be assumed known. Appendix A gives the details of estimating Δ_{it} .

There are some aspects of the estimation that are worth noting, however. The *Surrogate* and *Stealth* groups are used to estimate the parameters involved in the estimation of Δ_{it} . Age

specific natural escapements are needed, thus the unmarked fish in the escapement samples will have to be aged. Negative estimates of Δ_{it} may occur, which will be more likely when sampling rates are low. When assuming known age 3, 4, and 5 survival rates, data from the preceding three years, the current year, and three years into the future are needed (Appendix A). Assuming known survival rates is not desirable because there is likely natural variation between cohorts within the same stock, and between stocks. Newman (2000a) suggests that sampling from a known probability distribution of survival rates, then estimating the parameters used to estimate Δ_{it} will reflect some of the uncertainty in the parameter estimates. Finally, the assumption that *Stealth* fish are surrogates to the natural fish is crucial to the accurate estimation of Δ_{nt} . Appendix B shows how the straying of the *Stealth* group affects the estimate of Δ_{nt} and suggests a solution to account for that straying.

5.4 Alternative AS2

As in alternative AS1, a selective fishery is assumed to take place in all of the fisheries. However, in alternative AS2, the remaining 1- f % of the hatchery releases are unmarked and untagged, thus are subject to incidental mortality and are indistinguishable from natural fish.

5.4.1 AS2: Hatchery specific production

The estimates are similar to those of AS1, except that expansions for the 1- f % *Remainder* fish are not made in the catch estimates since unmarked fish are not kept.

$$\widehat{C}_{Ohi} = \frac{\widehat{C}_O}{n_O} (xa_i + xb_i + xc_i) \quad (36)$$

$$\widehat{C}_{Fhi} = \frac{\widehat{C}_F}{n_F} (ya_i + yb_i + yc_i) \quad (37)$$

$$\widehat{C}_{Thi} = \sum_{j=1}^k \frac{\widehat{C}_{Tj}}{n_{Tj}} (ta_{ij} + tb_{ij} + tc_{ij}). \quad (38)$$

However, as for NS2, estimation of hatchery-specific escapement requires that one estimate the contribution made by unmarked *Remainder* fish.

$$\begin{aligned} \widehat{E}_{hi} = & \sum_{j=1}^k \left[\frac{\widehat{E}_{j,\alpha}}{n_{E_{j,\alpha}}} (za_{ij,\alpha} + zb_{ij,\alpha} + zc_{ij,\alpha} + ze_{ij,\alpha}) + \sum_{a=2}^5 \widehat{E}_{dija,\alpha} \right] \\ & + \sum_{j=1}^k \left[\frac{\widehat{E}_{j,\beta}}{n_{E_{j,\beta}}} (za_{ij,\beta} + zb_{ij,\beta} + zc_{ij,\beta} + ze_{ij,\beta}) + \sum_{a=2}^5 \widehat{E}_{dija,\beta} \right] \end{aligned} \quad (39)$$

$\widehat{E}_{dija,\alpha}$ and $\widehat{E}_{dija,\beta}$ are estimates of unmarked *Remainder* fish from hatchery i , age a fish in the in-river and hatchery escapements to watershed j , and can be found using the escapement

of a *Stealth* group assumed to have the same life-history parameters. This stealth group does not have to be the same as that used as a surrogate to natural fish, which would require an additional release group. However, to reduce notation, a single *Stealth* group is assumed to be released from each hatchery. The escapements of age a *Stealth* fish from hatchery i fish, denoted $E_{eija,\alpha}$ and $E_{eija,\beta}$, can be estimated using the following equations.

$$\hat{E}_{eija,\alpha} = \frac{\hat{E}_{j,\alpha}}{n_{E_{j,\alpha}}} z_{eija,\alpha} \quad (40)$$

$$\hat{E}_{eija,\beta} = \frac{\hat{E}_{j,\beta}}{n_{E_{j,\beta}}} z_{eija,\beta} \quad (41)$$

The escapements of age a *Remainder* fish from hatchery i , $E_{dija,\alpha}$ and $E_{dija,\beta}$, are then estimated using the ratio of release numbers for this group and its surrogate stealth group.

$$\hat{E}_{dija,\alpha} = \frac{R_{di,t-a}}{R_{ei,t-a}} \hat{E}_{eija,\alpha} \quad (42)$$

$$\hat{E}_{dija,\beta} = \frac{R_{di,t-a}}{R_{ei,t-a}} \hat{E}_{eija,\beta} \quad (43)$$

where $R_{di,t-a}$ and $R_{ei,t-a}$ are the number of *Remaining* and *Stealth* fish released from hatchery i that are age a in the current year t .

5.4.2 AS2: Watershed specific natural production

The natural production consists of only escapement because no fishery is allowed to retain unmarked fish. To estimate escapement, the estimate of the *Remaining* escapement is subtracted from an estimate of the unmarked escapement. Let $E_{uj,\alpha}$ and $E_{uj,\beta}$ denote the in-river and hatchery escapement of unmarked fish to watershed j and $z_{uj,\alpha}$ and $z_{uj,\beta}$ be the number of unmarked fish observed in the escapement samples. The samples $z_{uj,\alpha}$ and $z_{uj,\beta}$ include *Remainder* fish and natural fish, but not the unmarked *Stealth* fish, which are separated by scanning for a CWT.

$$\hat{E}_{uj,\alpha} = \frac{\hat{E}_{j,\alpha}}{n_{E_{j,\alpha}}} z_{uj,\alpha} \quad (44)$$

$$\hat{E}_{uj,\beta} = \frac{\hat{E}_{j,\beta}}{n_{E_{j,\beta}}} z_{uj,\beta} \quad (45)$$

The escapement of natural fish to watershed j is estimated by subtracting the escapement estimates of *Remainder* fish (equations 42 and 43) from the estimated escapement of unmarked fish.

$$\hat{E}_{nj} = \hat{E}_{uj,\alpha} - \sum_{i=1}^r \sum_{a=2}^5 \hat{E}_{dija,\alpha} + \hat{E}_{uj,\beta} - \sum_{i=1}^r \sum_{a=2}^5 \hat{E}_{dija,\beta} \quad (46)$$

5.4.3 AS2: Effects of incidental fishing mortality

The procedures for estimating Δ_{it} are similar to those for AS1. Aged escapement samples are needed and age 3, 4, and 5 survival rates are assumed known. Additional details are given in Appendix A.

6 Variance estimation

Newman (2000a) suggested that the variances of the estimators, or at least suggestions for how to estimate the variances, is a topic for additional work. There are two approaches to calculating the variance of an estimated state. One is to calculate the variance of the estimates given a particular year, while the other is to incorporate the interannual variation in the life-history parameters to determine an overall variability in the estimates. Both approaches condition on the life-history parameters, but the difference is that these parameters are random variables when finding an overall variance.

The variance of the estimates for a given year are useful after finding point estimates for that year using actual collected data. Using the probability distributions defined in Section 4.3, analytical derivations of the variance are given for some of the more simple estimators. However, the more complex estimators incorporate multiple random variables that are not necessarily independent, resulting in complex analytical variances. Therefore, an approximate method to calculate the variances of all the estimators is ultimately suggested.

In contrast, when planning a CFM program or determining acceptable levels of releases and sampling rates, the overall variation in the estimates should be considered, taking into account the randomness in the life-history parameters. This is not an easy task to do analytically for a number of reasons. First, it may be difficult to determine the variability of the life-history parameters in order to incorporate it into the variance estimate. Second, the calculation of conditional variance can become complicated quickly due to nested expectations and variances. For example, assume X is the random variable of interest, and $\tilde{\mu}$ is a vector of the life-history parameters, which are all random variables. The variance of X can be calculated using the conditional variance formula (Rice 1995).

$$\text{Var}(X) = \text{E}[\text{Var}(X|\tilde{\mu})] + \text{Var}(\text{E}[X|\tilde{\mu}]) \quad (47)$$

In other words, the mean variation of X is found over the range of $\tilde{\mu}$ and added to the variation

of the expected value of X given values of $\tilde{\mu}$. Lastly, there are many life-history parameters making a joint distribution that can be difficult to find expectations and variances for. For these reasons, an analytical estimate of the overall variation in the state estimates is not given. However, described later in Section 8 is a simulation program called *CFM Sim* that incorporates the natural variability in life-history parameters and the sampling variability in observations to approximate the overall variation in production estimates.

6.1 Variances of catch and escapement estimates

6.1.1 Analytical variance calculation

The variabilities in catch and escapement estimates arise from sampling variation during the collection and identification of tagged fish. A simple example can be seen with the ocean catch estimates for hatchery i under alternative NS1 (equation 15). To begin, condition on the estimate of total ocean catch (which also can be thought of as assuming the total ocean catch is known). The variance of this estimate can be found using the multinomial sampling distribution given in equation 14.

$$\begin{aligned}
\text{Var}(\widehat{C}_{Ohi}|\widehat{C}_O) &= \text{Var}\left[\frac{\widehat{C}_O}{n_O}\left(xa_i + xb_i + \frac{xc_i}{f}\right)\right] \\
&= \left(\frac{\widehat{C}_O}{n_O}\right)^2 \text{Var}\left(xa_i + xb_i + \frac{xc_i}{f}\right) \\
&= \left(\frac{\widehat{C}_O}{n_O}\right)^2 \left[\text{Var}(xa_i) + \text{Var}(xb_i) + \frac{1}{f^2}\text{Var}(xc_i) \right] \\
&\quad + \left(\frac{\widehat{C}_O}{n_O}\right)^2 \left[2\text{Cov}(xa_i, xb_i) + \frac{2}{f}\text{Cov}(xa_i, xc_i) + \frac{2}{f}\text{Cov}(xb_i, xc_i) \right] \\
&= \left(\frac{\widehat{C}_O}{n_O}\right)^2 \left[n_O(pa_i)(1 - pa_i) + n_O(pb_i)(1 - pb_i) + \frac{n_O}{f^2}(pc_i)(1 - pc_i) \right] \\
&\quad - \left(\frac{\widehat{C}_O}{n_O}\right)^2 \left[2n_O(pa_i)(pb_i) + \frac{2n_O}{f}(pa_i)(pc_i) + \frac{2n_O}{f}(pb_i)(pc_i) \right] \\
&= \frac{\widehat{C}_O^2}{n_O} \left[(pa_i)(1 - pa_i) + (pb_i)(1 - pb_i) + \frac{1}{f^2}(pc_i)(1 - pc_i) \right] \\
&\quad - \frac{\widehat{C}_O^2}{n_O} \left[2(pa_i)(pb_i) + \frac{2}{f}(pa_i)(pc_i) + \frac{2}{f}(pb_i)(pc_i) \right]
\end{aligned}$$

The estimated proportions can be used to calculate the sample variance as follows

$$\begin{aligned}
\widehat{\text{Var}}(\widehat{C}_{Ohi}|\widehat{C}_O) &= \frac{\widehat{C}_O^2}{n_O} \left[\frac{xa_i}{n_O} \left(1 - \frac{xa_i}{n_O}\right) + \frac{xb_i}{n_O} \left(1 - \frac{xb_i}{n_O}\right) + \frac{1}{f^2} \left(\frac{xc_i}{n_O}\right) \left(1 - \frac{xc_i}{n_O}\right) \right] \\
&\quad - \frac{2\widehat{C}_O^2}{n_O} \left[\left(\frac{xa_i}{n_O}\right) \left(\frac{xb_i}{n_O}\right) - \frac{1}{f} \left(\frac{xa_i}{n_O}\right) \left(\frac{xc_i}{n_O}\right) - \frac{1}{f} \left(\frac{xb_i}{n_O}\right) \left(\frac{xc_i}{n_O}\right) \right] \quad (48)
\end{aligned}$$

The total catches and escapements are usually estimated, introducing additional variability into the estimate. This complicates the variance calculation because the probability distribution is now a product of random variables. However, assuming independence between the total catch estimate and the sample from the catch, the analytical variance can be calculated. For example, the true variance of the product of two independent random variables X and Y is

$$Var(XY) = [E(X)]^2Var(Y) + [E(Y)]^2Var(X) + Var(X)Var(Y) \quad (49)$$

Substituting in $X = \hat{C}_O$ and $Y = \frac{xa_i+xb_i+xc_i}{n_O}$, results in

$$\begin{aligned} \widehat{Var}(\hat{C}_{Ohi}) &= (\hat{C}_O)^2\widehat{Var}\left(\frac{xa_i + xb_i + xc_i}{n_O}\right) + \left(\frac{xa_i + xb_i + xc_i}{n_O}\right)^2\widehat{Var}(\hat{C}_O) + \\ &\quad \widehat{Var}(\hat{C}_O)\widehat{Var}\left(\frac{xa_i + xb_i + xc_i}{n_O}\right) \end{aligned} \quad (50)$$

where

$$\begin{aligned} \widehat{Var}\left(\frac{xa_i + xb_i + xc_i}{n_O}\right) &= \frac{1}{n_O} \left[\frac{xa_i}{n_O} \left(1 - \frac{xa_i}{n_O}\right) + \frac{xb_i}{n_O} \left(1 - \frac{xb_i}{n_O}\right) + \frac{1}{f^2} \left(\frac{xc_i}{n_O}\right) \left(1 - \frac{xc_i}{n_O}\right) \right] \\ &\quad - \frac{2}{n_O} \left[\left(\frac{xa_i}{n_O}\right) \left(\frac{xb_i}{n_O}\right) - \frac{1}{f} \left(\frac{xa_i}{n_O}\right) \left(\frac{xc_i}{n_O}\right) - \frac{1}{f} \left(\frac{xb_i}{n_O}\right) \left(\frac{xc_i}{n_O}\right) \right] \end{aligned}$$

As seen in this example, the variability in the estimate increases greatly when the total catch is not known without error. An unbiased sample variance subtracts the product of the variances instead of adding it (Goodman 1960).

The ideas presented here can be extended to other catch and escapement estimates that are also simple mean expansion estimates, such as the catch and escapement from hatchery stocks. Difficulty arises when estimating the ocean and mainstem catch of natural stocks (equations 23 and 24) because many sources of variation are present, and the ratio of random variables results in a complex solution. In addition, there is another level of variation in the aging samples. Therefore, instead of trying to find an analytical solution to the variance, approximate methods, such as the bootstrap, may be more applicable.

6.1.2 Approximate variance calculation

As explained in section 4.3, the underlying probability of the observations, given the total catch or escapement, can be approximated as multinomial with probabilities equal to the proportion of each group in the catch or escapement. Using this result, bootstrapping can be done to resample the data, supplying a bootstrapped estimate of the variance (Efron & Tibshirani 1998). For example, again using the estimate of ocean catch for hatchery i under alternative NS1 (equation 15), the multinomial model would be set up as having four groups to sample

from: xa_i , xb_i , xc_i , and an *Other* group, xo_i , that would represent fish not in the *Ad hoc*, *Surrogate*, or *CFM* group from hatchery i . In other words, the bootstrap samples would be as follows.

$$[xa_i^*, xb_i^*, xc_i^*, xo_i^*] \sim Mn\left(n_O, \frac{xa_i}{n_O}, \frac{xb_i}{n_O}, \frac{xc_i}{n_O}, 1 - \frac{xa_i + xb_i + xc_i}{n_O}\right) \quad (51)$$

The star superscripts indicate that the sample is not the original sample.

The use of the multinomial distribution is analogous to resampling the data directly. However, because there are typically many samples that originate from various sources and possibly watersheds, it may be easier to simply calculate the proportion of the group of interest in the sample for use in the multinomial distribution as explained above. This idea is adopted here instead of direct resampling of the data.

Also of consideration is the variability of the estimated total catch used to expand the recoveries. For each bootstrap sample, an estimated total catch will also have to be generated. How this is done is not discussed here because it is dependent on how the total catch was estimated. Nevertheless, to calculate the parametric bootstrapped variance, one would generate \hat{C}_{Ohi} for each bootstrap sample, and then determine the standard error of the estimates from those bootstrap samples. This procedure can be used for any estimators with the same form as equation 15.

It becomes more cumbersome when obtaining a bootstrapped estimate of the variance for the natural ocean and mainstem catches. These estimates in alternatives NS1 and NS2 (equations 23 and 24) use other estimates from natural and hatchery surrogate terminal catch and escapement. Plus, there is the underlying layer of sub-sampling to determine ages proportions. For simplicity, I assume that all fish in the escapement and terminal catch samples are aged without error. If they are not, the age proportions would have to be bootstrapped and included in the variance estimate.

Bootstrap samples must be generated at each step of the estimation sequence for ocean or mainstem catches. This consists of age specific samples for natural terminal catch, natural escapements, surrogate terminal catch and escapements, and surrogate ocean catch. Assuming B bootstrap samples are used, the bootstrapping process is

1. Generate B samples of tn_{ja}^* and tb_{ija}^* for watershed j and all ages $a = 2, \dots, 5$ using a multinomial distribution.

$$[tn_{j2}^*, tb_{ij2}^*, \dots, tn_{j5}^*, tb_{ij5}^*, to_j^*] \sim Mn\left(n_{Tj}, \frac{tn_{j2}}{n_{Tj}}, \frac{tb_{ij2}}{n_{Tj}}, \dots, 1 - \frac{tn_j + tb_{ij}}{n_{Tj}}\right)$$

to_j represents all of the other fish and can be discarded from each bootstrap sample.

2. Generate B samples of tb_{ila}^* for all watersheds $l \neq j$ and all ages $a = 2, \dots, 5$ using a multinomial distribution.

$$[tb_{il2}^*, tb_{il3}^*, tb_{il4}^*, tb_{il5}^*, to_l^*] \sim \text{Mn} \left(n_{T_l}, \frac{tb_{il2}}{n_{T_l}}, \frac{tb_{il3}}{n_{T_l}}, \frac{tb_{il4}}{n_{T_l}}, \frac{tb_{il5}}{n_{T_l}}, 1 - \frac{tb_{il}}{n_{T_l}} \right)$$

3. Generate B samples of $zn_{ja,\alpha}^*$ and $zb_{ija,\alpha}^*$ for watershed j and all ages $a = 2, \dots, 5$ using a multinomial distribution.

$$[zn_{j2,\alpha}^*, zb_{ij2,\alpha}^*, \dots, zn_{j5,\alpha}^*, zb_{ij5,\alpha}^*, zo_{j,\alpha}] \sim \text{Mn} \left(n_{T_j}, \frac{zn_{j2,\alpha}}{n_{E_j,\alpha}}, \frac{zb_{j2,\alpha}}{n_{E_j,\alpha}}, \dots, 1 - \frac{zn_{j,\alpha} + zb_{ij,\alpha}}{n_{E_j,\alpha}} \right)$$

$zo_{j,\alpha}$ represents all of the other fish and can be discarded from each bootstrap sample.

4. Generate B samples of $zb_{ila,\alpha}^*$ for all watersheds $l \neq j$ and all ages $a = 2, \dots, 5$ using a multinomial distribution.

$$[zb_{il2,\alpha}^*, zb_{il3,\alpha}^*, zb_{il4,\alpha}^*, zb_{il5,\alpha}^*, zo_l^*] \sim \text{Mn} \left(n_{T_l}, \frac{zb_{il2,\alpha}}{n_{T_l}}, \frac{zb_{il3,\alpha}}{n_{T_l}}, \frac{zb_{il4,\alpha}}{n_{T_l}}, \frac{zb_{il5,\alpha}}{n_{T_l}}, 1 - \frac{zb_{il,\alpha}}{n_{T_l}} \right)$$

5. Generate B samples of $zn_{ja,\beta}^*$ and $zb_{ija,\beta}^*$ using a multinomial distribution as in step 3, except with probabilities based on the hatchery escapement samples
6. Generate B samples of $zb_{ila,\beta}^*$ for all watersheds $l \neq j$ and all ages $a = 2, \dots, 5$ using a multinomial distribution as in step 4, except using hatchery escapement samples.
7. Generate B samples of xb_{ia}^* and yb_{ia}^* for all ages $a = 2, \dots, 5$ from multinomial distributions as above, except with respective probabilities $\frac{xb_{ia}}{n_O}$ and $\frac{yb_{ia}}{n_F}$.
8. Generate B samples of \widehat{C}_O^* , \widehat{C}_F^* , \widehat{C}_{Tj}^* , $\widehat{E}_{j,\alpha}^*$, and $\widehat{E}_{j,\beta}^*$.
9. Use these bootstrap samples to calculate \widehat{C}_{Onj}^* and \widehat{C}_{Fnj}^* using equations 23 and 24.
10. Calculate the standard errors of these estimates using the bootstrap results.

This is an example of how the variance can be calculated for the estimated natural ocean and mainstem catches in alternative NS1. Similar procedures can be developed for estimates in other alternatives, and are not given here.

6.2 Variance of Δ_{it}

Estimating the difference in the escapement due to incidental mortality is explained in Appendix A and depends on estimates of the natural process parameters initial survival and maturation. The variability of these parameters is discussed first, followed by a discussion of the variability in the Δ_{it} estimate. Following that, bootstrapping methods to estimate the variance of $\widehat{\Delta}_{it}$ are discussed.

6.2.1 Analytical variance calculation

Parameter estimates for four years are needed to estimate Δ_{it} and the variability of these estimates must be accounted for when determining the variance of Δ_{it} . By conditioning on the year specific estimates, the natural variation, or the year to year variation, does not need to be included, and an estimate of the variance for each estimated parameter from that year can be found.

The estimate of Δ_{it} is found using the estimates of the initial survival and maturation parameters (as seen in Appendix A). Therefore, by conditioning on the parameter estimates, the variance can be found using the conditional variance formula given in equation 47, but this time Y is a vector of the estimated parameters and X is Δ_{it} . The variance of Δ_{it} is the average variation in Δ_{it} over the range of the estimated parameters plus the variance of the expected values of Δ_{it} given the range of parameters. With multiple parameters making Y a joint distribution, the analytical variance can become quite complex and approximate methods to calculate the variance of Δ_{it} are recommended.

6.2.2 Approximate variance calculation

Again, bootstrapping can be done to estimate an approximate variance, but details are not given here. A possibility of what may be done is that the bootstrap estimates for the catch and escapement, as explained above, may be used to calculate initial survival and maturation parameters, and ultimately Δ_{it} . Also, as mentioned by Newman (2000a), the assumed known age 3, 4, and 5 survival parameters can be drawn from a reasonable distribution to introduce some of their variability. One hinderance to this method is that it is likely that a large number of bootstraps would be needed to accurately depict the total variation.

7 Assumptions

Before presenting an analysis of the methods described here, a few important assumptions for the tagging, sampling, and estimation procedures are discussed. First, Brownie et. al. (1985) mentions some general assumptions for tagging models that also apply here.

1. Samples are representative of the target population,
2. ages of individuals are correctly determined,
3. there is no tag loss,
4. survival rates are not affected by the tagging,
5. the year recovered is correctly tabulated,

6. the fate of each fish is independent of other fish, and
7. all individuals in the same cohort have the same life-history parameters.

In addition, simple random samples (SRS's) are assumed for the sampling of harvest and escapement. Thus, all temporal and spatial stratification is ignored. This oversimplification may result in the precision of catch and escapement data to be underestimated compared to the stratified samples usually taken from the fisheries and escapement.

A further assumption is that snouts are removed from every coded-wire tagged recovery. This is not controversial for harvest samples, but could be so for escapement samples of live fish. Also, when scanning for CWTs is necessary, it is assumed that the CWT is accurately detected and that the snout is removed for CWT extraction and reading. In other words, detection error, or sub-sampling of CWTs was not taken into account, although could be incorporated.

Because natural fish are unmarked and untagged, there must be an identifiable hatchery stock to serve as a surrogate for each natural stock. The surrogates are allowed to stray, but natural fish are assumed to have 100% fidelity. Per cohort, the surrogate hatchery stock and the natural stock are assumed to have the same overwinter survival rates, exploitation rates, and maturation probabilities. This is critical to the accuracy and precision of the production estimates and studies should be performed to determine if this is true. The initial survival rates may differ, however, because the estimation routines lead to a cancellation of these parameters. Conversely, the *Surrogate* and *Stealth* releases are assumed to have identical life history parameters.

Finally, maturation is not sex specific here, which may be an unreasonable assumption. Typically, only males return to freshwater as age 2 fish (commonly referred to as “jacks”).

8 Simulation and estimation program, *CFM Sim*

CFM Sim, short for Constant Fractional Marking Simulation, is an PC compatible computer program that simulates the following processes for multiple stocks of chinook salmon over multiple years:

1. the initial marking and tagging of fish, followed by natural mortality, fishing mortality, and maturation processes, each repeated for ages 2, 3, 4, and 5 (analogous to natural process variability),
2. the sampling of marine and freshwater catches and escapements, where the catches and escapements generally include a mixture of stocks (analogous to observation or sampling variability), and

3. the statistical estimation of catches and escapements for each stock separately, based on the catch and escapement sample data. The simulated estimates contain both natural process error and sampling error.

The program was designed to simulate these processes for hatchery-raised and naturally-spawned fall-run chinook salmon from California’s Central Valley using any of the four alternatives described above. The user can model hatchery and natural stocks in 14 different watersheds in the Central Valley simultaneously, with a limit of one hatchery and one natural stock per watershed. Multiple years can be simulated to determine trends over time, and any number of simulations can be performed to determine how the natural and sampling variability affect the production estimates.

CFM Sim is not a program that is to be used to determine the variability in the estimates after collecting and analyzing data. Instead it is a simulation program intended to be used as a planning tool when implementing a constant fractional marking program. The output allows one to predict the variability in the estimates and how that variability may change with changing parameters. Actually estimating the variances of the point estimates requires different methods, which were discussed in Section 6.

8.1 User input

The user must input parameters such as sampling, survival, and harvest rates, release sizes, and the constant fractional marking rate. A brief description of some parameters is provided here, and the reader is referred to Hicks & Newman (2000) for a more detailed look at *CFM Sim*. The main feature of *CFM Sim* is that the user may run simulations with different parameters to determine the effects on production estimates. For example, CFM rates of 20% and 40%, say, can be used, while keeping all other parameters constant. Other factors that can be manipulated within *CFM Sim* include the following:

- marine survival rates (ϕ_{iI} , ϕ_{i3} , ϕ_{i4} , and ϕ_{i5});
- age-specific ocean, freshwater, and terminal fishing harvest rates (μ_{Oia} , μ_{Fia} , μ_{Tiaj});
- sampling rate of catches in the marine fisheries (n_O/C_O);
- sampling rate of catches in the freshwater fisheries (n_F/C_F and n_{Tj}/C_{Tj});
- sampling rate of spawning escapement ($n_{E_{j,\alpha}}/E_{j,\alpha}$);
- sampling rate of the hatchery return ($n_{E_{j,\beta}}/E_{j,\beta}$);
- maturation rates (σ_{i2} , σ_{i3} , and σ_{i4}).

CFM Sim can also be used to analyze the effect of selective fisheries on the escapement of natural and hatchery bound Central Valley chinook salmon stocks. A user can vary any shaker mortality rate (Lawson and Sampson 1996), defined as a fraction of the harvest rate, to study the effects on both escapement and total shaker mortality. Drop-off mortality is not simulated in *CFM Sim* because it is not specifically a fraction of the harvest rate and is more difficult to simulate under the assumptions here.

A simple spreadsheet format is used to input the parameters which are specific to stock, age, and watershed, where applicable.

8.2 Simulation details

Each simulation begins by building up all age classes while holding recruitment constant. Once all of the age classes are present, *CFM Sim* begins the simulation of the observation processes and estimates production. A brief description follows, but the reader is referred to Hicks & Newman (2000) for a more thorough explanation.

Natural processes are mimicked in a number of ways. Hatchery releases are held constant over the years simulated, but given a starting recruitment value, the natural recruitment in each watershed is simulated using a constant stock-recruitment (S-R) relationship, a linear S-R relationship, a Ricker S-R curve, or a Beverton-Holt S-R curve, given user entered parameters (Quinn & Deriso 1999). A coefficient of variation is also supplied to the relationship and lognormal errors are applied each year. The life-history parameters, such as survival and harvest rates, vary each year, thus the user can enter a minimum, mode, and maximum for each parameter, which is then drawn from a triangle distribution. The triangle distribution was chosen because it is easy for a user to grasp the concept of a minimum, mode, and maximum, instead of asking them to enter a mean and variance, or similar, for each parameter. Related parameters are drawn from the same percentile in each of their distributions to simulate correlation between stocks, although this may make them more correlated than expected. For example, if a random number for survival is drawn in the lower part of its range, then the survival parameters for all stocks in that year will be low. After the parameters are drawn for that year, each event in the life-history sequence is simulated using a binomial distribution or multinomial distribution.

Observations are simulated once all of the age classes are present. This includes the simple random sampling of catches and escapement using a multinomial distribution, and the application of user entered aging errors. CWT detection rates are not yet modeled. Production estimates are then calculated, and Δ_{it} is calculated for alternatives AS1 and AS2 when enough years are present to estimate all of the necessary parameters.

8.3 Program output

The primary interest of *CFM Sim* is to determine the error and variability of estimated production. Therefore, estimated production is compared to “true” production in order to provide a measure of the quality of the estimates. *CFM Sim* uses the mean relative absolute error, labeled *MAE*, as this measure. To calculate the *MAE* for a given stock, the relative absolute error is calculated for each simulated year and then averaged across all the years of a single simulation.

$$MAE_s = \frac{1}{recYrs} \sum_{y=1}^{recYrs} \frac{|P_{s,y} - \hat{P}_{s,y}|}{P_{s,y}}, \quad (52)$$

where *recYrs* is the total number of recovery years, $P_{s,y}$ is the true production, and $\hat{P}_{s,y}$ is the estimated production, for simulation s and recovery year y . For example, suppose just two years were simulated. The true production was 10,000 and 11,000 for both years and the corresponding estimates were 9,500 and 11,200. Then

$$\begin{aligned} MAE &= \frac{1}{2} \left(\frac{|10,000 - 9,500|}{10,000} + \frac{|11,000 - 11,200|}{11,000} \right) \\ &= \frac{1}{2} (0.05 + 0.018) = 0.034 \end{aligned}$$

or a 3.4% relative absolute error for that single simulation. The two years would then be simulated again and another *MAE* would be calculated.

The average, median, standard deviation, minimum, and maximum of the *MAE*'s are calculated over all simulations and output by stock into new spreadsheets. These statistics give an idea as to the quality of the estimates of production and can be used to compare different marking, tagging, and sampling scenarios, for example. A relatively large average or median *MAE* indicates that production is not being estimated accurately, while a relatively large standard deviation shows that the estimates of production are not precise.

Also output to a spreadsheet are the summary statistics for the ratio of the last year's simulated production to the first year's simulated production for each natural stock. Ratios of true production and estimated production are reported and can be used to gauge the increase in natural production over the time period entered and if it is likely to be detected using the estimation procedures described here.

Some estimates may be negative in alternatives NS2 and AS2. When either of these alternatives are run, numbers of negative estimates that occurred are reported. This will help to determine how frequently absurd estimates may occur given the parameters entered.

When selective fishery alternatives are run and enough years are simulated, an additional worksheet outputs summary statistics for the mean relative absolute prediction error between the true Δ_{it} , and the estimated $\hat{\Delta}_{it}$ for each natural and hatchery stock.

A number of different external files are also created. The true and estimated production is reported for all stocks from the last simulation to give an insight into a simulation. The *MAE*'s for each stock and simulation are reported in case the summary statistics do not define the accuracy and precision well enough, and the total natural production for all recovery years and simulations is also reported. When requested, more detailed files for each stock are created that contain values such as true and estimated catches and escapement for each simulation.

9 Analysis of alternatives

CFM Sim was used to analyze the effects that various parameters have on the accuracy and precision of natural production estimates for the four marking alternatives. Newman (2000a) used a simulation program to study the effects of various parameters on the quality of production estimates, but he made some assumptions to simplify the simulations and did not study alternatives NS2 and AS2.

For the simulations presented here, the hypothetical system consisted of three watersheds. Hatcheries in watersheds 1 and 2 were meant to mimic Nimbus Fish Hatchery and Coleman National Fish Hatchery, respectively, at least in their release sizes. The third watershed contained only a natural stock and no hatchery. All stocks were susceptible to ocean and freshwater mainstem fisheries, and there was a terminal fishery only in watershed 1.

A single year was simulated to study the effects of certain parameters on the production estimates. Different designs were used for the no selective and all selective fisheries, thus they are explained separately. The results of the simulations are discussed for each alternative, as well as a comparison between alternatives NS1 and NS2, and a comparison between alternatives AS1 and AS2.

Seven years were simulated to study the effects of the treatment parameters on the estimate of Δ_{it} . These simulations were done separately from the single year simulations because at least seven years are necessary to estimate survival and maturation parameters.

The fixed parameters input into the simulations are discussed first, then the analyses.

9.1 Fixed parameters

The three natural stocks each had 300,000 initial recruits, where recruitment is defined as the point immediately prior to initial survival in the sequence of events. The hatchery in watershed 1 released 4 million fish and the hatchery in watershed 2 released 12 million fish. *Ad hoc* releases were set at 40,000 and 960,000 for the hatcheries in watershed 1 and 2, respectively, and the *Surrogate* levels were varied as treatment effects. The *CFM* and *Remaining* release groups were calculated based on the *CFM* rate, total release, *Ad hoc* release, and the number of *Surrogates*.

Hatchery 1 was a surrogate to the natural stock in watershed 1, while hatchery 2 was a surrogate to the natural stocks in watersheds 2 and 3. Each natural stock had the exact same parameters as its surrogate, except for straying rates. Natural stocks did not stray in these simulations.

The only parameters of interest in these analyses were the CFM rate, the escapement sampling rate, fishery sampling rates from the freshwater mainstem and terminal fisheries, number of *Surrogates* released, the shaker mortality rate in selective fishery alternatives, and the number of *Stealth* fish released in the selective fishery alternatives. All of the other parameters were fixed. Parameters that are subject to natural random variation, such as survival, maturation, and harvest rates, were given fixed minimum, mode, and maximum values to define a triangle distribution from which the year specific parameters were drawn.

The five stocks of fish had mostly the same parameters, with few differences. The minimum, mode, and maximum of the survival, ocean harvest, and freshwater mainstem harvest parameters were fixed at the same values for each stock (Table 5). There was no ocean harvest of age 2 fish, but freshwater mainstem harvest was equal for all ages. The harvest rates reflect the number of fish actually landed and do not include drop off mortality because *CFM Sim* does not simulate it. A shaker mortality is simulated in the selective fishery alternatives, however. Two different sets of maturity parameters were used for the two hatchery stocks and the natural stocks were given the same values as their surrogate stock (Table 6). Both hatchery stocks strayed to other watersheds (Table 7), but the natural stocks did not.

The terminal fishery in watershed 1 had min, mode, and max harvest rates of 10%, 20%, and 30%, respectively, for all ages, and the terminal harvest rates for watersheds 2 and 3 were set equal to zero. The sampling rates of the total ocean harvest and the ocean sport catch were fixed at 20%, and the percentage of ocean fish caught in the recreational fishery was set at $\frac{1}{3}$ based on recovery information from the Pacific Fishery Management Council (2001). The percentage of fish sampled for CWTs in the hatchery was set to 100%. In addition, no aging error was assumed. The simulations used to study the effects on production simulated only one year of production, thus spawning parameters were not necessary. However, separate simulations done to study the estimates of Δ_{it} required seven years of data to be simulated, and used constant recruitment.

The main difference between the three natural stocks is the surrogate stock they use. Throughout the analysis, each natural stock may be referred to by its watershed number. For example, the natural stock from watershed 1 will simply be called Watershed 1.

9.2 No selective fishery alternatives

9.2.1 Design

The effects of five parameters on the production estimates for the no selective fisheries alternatives were studied using a random effects 2^5 factorial design. These parameters were:

CFM, the constant fractional marking rate,

ESR, the escapement sampling rate,

FSR, the freshwater mainstem harvest sampling rate,

Sur, the number of surrogate fish released, and

TSR, the terminal harvest sampling rate.

The levels for each parameter in the 32 experiments is given in Table 8.

Each scenario was simulated 1000 times within each alternative. The response variable was the mean relative absolute prediction error ($MAE_{i,j}$), calculated using equation 52. However, since only one recovery year (one year of measured production) was simulated in this study, the MAE becomes

$$MAE_s = \frac{|P_s - \hat{P}_s|}{P_s},$$

where s refers to the simulation. MAE without the subscripts, as used here, refers to the general absolute error for a stock over all simulations. Only the production estimates and the MAE of the natural stocks were used in the analyses, and each stock is always analyzed separately.

For each watershed, the bias of the estimates is considered first using the relative error of the production estimates. With only one recovery year simulated the relative error is

$$RE_s = \frac{P_s - \hat{P}_s}{P_s},$$

where s denotes the stock. Non-parametric densities of the relative error are used to display the distribution of the errors, and a one-sample t-test was done for each treatment level to test if the mean was significantly different from zero. The significance level of 0.05 was adjusted using the Bonferroni correction to control for experiment-wise error rate.

Then, to determine the size and significance of the effects, an analysis of variance (ANOVA) using a Box-Cox recommended transformation was adopted to deal with heteroskedasticity and

non-normal residuals. The Box-Cox transformation on the MAE was

$$MAE^* = \frac{MAE^\lambda - 1}{\lambda} \quad (53)$$

where λ is the recommended transformation parameter. The purpose of this analysis was to identify the effects that are most important to the precision of the production estimates. Therefore, only second-order terms were included in the model, so that a better understanding of individual effects and their direct interactions with other variables could be understood. It was felt that the inclusion of higher order terms would interfere with a simple understanding of the variables and their importance. Any interactions that are important to understanding the variables are likely to appear in the second-order terms.

The full model including all second-order interactions was first tested and insignificant interactions were pooled into the error term following the methods described by Kirk (1995). That is, given the full model, the two-way interactions with a p-value less than 0.1 were removed. Each variable in the reduced model was then determined to be significant or not at the $\alpha = 0.05$ level by using the correct denominator to calculate the F-statistic. A random effects model was assumed, thus the denominator is not necessarily the mean squared error when second-order terms are present. Therefore, if necessary, a quasi-F statistic was calculated by adding and subtracting appropriate mean squares to create the correct error term.

Simple main effects were tested when interactions were present. These are tests on the difference in means at a fixed level of the other variable in the interaction. Not all tests were performed, but interesting ones were chosen based on the ANOVA results and interaction plots. To also help explain the important variables, effect coefficients, or slopes in a regression context, were calculated to indicate the magnitude and direction of the significant effects.

After the results of alternatives NS1 and NS2 are discussed, the two are compared.

9.2.2 Production estimates using alternative NS1

The relative error for alternative NS1 ranged from -1 to greater than 8, and cannot be less than -1 because negative estimates of production are not possible. Because of this, the distribution can become highly skewed when estimates are much greater than the true natural production. Figure 3 shows smoothed non-parametric density plots of the relative error for each watershed, which were all right skewed. Watershed 1 had lower estimation errors than the other two watersheds, on average, which may be because the terminal fishery adds extra information to the estimation of ocean and mainstem catches. For most treatment combinations, the t-test showed that the mean was significantly not equal to zero, and in all cases for all watersheds the estimated mean was greater than zero. Also, as Table 9 shows, there was a greater tendency to overestimate the natural production.

Before giving the results of the ANOVA, the plots in Figure 4 show that the escapement sampling rate and the number of surrogates released seemed to have the most influence on the precision of the estimates. The skewness of the MAE can also be seen in these plots. The mean and median MAE 's (in percent) are given in Table 11, showing the dramatic decrease due to an increase in the escapement sampling rate.

NS1: Watershed 1 production

The Box-Cox recommended transformation was 0.32 for the data from Watershed 1. The residuals appeared mostly normal, but were still slightly skewed to the right, and some heteroskedasticity was apparent, which was not removed using a weighted regression approach. Nevertheless, there were two significant interactions: $ESR-Sur$ and $ESR-TSR$. Using the correct error term to test the significance of the main effects resulted in only ESR being significant.

This is the only watershed with a terminal fishery and it was thought that there may be an interaction between the terminal and escapement sampling rates, because the terminal and escapement estimates are both used to estimate the natural ocean and mainstem catch (equations 23 and 24). Therefore, before beginning the ANOVA, it was determined that any interaction between ESR and TSR would be further tested. Significant simple main effects showed that an increase in ESR at either level of TSR decreased the MAE , although a lesser effect was seen at a low TSR . A change in TSR was not significant at either level of ESR .

Because Sur was a nearly significant main effect, tests were performed to see if the number of surrogates affected the MAE at separate levels of the escapement sampling rate. After applying a Bonferroni correction to the level of significance for doing two a posteriori tests, no significant differences were found in the mean MAE when increasing Sur . However, the difference was nearly significant at the high escapement sampling rate.

The coefficients of all main effects, although only ESR was significant, were negative, indicating that an increase in any variable lowers the MAE . The coefficient for ESR was by far the largest.

NS1: Watershed 2 production

The Box-Cox transformation used for Watershed 2 was 0.3. The residuals were adequately normalized and stabilized, and a weighted model did not change the results. Therefore, the analysis followed a normal ANOVA approach using the transformed MAE .

All but four second-order terms were pooled into the error. The interactions kept in the reduced model were: $ESR-Sur$, $ESR-FSR$, $CFM-FSR$, and $CFM-Sur$. However, only the interactions involving CFM were significant with the p-values for the ESR interactions near

0.09. The only significant main effect was *ESR*, which had a p-value of 0.047. The *Sur* main effect was again slightly not significant with a p-value of 0.08. In summary, the significant effects were: *ESR*, *CFM-Sur*, and *CFM-FSR*. Tests on the simple main effects suggested that the variables involved in the interactions (*CFM*, *Sur*, and *FSR*) were not significant when fixing on a level of the other variable in the interaction.

The coefficients for all of the significant effects were negative, indicating that an increase in the variables results in a decrease in the estimation error. *ESR* and *Sur* had the two largest coefficients.

NS1: Watershed 3 production

Similar to Watershed 2, the Box-Cox transformation for Watershed 3 was 0.3, which appeared to stabilize and normalize the residuals. Two second-order effects were significant: *ESR-FSR* and *Sur-TSR*. The only significant main effect was *ESR*, with a p-value of 0.04, and as before, the main effect *Sur* was slightly not significant with a p-value of 0.07. Therefore, in this watershed, all of the variables, except for *CFM*, had some significant influence on the precision in production estimates, whether directly or through an interaction.

The coefficient for the significant *ESR* main effect was again negative, but the interaction coefficients were both positive, although barely significant with p-values greater than 0.029. Interaction plots in Figure 5 show that the *MAE* is not reduced as much by *ESR* at a high *FSR*, and similarly, the *MAE* is not reduced as much by the number of surrogates at a high *TSR*. Simple main effects tests showed that the decrease in the *MAE* was not significant when holding *Sur* at 50,000 and increasing *TSR* from 10 to 30%. However, the decrease in *MAE* due to an increase in the number of surrogates was slightly significant when the terminal sampling rate was low, and not significant when *TSR* was high.

The estimated coefficients for the main effects were negative, except for the *CFM* rate, which was slightly positive, but not significant. Again, *ESR* was the largest in magnitude, followed by *Sur*.

9.2.3 Production estimates using alternative NS2

The estimated production estimates for alternative NS2 showed slightly different results between Watershed 1 and Watersheds 2 and 3. The range of the relative error was -0.8 to 2.1 for Watershed 1, but increased greatly for Watersheds 2 and 3 with a minimum of -1 for both and a maximum of 6.3 and 8.1 for them, respectively. Non-parametric densities of the relative error for each watershed are shown in Figure 6.

The estimated production for all three natural stocks was greater than the true production slightly more often than not (Table 9). And, as with alternative NS1, the mean relative error

over all simulations for a set of treatment levels was always positive. The t-tests to determine if the mean relative error was significantly different from zero showed significance in all but two experiments for Watershed 1. Bias was seen in 22 and 23 scenarios for Watersheds 2 and 3, respectively. Specific patterns to which scenarios were rejected in the t-test could not be found.

The number of negative escapement and terminal catch estimates are also of interest in alternative NS2 (Table 10). The escapement estimate in Watershed 1 was negative in only 1 simulation of the 32000, but the terminal catch was more often estimated as a negative number, especially when the terminal sampling and constant fractional marking rates were low. A generalized linear model (GLM) with a log link for Poisson data (McCullagh & Nelder 1989) confirmed that *CFM* and *TSR* were the most important variables determining the number of negative terminal catch estimates for Watershed 1.

The number of negative escapement estimates in Watershed 2 had a high occurrence when the constant fractional marking rate was low. A GLM confirmed this, and also showed that the escapement sampling rate had some significance on the number of negative estimates, while *Sur* and *TSR* showed a slight significance. Freshwater mainstem sampling, as expected did not seem to affect the number of negative estimates.

Watershed 3 also had a large number of negative escapement estimates when the CFM rate was low, but was considerably reduced when the *ESR* was high. A GLM showed that only *CFM* and *ESR* were significant in reducing the number of negative estimates of escapement for the natural stock from this watershed.

The median *MAE* as well as the 10th and 90th percentiles from the simulated distribution are plotted in Figure 7 to show how each variable alone affected the *MAE*. An increase in the escapement sampling rate appeared to lower the *MAE*. However, in contrast to alternative NS1, the CFM rate seemed to have some importance on the precision of the natural production estimates. The mean and median percent *MAE*'s, conditioned on each variable level, are given in Table 11.

NS2: Watershed 1 production

Although the recommended Box-Cox transformation of 0.32 satisfactorily normalized the residuals, a slight amount of heteroskedasticity was observed when plotted against fitted values. Including weights in the model did not change the observed pattern, and an analysis of variance without weights was used to gain insight into the important effects.

The ANOVA found a very strong interaction between *TSR* and *ESR*, as expected, and simple main effects tests were decided a priori. In addition to this interaction, four others were kept in the model: *CFM-ESR*, *CFM-Sur*, *CFM-TSR*, and *Sur-TSR*. The *Sur-TSR* interaction was not significant but had a p-value less than 0.1. Accounting for these interactions,

no main effects were significant. However, some insight into the interactions will display the important variables.

Figure 8 shows the six possible significant interaction plots. The interactions $ESR-CFM$, $CFM-Sur$ and $ESR-TSR$ showed greater differences in means when both variables were at high levels. The interaction $CFM-TSR$ did not have as great of an effect when the other variable was at a high level. A priori tests on the difference of means were done for the four scenarios of the $ESR-TSR$ interaction. A significant difference in the means was not found for an increase in TSR at either level of ESR . A posteriori tests of the simple main effects were done for the following combinations of the other interactions: CFM at $ESR = 30$, ESR at $CFM = 50$, TSR at $CFM = 20$, and CFM at $TSR = 10$. The significance level was adjusted to 0.0125, and none of these four tests were significant. However, all but TSR at $CFM = 20$ would have been significant at the uncorrected $\alpha = 0.05$ level.

NS2: Watershed 2 production

The recommended transformation was 0.31, which normalized and stabilized the residuals. The model was reduced to contain the main effects and two second-order effects: $ESR-Sur$ and $ESR-FSR$. The interaction $ESR-FSR$ was not significant at the 0.05 significance level, but was kept in the model because the p-value was less than 0.10.

The only significant effects were CFM and the $ESR-Sur$ interaction. The escapement sampling rate and number of surrogates were nearly significant with negative estimated effects much smaller than the significant effect for CFM . The $ESR-Sur$ interaction was slight and showed that the escapement sampling rate decreased the MAE more when the surrogate level was low. However, simple main effects tests did not find any significant differences in means when either variable was held constant.

All of the estimated coefficients for the main effects were negative, except for FSR .

NS2: Watershed 3 production

With a Box-Cox transformation of 0.30, two second-order terms were significant: $CFM-ESR$ was slightly significant and $Sur-TSR$ was highly significant. CFM and ESR were highly significant main effects, and surprisingly FSR was slightly significant with a positive coefficient. Because no interaction terms appeared in the model with FSR , the denominator degrees of freedom for the F-test was very large resulting in the detection of a significant difference. It seems unlikely that an increase in FSR would worsen the precision, and the main effects CFM and ESR appeared to be the most influential variables, each with negative coefficients.

The interactions plotted in Figure 9 showed a larger decrease in the MAE when the terminal sampling rate was high and the number of surrogates is increased. This may have to do with

a small number of straying surrogates caught in the terminal fishery of watershed 1 not being adequately sampled until the sampling rate is large enough. The *CFM-ESR* interaction was very slight and almost unnoticeable, but the main effects were significant.

Tests on the difference in means were done for the four combinations of the *Sur-TSR* interaction and no significant differences in the means were found.

9.2.4 Comparison of NS1 and NS2

The simulated true production for both alternatives were the same within each stock, and because the life-history parameters were not varied, the true production was similar for each scenario. Watershed 1 had a mean true production of 6728 while Watersheds 2 and 3 simulated similar mean true productions of about 6380. However, the two alternatives showed prominent differences in the quality of production estimates. First, the error in the estimates were larger when using alternative NS2 (see Table 11 and Figures 3 and 6). Although the ranges of the MAE for the watersheds were nearly the same when using either alternative, the median MAE was over 4 times greater when simulating alternative NS2 at a low CFM rate. However, a high CFM rate significantly reduced the median *MAE* making the difference between the two alternatives smaller. Second, the CFM rate was a significant effect in alternative NS2. The increased number of marked and tagged fish made a difference because all hatchery fish are not marked. Finally interactions were more common in the analysis for NS2. It appears that many variables interact with each other, possibly because each has a greater effect on the *MAE*.

There were also some similarities between the two alternatives. The escapement sampling rate was a significant factor on the precision, and as expected for Watershed 1, the terminal sampling rate interacted with the escapement sampling rate. Also, the estimates of natural production were typically most precise for Watershed 1 in both alternatives, while Watersheds 2 and 3 had similar errors.

9.3 All selective fishery alternatives

The same hypothetical stocks as before were used to generate production estimates under the selective fishery alternatives AS1 and AS2. An additional consideration for the selective fishery alternatives is the shaker mortality rate, which was included as a factor in the design. Estimating Δ_{nt} was also considered, which is discussed after the analysis on the precision of the production estimates.

9.3.1 Design

Seven parameters were simulated at two levels, but the three freshwater sampling rates (terminal and freshwater fisheries, along with escapement) were combined into one overall sampling rate, resulting in a full 2^5 factorial design (Table 12). The five parameters considered were:

CFM, the constant fractional marking rate,

SR, the freshwater, terminal, and escapement sampling rates,

SM, the shaker mortality rate,

Sur, the number of surrogate fish released, and

Stl, the number of stealth fish released.

The shaker mortality rate was the same for all age classes and the three fisheries. Each possible combination of treatment levels is labeled 1 through 32, as shown in Table 12.

Again, 1000 simulations were performed for each of the 32 experiments within the two alternatives. An analysis of variance following the procedures described in Section 9.2.1 was also done for the three watersheds in each of these alternatives. The response variable was the mean relative absolute prediction error ($MAE_{i,j}$) from equation 52.

Separate sets of simulations were performed using alternatives AS1 and AS2 to study the estimates of Δ_{nt} and how the precision is affected by the five variables described above. The previous simulations used to study production estimates simulated only one recovery year. However, seven years were simulated to get one estimate of Δ_{nt} . Constant recruitment of 300,000 was used for each of the seven years to avoid natural population crashes due to poorly picked stock-recruitment parameters or unusually poor recruitment. The same design was used as above and 1000 simulations were performed for each scenario. The output of *CFM Sim* limited the analysis to only studying the *MAE*.

The two alternatives, AS1 and AS2, are presented separately, with the bias studied first. Then, the results of the ANOVAs are discussed. Finally, a comparison between the two alternatives is made.

9.3.2 Production estimates using alternative AS1

The natural production in alternative AS1 consists only of the escapement because all fisheries are selective. Also, because all hatchery released fish are marked, the escapement estimates of natural fish are simple expansions of the sample values (equation 22) and were unbiased. The relative error also had a more symmetric shape than in the no selective fishery alternatives (Figure 10). Only one t-test of the 96 showed that the relative error was significantly different

from zero. The relative error ranged from -0.40 to 0.46 for all the natural stocks, with the natural stock from watershed 1 having the largest range, due to some outliers.

Figure 11 shows the median MAE with the 10th and 90th percentiles plotted for each variable and natural stock. It appears that an increase in SR results in a reduction in the MAE , and an increase in SM will increase the MAE . The mean and median MAE 's (in percent) are given in Table 14, conditioned on the level of each variable. The ANOVAs discussed next will give some insight into the significance of the variables.

AS1: Watershed 1 production

The recommended Box-Cox transformation was 0.14 for this natural stock, which left the residuals slightly non-normal and increasing over the range of fitted values. ANOVA was still done to determine influential variables.

The only significant interaction was between SR and SM , and given this interaction, no main effects were significant, although the p-value for SR was 0.057. Therefore, simple main effect tests were done at each of the two levels of each variable, resulting in four tests. The tests on the sampling rates were significant at each level of the shaker mortality rate, but SM did not have any significant effects on the MAE , although a increase in SM would have significantly increased the MAE at the 10% SR level with an uncorrected α level. As expected, an increase in SR decreased the MAE , and this increase was greater when the shaker mortality rate was high.

The coefficients of the SR and SM main effects were negative and positive, respectively. The interaction coefficient was also positive and adding the SR , SM , and SR - SM coefficients together (as when SR and SM are at their respective high values), the overall value was still negative, indicating that increasing both variables simultaneously will result in an increase in the precision.

AS1: Watershed 2 production

A transformation of 0.31 adequately normalized and standardized the residuals for the natural stock in Watershed 2. The SR main effect was significant in this analysis, but the SM main effect was not. However, as with Watershed 1, the SR - SM interaction was again significant, although slight.

The four simple main effects tests were again done on the significant SR - SM interaction. Using the corrected significance level for the four tests, only the test on SR at $SM = 30$ was significant, although slightly. If the correction was not done, the only test that would have been not significant was SM at $SR = 30$. The best interpretation of this interaction is that when the shaker mortality rate is high, the sampling rates have a greater affect on the precision of

the natural production estimates.

The coefficient for the SR effect was negative, and when both SR and SM are at their respective high levels, the overall coefficient is still negative.

AS1: Watershed 3 production

Using a transformation of 0.32 adequately normalized and standardized the residuals, and resulted in a reduced model containing only the SR – SM interaction, as with the analysis for the other two stocks. The SR main effect was significant and the SM main effect was nearly significant with a p-value of 0.077. The interaction was similar to the interactions in Watersheds 1 and 2.

None of the simple main effects tests were significant in the SR – SM interaction, but as with Watershed 2, if the significance level correction was not made, the only difference in means that would be not significant was SM at $SR = 30$.

As before, when SR and SM are at their respective high levels, the overall coefficient would be negative, suggesting that increasing sampling rates influence the production estimates by increasing the precision, and high shaker mortality rates take away some precision.

9.3.3 Δ_{nt} estimates under alternative AS1

The medians and 10th and 90th percentiles for the MAE when estimating Δ_{nt} under alternative AS1 are shown in Figure 12. The mean and median MAE 's for each watershed are given in Table 15. The variables SR , SM , and Sur all seemed to have a positive effect on the precision of the estimates.

AS1: Watershed 1 Δ_{nt}

A Box-Cox transformation of 0.29 was used for Watershed 1 and the reduced model for the MAE of Δ_{nt} included three interactions: SR – SM , SR – Stl , and Sur – Stl . The Sur – Stl interaction was barely significant with a p-value of 0.049. The residuals slightly increased over the range of fitted values, and weighting each cell by the inverse of the variance did not help. The main effect SM was significant and Sur was nearly significant. Four simple main effects were tested to see if a significant reduction in the MAE occurred for any of the variables. These tests were SR at $SM = 10$, Sur at $Stl = 100,000$, SR at $Stl = 50,000$, and Stl at $SR = 10$. None of the tests were significant, however.

AS1: Watershed 2 Δ_{nt}

Transforming the data with $\lambda = 0.31$ normalized the residuals, but left them slightly increasing when plotted against the fitted values. The interaction $Sur-Stl$ was slightly significant and the interaction $SR-Stl$ was not significant, but had a p-value less than 0.10. The only significant main effect was SM , although SR and Sur were nearly significant. Two simple main effects were tested: Sur at $Stl = 100,000$ and SR at $Stl = 50,000$. Neither were significant when adjusting the significance level for the two tests, but they both would have been without the adjustment.

AS1: Watershed 3 Δ_{nt}

The MAE of Δ_{nt} for Watershed 3 was also transformed with $\lambda = 0.31$ and the residuals seemed to slightly increase with fitted value. The only interaction kept in the model was $SR-Stl$, and because no interaction with surrogates was present, Sur was a significant main effect along with SM . SR was nearly significant. One simple main effect tested for SR when Stl was low was not significant.

9.3.4 Production estimates using alternative AS2

Alternative AS2 has selective fisheries for all fisheries, thus the natural production consists solely of escapement. In this alternative, however, all the hatchery fish are not distinguishable from natural fish, resulting in more complicated estimation procedures.

The difference from alternative AS1 in terms of production estimates can be immediately seen in Figure 13. These non-parametric density plots show that the relative errors are larger than in alternative AS1, ranging from -1 to 1.9, and that there are quite a few estimates of zero production. Watershed 2 shows the most unusual distribution of relative errors, with a large number of zero production estimates and some very large overestimates. However, even with these abnormal densities, bias was typically non-existent. Few t-tests on the mean relative error being different from zero were significant for Watersheds 1 and 3. About half of the t-tests were significant for Watershed 2. Table 13 shows the number of times that the production was underestimated for each watershed and scenario. The production is underestimated slightly more often than overestimated (Table 13), which may have been a result of zero estimates (typically arising from negative estimates set equal to zero).

Negative estimates are possible in this alternative and Table 16 shows the number of negative escapement estimates for each watershed. All watersheds had some negative estimates with Watershed 2 having the most. A GLM using a log-link containing no interactions was used to determine if any variables had an influence on the number of negative estimates. In Watersheds 1 and 2, all but the number of surrogates released had an influence. In Watershed 2 all five

variables were significant. The number of negative estimates were reduced by increasing *CFM*, *SR*, or *Stl*, and were inflated by increasing *Sur* and *SM*.

The median *MAE* with the 10th and 90th percentiles, plotted for each variable and natural stock, are shown in Figure 14. As with alternative AS1, *SR* and *SM* appeared to have opposite effects on the *MAE*. However, in this alternative *CFM* and *Stl* seemed to be influential also. The mean and median *MAE* (in percent) for each level of each variable is given in Table 14.

AS2: Watershed 1 production

The Box-Cox transformation used for Watershed 1 was 0.36 and the residuals appeared to satisfy the assumptions. All but three interactions were significant, leaving seven in the model. With these interactions in the model, the main effects were not significant, but they affected the *MAE* in complicated ways through the interactions with other variables. It is important to note that the *SR-SM* interaction that was significant in the AS1 alternatives was barely significant here with a p-value of 0.049.

It can be difficult to identify the important variables when many interactions are present thus I focused on the two significant interactions with *CFM* and the two significant interactions with *Sur* (*CFM-SR*, *CFM-Sur*, *SR-Sur*, and *Sur-Stl*). They were the strongest interactions (as defined by the lowest p-values) in the model. The plots in Figure 15 show these four interactions.

Tests for differences in the means were done on *SR* at each level of *Sur*, *Stl* at each level of *Sur*, *CFM* at each level of *Sur*, and *SR* at each level of *CFM*. None of the tests were significant when adjusting for the eight multiple comparisons. However, most would have been significant as a priori tests. Nevertheless, the plots in Figures 14 and 15 show that *CFM*, *SR*, and *Stl* were influential variables on the precision of natural production estimates. Coefficient estimates for the main effects are not reported here because they have little meaning with all of the second-order effects present in the model.

AS2: Watershed 2 production

A λ of 0.42 was used to transform the *MAE* for Watershed 2. This residuals were somewhat normalized, but their variance increased slightly when plotted against the fitted value. Using weights in a regression approach did not stabilize them any more than the transformation had already done. Nevertheless, analysis of variance was done to tease out the influential variables and their interactions.

All but three interactions were significant: *SR-SM*, *SR-Stl*, and *SM-Stl*. This left seven interactions in the model and *CFM* had an interaction with every other variable. No main effects were significant, although *CFM* and *Stl* were nearly significant (p-values = 0.09 and 0.06, respectively). As with Watershed 1, four of the most significant interactions were chosen

to see how they influence the MAE . These were $CFM-SR$, $CFM-SM$, $CFM-Sur$, and $SR-Sur$, which are plotted in Figure 16. Four a posteriori tests were done to determine the significance of specific main effects: CFM at the low levels of SR , SM , and Sur ; and SR at the low level of Sur . None of the tests showed significance, but the plots in Figure 16 indicate that CFM and SR may have some influence on the MAE . Also, the plot in Figure 14 and the near significance of Stl as a main effect indicate that increasing the number of *Stealth* fish may increase the precision of the production estimates.

AS2: Watershed 3 production

Watershed 3 used a transformation of 0.39, which normalized, but did not stabilize the residuals. A weighted regression approach also did not adequately stabilize the residuals when used in conjunction with the transformation. Going ahead with an analysis of variance, the results suggested that all but three interactions were significant. The three interactions that were not significant were: $CFM-SM$, $CFM-Stl$, and $SR-Sur$. Again, none of the main effects were significant, but CFM , SR , and Stl all had p-values less than 0.1.

The four most significant interactions ($CFM-SR$, $CFM-Sur$, $SR-Stl$, and $Stl-Sur$) were chosen to study further and are displayed in Figure 17. Three a priori simple main effects tests were chosen to determine if CFM , SR , or Stl had some significance on the production estimates. The specific tests were SR at a low CFM , CFM at low Sur , and Stl at high Sur . None were significant. The sampling rates at a low CFM rate were nearly significant, however, and they all would have been significant without adjusting the significance level. Along with Figure 14, these results suggest that CFM , SR , and Stl influence the MAE , and they each interact with many other variables.

9.3.5 Δ_{nt} estimates under alternative AS2

The MAE 's for Δ_{nt} when estimated with alternative AS2 were very large, ranging from near zero to over 50 for Watershed 2. That is a 5000% error! Table 15 gives the mean and median percent MAE 's for the different levels of the variables. Figure 18 shows that SM affects the MAE of Δ_{nt} more than any other variable, although slight effects may be present due to CFM , SR , and Stl .

AS2: Watershed 1 Δ_{nt}

The Box-Cox transformation of 0.36 did not adequately normalize or stabilize the residuals, but analysis of variance was still done. Five interactions were significant: $CFM-SM$, $CFM-Sur$, $CFM-Stl$, $SR-SM$, and $SM-Stl$. The only significant main effect was SM . Looking at interaction plots, three simple main effects tests were done to determine if CFM or Stl had

any significance at specific levels of other parameters. These were *CFM* at high *Stl*, *Stl* at high *CFM*, and *CFM* at low *Sur*. None were significant. The violation of assumptions in this analysis makes the tests questionable, but it does appear that *SM* had a large influence on the precision of Δ_{nt} estimates.

AS2: Watershed 2 Δ_{nt}

For the most part, transforming the data with $\lambda = 0.19$ stabilized and normalized the residuals. Five interactions were kept in the model (*CFM-SR*, *SR-SM*, *SR-Stl*, *SM-Sur*, and *SM-Stl*) and all were significant. The main effect *SM* was just significant with a p-value of 0.05. Three simple main effects tests for *CFM* at low *SR*, *SR* at low *CFM*, and *SR* at low *Stl* were not significant. *SM* was the main influence on the error in estimates of Δ_{nt} by decreasing the prediction error when *SM* increases.

AS2: Watershed 3 Δ_{nt}

The residuals were mostly stabilized and normalized by a Box-Cox transformation where $\lambda = 0.29$. The six interactions *CFM-SR*, *CFM-SM*, *SR-SM*, *SR-Sur*, *SR-Stl*, and *SM-Stl* were all significant in the reduced model, but no main effects were significant (*SM* was close with a p-value of 0.065). Of four simple main effects tests, *SR* at low *Sur*, *SR* at low *CFM*, *SM* at high *SR*, and *SM* at high *Stl*, none were significant. Even though no effects were significant for this watershed, the variables interacted with each other and *SM* appeared to influence the *MAE*, although not significantly.

9.3.6 Comparison between alternatives AS1 and AS2

The simulated true production varied for these alternatives because the shaker mortality was a treatment. It remained the same between alternatives, however. When the shaker mortality was 30% the mean simulated true production was about 4300 for Watershed 1 and about 3450 for Watershed 2 or 3. Decreasing the shaker mortality rate to 10% increased the true production to 5450 and 4670 for Watershed 1 and Watershed 2 or 3, respectively.

The precision in the natural production estimates differed greatly for the two all selective fisheries alternatives (Figures 10 and 13, and Table 14). The natural production in these alternatives consisted of only escapement because the selective fisheries do not keep unmarked fish. Because of this, alternative AS1 used a simple expansion of the sample means to estimate natural production, resulting in a high precision for the natural production estimates. On the other hand, alternative AS2 estimates the unmarked hatchery fish in the escapement, and subtracts it from the total escapement, resulting in much less precise estimates than in alternative

AS1. When using a low CFM rate, the median *MAE* in Watershed 2 under alternative AS2 was almost 50 times greater than the median *MAE* under alternative AS1.

Another difference between alternatives AS1 and AS2 was that the CFM rate became an important, although not significant, factor on the precision of the production estimates. Also, the number of stealth fish released showed some importance in alternative AS2. The analysis of the *MAE* for natural production in AS1 was straight forward and only one interaction was significant for all three watersheds. In contrast, many interactions were significant in the analysis of alternative AS2, indicating that other variables have strong influences on the precision. The sampling rates were important in both alternatives, but other factors became important when using alternative AS2.

The estimates of Δ_{nt} were much less precise than the production estimates and the *MAE* differed for the two alternatives (Table 15). The precision of the Δ_{nt} estimate was very poor, with a median *MAE* of up to 749 at a low shaker mortality rate using alternative AS2. The *MAE*'s for alternative AS1 were much smaller. The shaker mortality rate was the most significant effect, regardless of the alternative used. However, under alternative AS2, the number of surrogates and the sampling rates were more important to reducing the *MAE* of the Δ_{nt} estimates than in alternative AS1.

Overall, the sampling rates were again influential variables on the *MAE*. However, the range of sampling rates used here did not detect a significant effect. Also, the shaker mortality rates affected the precision of the natural production and Δ_{nt} estimates differently. However, shaker mortality rates may be of less use than sampling rates, say, because they cannot be as easily controlled.

10 Discussion

A constant fractional marking program is a standardized method to estimate the hatchery and natural production of the Sacramento–San Joaquin river system. The estimation equations given here were based on simple means expansion and method of moments, which resulted in some shortcomings. The most obvious was the possibility of negative catch or escapement estimates when all hatchery fish are not marked or tagged. The equations also became quite complicated when estimating the effects of incidental mortality.

In addition, the assumptions used here may not be valid. Samples from the catches and escapements are actually stratified and would result in a lower variability in the estimates. The binary sequence of events may not be realistic (e.g., mortality may occur throughout the year). And, there may be natural mortality in the freshwater system, which was not accounted for. However, these assumptions were necessary to display the estimation routines in a constant fractional marking program, which can be modified to specific situations, if the need should

arise.

One of the most important assumptions, however, is that a hatchery release group will act as a surrogate for natural fish. This assumption is critical and any difference between a natural stock and its surrogate can lead to large errors. There are only five hatcheries in the Sacramento–San Joaquin river system, compared to thirteen natural fall-run stocks of chinook salmon. A surrogate release group may behave more like the natural stock it represents if they are released in the same watershed that the natural stock is native to. This means off-site releases will most likely be necessary. Nevertheless, it is reiterated that studies should be done on natural stocks and their surrogates to make sure they behave similarly.

Nevertheless, given the model and estimation routines supplied here, probability distributions were assigned to observation errors and unobservable processes. Because the model was sequential with binary events occurring at each phase, the probability models were either binomial or multinomial, and the sampling distributions can be approximated using a multinomial distribution. Using these probability models, the variance formulas for some estimators were found. However, because some equations involved multiple estimated states, the variance formulas became complicated rather quickly, and bootstrapping methods were recommended. The downside to this is that the process becomes computer intensive and there are sometimes many layers to the sampling. However, it can be done using careful thought and programming.

A scenario that was not described here is where the fisheries in freshwater are the only selective fisheries, thus maximizing the ocean harvest. The estimation routines for this scenario can be easily derived using the same methods as in the existing alternatives. In fact, using the sequence of events given here, it would be possible to create many different scenarios of selective and non-selective fisheries.

The simulations for the no selective and all selective fisheries alternatives found that the important variables varied depending on whether or not all of the hatchery fish were marked. When all hatchery fish were distinguishable from natural fish, the escapement sampling rate was the most important variable and few interactions were significant. The design in alternative AS1 simultaneously changed the levels of the mainstem, terminal, and escapement sampling rates, but because the natural production consists of only escapement, the escapement sampling rate is most likely the largest influence. It isn't surprising that the constant fractional marking rate was not influential when all fish were marked because the natural fish are always distinguishable from the hatchery fish, regardless of the rate at which they are tagged. The CFM rate, however, may affect the precision of the estimates for hatchery specific production. Overall, an increase in the escapement sampling rate significantly increased the precision of the natural production estimates, especially when all hatchery fish were distinguishable from natural fish. A quality estimate of the escapement is required because it is a large portion of the production and the escapement estimates are used to calculate ocean and mainstem catches in the no selective

fisheries alternatives.

Interestingly, an increase in the shaker mortality rate resulted in less precise natural production estimates when simulating alternative AS1. An explanation for this is that there would be less fish in the escapement, and because *CFM Sim* uses a percentage sampled to determine the sample size, the sample size would be smaller at higher shaker mortality rates. With a smaller sample size, the variability would be higher, as can be seen in equation 48. This also explains the significant interaction between shaker mortality rate and sampling rate for all three natural stocks in alternative AS1.

Conversely, when some hatchery fish were unmarked and untagged, other variables became influential to the precision of the estimates. Particularly, the constant fractional marking rate was a positive influence on the precision, because with a higher number of fish tagged, there are less fish indistinguishable from the natural fish, and a better estimate of unmarked hatchery fish can be found (see equations 28 and 29, and equation 46). The difficulty with these alternatives is that many interactions were significant, clouding the interpretability of a single parameter.

The median *MAE*'s for each alternative were much greater when all the hatchery fish were not marked, but the range stayed nearly the same for each. The most dramatic difference between marking or not marking all of the fish occurred in the selective fisheries. For the stock in watershed 2, the *MAE* was almost 50 times higher in alternative AS2 than in alternative AS1. In addition, although alternative AS1 had the best precision in the estimates of natural production, the worst prediction errors occurred in alternative AS2.

The number of negative estimates was higher in the production estimates for Watersheds 2 and 3 when using alternative NS2 or AS2. These watersheds used the larger hatchery as surrogates, which also released many unmarked fish, resulting in more variable estimates of the unmarked *Remaining* group and more negative estimates. Higher marking and sampling rates reduced the number of negative estimates.

The no selective fisheries alternatives are an interesting case because the ocean and mainstem catches are estimated using *Surrogate* and natural terminal catch and escapement estimates (equations 23 and 24). These alternatives showed a slight increase in precision when increasing the number of *Surrogate* fish, which leads to some important conclusions about the *Surrogate* fish. First, the difference between the release sizes in this design may not have been large enough to adequately detect a difference. It appears that the number of *Surrogate* fish lowers the *MAE* (Table 11), but significant main effects were not detected. Second, the number of *Surrogate* fish released from a hatchery should be large enough such that adequate quantities appear in the catch of all fisheries and in the escapement. For example, if a very small number of *Surrogates* were released, and very few fish were alive when they entered freshwater at a certain age, it may be unlikely to sample the few fish in the freshwater catches and escapement, resulting in zero denominators for some estimators. Because the ocean and mainstem natural

catch estimates depend on the *Surrogates*, as do the life history parameters in the selective fishery alternatives, it is important that adequate numbers of *Surrogate* fish are available to be used in these estimates.

The estimates of Δ_{nt} were extremely inaccurate, especially in alternative AS2 (Table 15). The natural stock from Watershed 1 showed the highest *MAE* for Δ_{nt} , which is most likely due to the additional shaker mortality from the terminal fishery. The large variability in the *MAE* made it difficult to find significant effects in the analysis of variance. As the shaker mortality rate increased the precision of the estimates increased, but shaker mortality is not as easily controlled, thus may not be of much interest. Other factors, such as the numbers of *Stealth* and *Surrogate* fish, seemed to be slightly influential on the precision of the estimate of Δ_{nt} , and would be more useful in trying to reduce the estimation error.

It was of concern that the assumption of normal and independent, identically distributed residuals was violated in the tests using analysis of variance. Therefore, a different analysis of variance using the median from each experiment as the response variable was done for the four alternatives. The degrees of freedom were considerably reduced due to collapsing the 1000 simulations in each scenario down to one, but the results were very similar to those reported here. The models typically included the same terms as the full ANOVA models and the overall trends described here were the same. However, all of the analyses performed here ignored third-order and higher terms. Given the number of interactions present in some of the analyses, it may be worthwhile to explore these higher order interactions more, although it further complicates the interpretation of a single variable.

Even though precision for the natural production estimates was much better when marking all of the fish, a study of the costs for such a program should be done to determine if a “mass marking” program is feasible. Marking all of the fish from all hatcheries would be costly and time consuming, and increased sampling costs may occur due to the need for extra personnel to scan for CWTs in sampled fish. However, in the presence of selective fisheries, it would be beneficial to commercial and recreational fishers to mark all of the hatchery fish. The planning of a constant fractional marking program requires more thought than just determining which alternatives result in best precision. For example, incorporating costs can be thought of as a constrained optimization problem as mentioned by Newman (2000a).

The conclusions of this study may be expanded upon by performing additional analyses. Given the results shown here, and knowledge of the system for which a constant fractional marking program is to be applied, well thought out variables, that would likely be controlled in practice, would make good candidates for an analysis. Thought should also be given to determine if release levels used in the analysis would be large enough to obtain adequate samples from harvest or escapement, and if the treatment levels are far enough apart to detect significant differences. In addition, more than two levels should be used to detect non-linear trends and

add degrees of freedom to the significant interactions, which are used to test the significance of the main effects.

Also, an analysis on the estimates of hatchery production may prove worthwhile. The CFM rate did not show much influence on the *MAE* in the alternatives where all hatchery fish were marked, because natural fish could be easily separated from the hatchery fish. However, the median *MAE* for hatchery estimates was considerably less when the CFM rate was higher in alternatives NS1 and AS1. Therefore, the CFM rate may be an important variable to consider in a CFM program if accurate hatchery production estimates are desirable.

Overall, alternative AS1 showed the smallest errors in estimated production, because natural production consisted of only escapement. However, production maybe defined to include natural fish that suffered incidental mortality in the selective fisheries. In this case, the production estimates from alternative AS1 would be considerably worse because incidental mortality estimates would need to be added to the escapement estimates. These estimates would be similar to Δ_{nt} , although not the same, because Δ_{nt} represents a difference in escapement due to incidental mortality, and not the numbers that specifically died in the selective fisheries.

The simulation process and analysis done here is intended as a planning tool when implementing a constant fractional marking program to determine how different variables affect the production estimates. The estimates of error include natural variability in the life-history parameters as well as the sampling error, and are examples of the overall estimation errors likely to be encountered when using each of the alternatives. When estimating the production after the implementation of a CFM program, the bootstrapping methods described in Section 6.1.2 should be used to determine the variance of the estimates. Analytical variances may be calculated for the estimates of catches and escapements in alternative AS1 because they are all simple expansions of the samples collected, but assumptions of independence may not be met for some estimates. To calculate the variance of Δ_{it} in this alternative, the bootstrap method should be used.

The model presented here is not a standard release-recovery situation as described by Brownie et. al. (1985), and applying maximum likelihood theory requires additional thought and research. Brownie et. al. present many different models with releases in multiple years, and in the most unrestricted model, survival and recovery rates are assumed year independent. However, the model used here uses survival and harvest rates that are year and age independent, and because only young are tagged, there are identifiability problems in the parameters (see Brownie et. al. 1985, Section 3.9). In addition, inferences are made for natural stocks and non-tagged release groups using tagged release groups. Because some states are not specifically identifiable for non-tagged groups, it is unlikely that simple maximum likelihood estimates using a multinomial distribution can be formulated without some strict assumptions.

A possible solution, however, may be to use a more complicated probabilistic model to cal-

culate maximum likelihood estimates. A state-space model, conditioned on N_{a-1} , would be able to incorporate the observation and process error (Newman 1998, Schnute 1994). Also, an extra level of parameterization could be added to make a hierarchical state-space model (Newman 2000b). For example, in the model provided here, there are multiple layers of states and parameters. The bottom layer consists of the observations, which are sampled from catches, and have observation error as described in Section 4.3.2. The natural states, or catches and escapement, from which these observations are derived are dependent on life-history parameters, such as survival and maturation, as well as harvest parameters. These parameters can be characterized by probability distributions containing a third level of parameters called hyperparameters. For example, initial survival varies from year to year, and it may be that a beta distribution with hyperparameters α and β can be used to describe this natural variation. Using such a model would eliminate nonsensical, negative estimates, and other variance estimating procedures, such as inverting the Hessian, could be used for the likelihood estimates.

11 References

- Brooks EN, Pollock KH, Hoenig JM, Hearn WS. 1998. Estimation of fishing and natural mortality from tagging studies on fisheries with two user groups. *Canadian Journal of Fisheries and Aquatic Sciences* 55:2001–2010.
- Brownie C, Anderson DR, Burnham KP, and Robson DS. 1985. *Statistical inference from band recovery data – a handbook*. Second edition. U.S. Fish and Wildlife Service Resource Publication 156. 305 p.
- Committee on Protection and Management of Pacific Northwest Anadromous Salmonids, Board on Environmental Studies and Toxicology, Commission of Life Sciences. 1996. *Upstream: salmon and society in the Pacific Northwest*. National Academy Press. Washington, D.C. 452 p.
- Efron B, Tibshirani RJ. 1993. *An introduction to the bootstrap*. Chapman & Hall/CRC. Boca Raton. 436 p.
- Evans M, Hastings N, Peacock B. 1993. *Statistical distributions*. Second edition. John Wiley & Sons, Inc. New York.
- Goodman LA. 1960. On the exact variance of products. *Journal of the American Statistical Association*. 55:708–713.
- Hankin DG. 1982. Estimating escapement of Pacific salmon: marking practices to discriminate wild and hatchery fish. *Transactions of the American Fisheries Society* 111: 286–298.
- Hankin DG, Healey MC. 1986. Dependence of exploitation rates for maximum yield and stock collapse on age and sex structure of chinook salmon (*Oncorhynchus tshawytscha*) stocks. *Canadian Journal of Fisheries and Aquatic Sciences* 43:1746–1759.
- Healey MC. 1991. Life history of chinook salmon (*Oncorhynchus tshawytscha*). In *Pacific salmon life histories*. ed. Groot C and Margolis L. UBC Press, Vancouver, BC, Canada.
- Hearn WS, Pollock KH, Brooks EN. 1998. Pre- and Post-season tagging models: estimation of reporting rate and fishing and natural mortality rates. *Canadian Journal of Fisheries and Aquatic Sciences* 55:199–205.
- Hicks AC, Newman KB. 2001. *CFM Sim*. Constant Fractional Marking Simulation Program and User Guide. Version 1.0. Developed for Bailey Environmental under subcontract to Northwest Marine Technology under contract to U.S. Bureau of Reclamation.
- Kirk RE. 1995. *Experimental design: procedures for the behavioral sciences*. Brooks/Cole Publishing Company. Pacific Grove. 921 p.
- Lawson PW, Sampson DB. 1996. Gear-related mortality in selective fisheries for ocean salmon. *North American Journal of Fisheries Management* 16:512–520.
- Lebreton J, Burnham KP, Clobert J, and Anderson DR. 1992. Modeling survival and testing biological hypotheses using marked animals: a united approach with case studies. *Ecological Monographs* 62:67–118.

- McCullagh P, Nelder JA. 1989. Generalized linear models. Second edition. Chapman & Hall. London. 511 p.
- Newman KB. 1998. State-space modeling of animal movement and mortality with application to salmon. *Biometrics* 54: 274–297.
- Newman KB. 2000a. Estimating the chinook salmon production for watersheds within the Sacramento–San Joaquin river system. **In** *Development of a Marking and Tagging Program for Juvenile Chinook Salmon Reared in California’s Central Valley Fish Production Facilities*, Newman, K.B., Bailey, R. and Monroe, J. (2000) a report by Bailey Environmental under subcontract to Northwest Marine Technology under contract to U.S. Bureau of Reclamation.
- Newman KB. 2000b. Hierarchic modeling of salmon harvest and migration. *Journal of Agricultural, Biological, and Environmental Statistics* 5:98–123.
- Pacific Fishery Management Council. 2001. Review of 2000 ocean salmon fisheries. A report of the Pacific Fishery Management Council pursuant to National Oceanic and Atmospheric Administration award number NA17FC1048.
- Pollock KH, Hoenig JM, Jones CM. 1991. Estimation of fishing and natural mortality when a tagging study is combined with a creel survey or port sampling. *American Fisheries Society Symposium* 12:423–434.
- Quinn TJ II, Deriso RB. 1999. *Quantitative Fish Dynamics*. Oxford University Press. New York. 542 p.
- Rice JA. 1995. *Mathematical statistics and data analysis*. Second Edition. Duxbury Press. Belmont, CA. 602 p.
- Ross SM. 1998. *A first course in probability*. Fifth edition. Prentice-Hall, Inc. New Jersey. 514 p.
- Schnute J. 1994. A general framework for developing sequential fisheries models. *Canadian Journal of Fisheries and Aquatic Sciences* 51:1676–1688.
- Thompson SK. 1992. *Sampling*. John Wiley & Sons, Inc. New York. 343 p.
- Venables WN, Ripley BD. 1997. *Modern applied statistics with S-PLUS*. Second edition. Springer-Verlag. New York. 548 p.
- Youngs WD, Robson DS. 1975. Estimating survival rate from tag returns: model tests and sample size determination. *Journal of the Fisheries Research Board of Canada* 32:2365–2371.

Table 1: The 14 watersheds and the 5 hatcheries in the Central Valley and the total production of chinook salmon for each hatchery.

| Watershed | Hatchery | Production |
|-------------------|-----------------|--|
| American River | Nimbus | 4.6 million fall run |
| Battle Creek | Coleman NFH | 12.0 million fall run |
| Butte Creek | | |
| Clear Creek | | |
| Deer Creek | | |
| Feather River | Feather River | 8.0 million fall run 5.0 million spring run |
| Merced River | Merced | 1.0 million fall run |
| Mill Creek | | |
| Mokelumne River | Mokelumne | 5.4 million fall run |
| Sacramento River | | |
| San Joaquin River | | |
| Stanislaus River | | |
| Tuolumne River | | |
| Yuba River | | |

Table 2: Different hatchery fish adipose fin clip and tagging schemes. NS and AS denote no selective fisheries and all selective fisheries, and f is the CFM rate.

| Release Groups | NS1 | NS2 | AS1 | AS2 |
|-----------------------|------------|------------|------------|------------|
| <i>Ad hoc</i> | Optional | Optional | Optional | Optional |
| <i>Surrogate</i> | Required | Required | Required | Required |
| <i>CFM</i> | Fixed f | Fixed f | Fixed f | Fixed f |
| <i>Remainder</i> | Marked | Unmarked | Marked | Unmarked |
| <i>Stealth</i> | None | None | Required | Required |

Table 3: The notation for abundance and life history parameters.

| | |
|------------------|--|
| R_i | number of juvenile fish released from hatchery i or naturally produced juvenile fish (recruits) from watershed i |
| N_{ia} | abundance of age a fish from stock i immediately prior to ocean harvest |
| C_{Oia} | number of age a fish from stock i harvested in the ocean |
| C_{Fia} | number of age a fish from stock i harvested in the freshwater mainstem, <i>prior</i> to reaching a spawning area |
| C_{Tiaj} | number of age a fish from stock x harvested in the terminal fishery of watershed j |
| $E_{iaj,\alpha}$ | total in-river escapement of age a fish from stock i . Natural stocks include strays, while hatchery stocks do not |
| $E_{iaj,\beta}$ | total hatchery freshwater escapement of age a fish from stock i . |
| ϕ_{iI} | probability of surviving from time of release to beginning of ocean harvest, where I stands for initial |
| ϕ_{ia} | probability an unharvested, immature, age $a - 1$ fish from stock i is alive at age a prior to fishing, $a = 3, 4, 5$ |
| μ_{Oia} | ocean fishery harvest rate on age a fish |
| μ_{Fia} | freshwater mainstem (pre-watershed) fishery harvest rate on age a fish |
| μ_{Tiaj} | freshwater terminal (watershed specific) fishery harvest rate on age a fish |
| σ_{ia} | probability an age a fish from stock i matures at age a , given survival to age a without maturing earlier |
| ρ_{iaj} | probability a fish from stock i goes to watershed j conditional on surviving the freshwater mainstem fishery |
| η_{iaj} | probability that a fish in watershed j will escape to the hatchery The complement of this, $1 - \eta$, is the probability that a fish will spawn naturally in the watershed and contribute to next year's natural recruitment |

Table 4: Notation for samples and sample recoveries.

| | |
|------------------|--|
| n_O | size of a simple random sample taken from C_O |
| n_F | size of a simple random sample taken from C_F |
| n_{Tj} | size of a simple random sample taken from C_{Tj} |
| $n_{Ej,\alpha}$ | size of a simple random sample taken from $E_{j,\alpha}$, the watershed j in-river escapement |
| $n_{Ej,\beta}$ | size of a simple random sample taken from $E_{j,\beta}$, the watershed j hatchery escapement |
| xa_i | # of hatchery stock i <i>Ad hoc</i> recoveries in ocean catch sample |
| xb_i | # of hatchery stock i <i>Surrogate</i> recoveries in ocean catch sample |
| xc_i | # of hatchery stock i <i>CFM</i> recoveries in ocean catch sample |
| xd_i | # of hatchery stock i <i>Remainder</i> recoveries in ocean catch sample |
| xe_i | # of hatchery stock i <i>Stealth</i> recoveries in ocean catch sample |
| ya_i | # of hatchery stock i <i>Ad hoc</i> recoveries in freshwater mainstem catch sample |
| yb_i | # of hatchery stock i <i>Surrogate</i> recoveries in freshwater mainstem catch sample |
| yc_i | # of hatchery stock i <i>CFM</i> recoveries in freshwater mainstem catch sample |
| yd_i | # of hatchery stock i <i>Remainder</i> recoveries in freshwater mainstem catch sample |
| ye_i | # of hatchery stock i <i>Stealth</i> recoveries in freshwater mainstem catch sample |
| ta_{ij} | # of hatchery stock i <i>Ad hoc</i> recoveries in terminal area j catch sample |
| tb_{ij} | # of hatchery stock i <i>Surrogate</i> recoveries in terminal area j catch sample |
| tc_{ij} | # of hatchery stock i <i>CFM</i> recoveries in terminal area j catch sample |
| td_{ij} | # of hatchery stock i <i>Remainder</i> recoveries in terminal area j catch sample |
| te_{ij} | # of hatchery stock i <i>Stealth</i> recoveries in terminal area j catch sample |
| tn_j | # of natural stock j recoveries in terminal area j catch sample |
| $za_{ij,\alpha}$ | # of hatchery stock i <i>Ad hoc</i> recoveries in watershed j in-river escapement sample |
| $zb_{ij,\alpha}$ | # of hatchery stock i <i>Surrogate</i> recoveries in watershed j in-river escapement sample |
| $zc_{ij,\alpha}$ | # of hatchery stock i <i>CFM</i> recoveries in watershed j in-river escapement sample |
| $zd_{ij,\alpha}$ | # of hatchery stock i <i>Remainder</i> recoveries in watershed j in-river escapement sample |
| $ze_{ij,\alpha}$ | # of hatchery stock i <i>Stealth</i> recoveries in watershed j in-river escapement sample |
| $zn_{j,\alpha}$ | # of natural stock j recoveries in watershed j in-river escapement sample |
| $zu_{j,\alpha}$ | # of unmarked fish in watershed j in-river escapement sample |
| $za_{ij,\beta}$ | # of hatchery stock i <i>Ad hoc</i> recoveries in watershed j hatchery escapement sample |
| $zb_{ij,\beta}$ | # of hatchery stock i <i>Surrogate</i> recoveries in watershed j hatchery escapement sample |
| $zc_{ij,\beta}$ | # of hatchery stock i <i>CFM</i> recoveries in watershed j hatchery escapement sample |
| $zd_{ij,\beta}$ | # of hatchery stock i <i>Remainder</i> recoveries in watershed j hatchery escapement sample |
| $ze_{ij,\beta}$ | # of hatchery stock i <i>Stealth</i> recoveries in watershed j hatchery escapement sample |
| $zn_{j,\beta}$ | # of natural stock j recoveries in watershed j hatchery escapement sample |
| $zu_{j,\beta}$ | # of unmarked fish in watershed j hatchery escapement sample |

Table 5: The fixed minimum, mode, and maximum of the survival, ocean harvest, and freshwater harvest parameters (in percent) for all stocks.

| | Survival | | | Ocean Harvest | | | Mainstem Harvest | | |
|-------|----------|------|-----|---------------|------|-----|------------------|------|-----|
| | Min | Mode | Max | Min | Mode | Max | Min | Mode | Max |
| Age 2 | 1 | 3.5 | 8 | 0 | 0 | 0 | 10 | 20 | 30 |
| Age 3 | 45 | 50 | 70 | 55 | 60 | 65 | 10 | 20 | 30 |
| Age 4 | 75 | 80 | 90 | 40 | 50 | 60 | 10 | 20 | 30 |
| Age 5 | 75 | 80 | 90 | 40 | 50 | 60 | 10 | 20 | 30 |

Table 6: Maturity parameters, in percent, for the two hatchery stocks. The natural stocks have the same parameters as its surrogate.

| Stock | Age 2 | | | Age 3 | | | Age 4 | | | Age 5 |
|------------|-------|------|-----|-------|------|-----|-------|------|-----|-------|
| | Min | Mode | Max | Min | Mode | Max | Min | Mode | Max | |
| Hatchery 1 | 1 | 4 | 7 | 20 | 30 | 40 | 70 | 80 | 85 | 100 |
| Hatchery 2 | 0.3 | 0.5 | 1 | 5 | 10 | 20 | 40 | 50 | 70 | 100 |

Table 7: Straying rates for the two hatchery stocks. The natural stocks do not stray from their native watershed.

| Stock | Watershed 1 | Watershed 2 | Watershed 3 |
|------------|-------------|-------------|-------------|
| Hatchery 1 | 20 | 60 | 20 |
| Hatchery 2 | 15 | 80 | 5 |

Table 8: The experimental design for the simulation studies of alternatives NS1 and NS2. *CFM* is the constant fractional marking rate, *ESR* is the escapement sampling rate, *FSR* is the freshwater sampling rate, *Sur* is the number of surrogates released, and *TSR* is the terminal area sampling rate. The letters “l” and “h” denote low and high levels for the factors *CFM*, *ESR*, *FSR*, *Sur*, and *TSR*, respectively.

| Expt | Levels | <i>CFM</i> | <i>ESR</i> | <i>FSR</i> | <i>Sur</i> | <i>TSR</i> |
|------|--------|------------|------------|------------|------------|------------|
| 1 | lhhhh | 20 | 30 | 30 | 100000 | 30 |
| 2 | hhhhh | 50 | 30 | 30 | 100000 | 30 |
| 3 | llhhh | 20 | 10 | 30 | 100000 | 30 |
| 4 | hlhhh | 50 | 10 | 30 | 100000 | 30 |
| 5 | lhlhh | 20 | 30 | 10 | 100000 | 30 |
| 6 | hhlhh | 50 | 30 | 10 | 100000 | 30 |
| 7 | lllhh | 20 | 10 | 10 | 100000 | 30 |
| 8 | hllhh | 50 | 10 | 10 | 100000 | 30 |
| 9 | lhhlh | 20 | 30 | 30 | 50000 | 30 |
| 10 | hbhlh | 50 | 30 | 30 | 50000 | 30 |
| 11 | llhlh | 20 | 10 | 30 | 50000 | 30 |
| 12 | hlhlh | 50 | 10 | 30 | 50000 | 30 |
| 13 | lhllh | 20 | 30 | 10 | 50000 | 30 |
| 14 | hbllh | 50 | 30 | 10 | 50000 | 30 |
| 15 | llllh | 20 | 10 | 10 | 50000 | 30 |
| 16 | hlllh | 50 | 10 | 10 | 50000 | 30 |
| 17 | lhhhl | 20 | 30 | 30 | 100000 | 10 |
| 18 | hbhhhl | 50 | 30 | 30 | 100000 | 10 |
| 19 | llhhl | 20 | 10 | 30 | 100000 | 10 |
| 20 | hlhhl | 50 | 10 | 30 | 100000 | 10 |
| 21 | lhhlh | 20 | 30 | 10 | 100000 | 10 |
| 22 | hbhlh | 50 | 30 | 10 | 100000 | 10 |
| 23 | lllhl | 20 | 10 | 10 | 100000 | 10 |
| 24 | hllhl | 50 | 10 | 10 | 100000 | 10 |
| 25 | lhhll | 20 | 30 | 30 | 50000 | 10 |
| 26 | hbhll | 50 | 30 | 30 | 50000 | 10 |
| 27 | llhll | 20 | 10 | 30 | 50000 | 10 |
| 28 | hlhll | 50 | 10 | 30 | 50000 | 10 |
| 29 | lhlll | 20 | 30 | 10 | 50000 | 10 |
| 30 | hblll | 50 | 30 | 10 | 50000 | 10 |
| 31 | lllll | 20 | 10 | 10 | 50000 | 10 |
| 32 | hllll | 50 | 10 | 10 | 50000 | 10 |

Table 9: The number of simulations where the estimated production was less than the true production out of 1000 simulations for each natural stock and alternatives NS1 and NS2. “WS” refers to the watershed that the natural stock is natal to. The levels of the variables are given for reference. Refer to Table 8 for an explanation of the levels.

| Expt | Levels | NS1 | | | NS2 | | |
|------|--------|------|------|------|------|------|------|
| | | WS 1 | WS 2 | WS 3 | WS 1 | WS 2 | WS 3 |
| 1 | lhhhh | 471 | 461 | 457 | 472 | 472 | 483 |
| 2 | hhhhh | 486 | 487 | 485 | 473 | 511 | 495 |
| 3 | llhhh | 475 | 479 | 473 | 455 | 485 | 497 |
| 4 | hlhhh | 474 | 484 | 519 | 492 | 465 | 472 |
| 5 | lhlhh | 465 | 486 | 515 | 482 | 482 | 519 |
| 6 | hhlhh | 469 | 497 | 502 | 496 | 523 | 486 |
| 7 | lllhh | 489 | 480 | 495 | 476 | 491 | 513 |
| 8 | hllhh | 481 | 462 | 485 | 476 | 495 | 499 |
| 9 | lhhlh | 459 | 501 | 485 | 440 | 479 | 493 |
| 10 | hhlhl | 467 | 483 | 475 | 479 | 484 | 474 |
| 11 | llhlh | 456 | 459 | 479 | 467 | 487 | 459 |
| 12 | hlhlh | 465 | 477 | 477 | 470 | 515 | 468 |
| 13 | lhllh | 448 | 484 | 488 | 463 | 472 | 493 |
| 14 | hhllh | 489 | 474 | 486 | 461 | 499 | 475 |
| 15 | llllh | 454 | 445 | 456 | 447 | 500 | 470 |
| 16 | hlllh | 489 | 461 | 438 | 443 | 485 | 468 |
| 17 | lhhlh | 485 | 451 | 472 | 428 | 503 | 480 |
| 18 | hhhlh | 488 | 493 | 468 | 476 | 491 | 508 |
| 19 | llhhl | 484 | 482 | 515 | 452 | 464 | 468 |
| 20 | hlhhl | 469 | 467 | 461 | 484 | 487 | 492 |
| 21 | lhhlh | 491 | 504 | 510 | 441 | 488 | 475 |
| 22 | hhlhl | 488 | 497 | 501 | 435 | 489 | 489 |
| 23 | lllhl | 484 | 500 | 460 | 455 | 487 | 491 |
| 24 | hllhl | 489 | 501 | 450 | 516 | 531 | 502 |
| 25 | lhhlh | 472 | 437 | 460 | 411 | 477 | 499 |
| 26 | hhhlh | 472 | 465 | 481 | 456 | 516 | 494 |
| 27 | llhhl | 472 | 470 | 464 | 433 | 473 | 491 |
| 28 | hlhhl | 451 | 458 | 429 | 463 | 507 | 484 |
| 29 | lhllh | 461 | 456 | 470 | 404 | 473 | 476 |
| 30 | hhllh | 467 | 467 | 473 | 475 | 480 | 469 |
| 31 | llllh | 465 | 432 | 434 | 422 | 515 | 487 |
| 32 | hlllh | 487 | 442 | 460 | 471 | 476 | 482 |

Table 10: Number of negative estimates for each scenario out of 1000 simulations using alternative NS2. The levels of the variables are given for reference. Refer to Table 8 for an explanation of the levels.

| Expt | Levels | Watershed 1 | | Watershed 2 | Watershed 3 |
|------|--------|-------------|----------------|-------------|-------------|
| | | Escapement | Terminal Catch | Escapement | Escapement |
| 1 | lhhhh | 0 | 35 | 25 | 3 |
| 2 | hhhhh | 0 | 0 | 0 | 0 |
| 3 | llhhh | 0 | 25 | 48 | 43 |
| 4 | hlhhh | 0 | 1 | 2 | 0 |
| 5 | hlhhl | 0 | 29 | 32 | 2 |
| 6 | hhlhh | 0 | 0 | 0 | 0 |
| 7 | lllhh | 0 | 22 | 51 | 42 |
| 8 | hllhh | 0 | 1 | 5 | 2 |
| 9 | lhhhl | 0 | 32 | 31 | 9 |
| 10 | hhhlh | 0 | 2 | 0 | 0 |
| 11 | llhlh | 0 | 46 | 46 | 49 |
| 12 | hlhlh | 0 | 0 | 3 | 0 |
| 13 | lhllh | 0 | 39 | 21 | 3 |
| 14 | hhlhh | 0 | 0 | 1 | 0 |
| 15 | llllh | 0 | 35 | 61 | 35 |
| 16 | hlllh | 0 | 0 | 4 | 1 |
| 17 | lhhlh | 0 | 109 | 27 | 4 |
| 18 | hhhlh | 0 | 27 | 0 | 0 |
| 19 | llhhl | 0 | 140 | 56 | 36 |
| 20 | hlhhl | 0 | 25 | 4 | 3 |
| 21 | hlhlh | 0 | 123 | 25 | 3 |
| 22 | hhlhl | 0 | 14 | 0 | 0 |
| 23 | lllhl | 0 | 110 | 57 | 46 |
| 24 | hllhl | 0 | 28 | 0 | 0 |
| 25 | lhhlh | 0 | 121 | 31 | 3 |
| 26 | hhhlh | 0 | 25 | 1 | 0 |
| 27 | llhll | 0 | 133 | 74 | 42 |
| 28 | hlhll | 0 | 22 | 1 | 2 |
| 29 | lhlll | 0 | 116 | 43 | 1 |
| 30 | hhlll | 0 | 20 | 1 | 0 |
| 31 | lllll | 1 | 142 | 74 | 36 |
| 32 | hllll | 0 | 27 | 1 | 1 |

Table 11: Means and medians of the simulated natural production *MAE*'s (in %) for the high (H) and low (L) levels of the variables in the factorial design for the three natural stocks and alternatives NS1 and NS2. Section 9.2.1 defines the variables and their acronyms.

| | Watershed 1 | | | | | | | |
|------------|-------------|------|--------|-----|------|------|--------|------|
| | NS1 | | | | NS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 12.4 | 12.2 | 9.1 | 9.1 | 21.3 | 15.4 | 16.8 | 11.8 |
| <i>ESR</i> | 15.9 | 8.6 | 12.6 | 6.9 | 21.5 | 15.2 | 16.9 | 11.8 |
| <i>FSR</i> | 12.4 | 12.2 | 9.2 | 9.1 | 18.4 | 18.2 | 14.1 | 14.0 |
| <i>Sur</i> | 13.5 | 11.1 | 10.2 | 8.2 | 19.3 | 17.4 | 14.8 | 13.4 |
| <i>TSR</i> | 12.6 | 12.0 | 9.4 | 8.9 | 19.5 | 17.2 | 15.0 | 13.3 |

| | Watershed 2 | | | | | | | |
|------------|-------------|------|--------|------|------|------|--------|------|
| | NS1 | | | | NS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 14.6 | 14.7 | 11.1 | 11.3 | 61.2 | 36.3 | 50.9 | 28.9 |
| <i>ESR</i> | 17.0 | 12.4 | 13.3 | 9.5 | 53.6 | 43.8 | 42.4 | 34.6 |
| <i>FSR</i> | 14.7 | 14.6 | 11.2 | 11.2 | 48.3 | 49.2 | 37.7 | 38.5 |
| <i>Sur</i> | 16.5 | 12.8 | 12.5 | 10.0 | 50.3 | 47.1 | 38.4 | 37.7 |
| <i>TSR</i> | 14.7 | 14.6 | 11.2 | 11.2 | 49.0 | 48.5 | 38.2 | 37.9 |

| | Watershed 3 | | | | | | | |
|------------|-------------|------|--------|------|------|------|--------|------|
| | NS1 | | | | NS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 15.3 | 15.4 | 11.7 | 11.8 | 52.2 | 31.6 | 42.8 | 23.9 |
| <i>ESR</i> | 18.0 | 12.7 | 14.0 | 9.9 | 51.2 | 32.6 | 40.8 | 24.8 |
| <i>FSR</i> | 15.3 | 15.4 | 11.7 | 11.8 | 41.5 | 42.3 | 31.4 | 31.6 |
| <i>Sur</i> | 17.1 | 13.6 | 13.0 | 10.6 | 43.3 | 40.6 | 32.0 | 31.0 |
| <i>TSR</i> | 15.5 | 15.1 | 11.9 | 11.6 | 42.0 | 41.8 | 31.8 | 31.3 |

Table 12: All possible treatment combinations in the 2^5 factorial design used to analyze alternatives AS1 and AS2. *CFM* is the constant fractional marking rate, *SR* is the escapement and freshwater fisheries sampling rates, *SM* is the shaker mortality rate, *Sur* is the number of surrogates released, and *Stl* is the number of stealth fish released. The letters “l” and “h” denote low and high levels for the factors *CFM*, *SR*, *SM*, *Sur*, and *Stl*, respectively.

| Expt | Levels | <i>CFM</i> | <i>SR</i> | <i>SM</i> | <i>Sur</i> | <i>Stl</i> |
|------|---------|------------|-----------|-----------|------------|------------|
| 1 | lhhhh | 20 | 30 | 30 | 100000 | 100000 |
| 2 | hhhhh | 50 | 30 | 30 | 100000 | 100000 |
| 3 | llhhh | 20 | 10 | 30 | 100000 | 100000 |
| 4 | hlhhh | 50 | 10 | 30 | 100000 | 100000 |
| 5 | lhlhh | 20 | 30 | 10 | 100000 | 100000 |
| 6 | hhlhh | 50 | 30 | 10 | 100000 | 100000 |
| 7 | lllhh | 20 | 10 | 10 | 100000 | 100000 |
| 8 | hllhh | 50 | 10 | 10 | 100000 | 100000 |
| 9 | lhhlh | 20 | 30 | 30 | 50000 | 100000 |
| 10 | hhhlh | 50 | 30 | 30 | 50000 | 100000 |
| 11 | llhlh | 20 | 10 | 30 | 50000 | 100000 |
| 12 | hlhlh | 50 | 10 | 30 | 50000 | 100000 |
| 13 | lhllh | 20 | 30 | 10 | 50000 | 100000 |
| 14 | hlllh | 50 | 30 | 10 | 50000 | 100000 |
| 15 | llllh | 20 | 10 | 10 | 50000 | 100000 |
| 16 | hlllh | 50 | 10 | 10 | 50000 | 100000 |
| 17 | lhhdl | 20 | 30 | 30 | 100000 | 50000 |
| 18 | hhhdl | 50 | 30 | 30 | 100000 | 50000 |
| 19 | llhdl | 20 | 10 | 30 | 100000 | 50000 |
| 20 | hldhl | 50 | 10 | 30 | 100000 | 50000 |
| 21 | lhdhl | 20 | 30 | 10 | 100000 | 50000 |
| 22 | hdhdh | 50 | 30 | 10 | 100000 | 50000 |
| 23 | llldh | 20 | 10 | 10 | 100000 | 50000 |
| 24 | hdldh | 50 | 10 | 10 | 100000 | 50000 |
| 25 | lhdll | 20 | 30 | 30 | 50000 | 50000 |
| 26 | hdhdll | 50 | 30 | 30 | 50000 | 50000 |
| 27 | lldll | 20 | 10 | 30 | 50000 | 50000 |
| 28 | hdldll | 50 | 10 | 30 | 50000 | 50000 |
| 29 | lhdlll | 20 | 30 | 10 | 50000 | 50000 |
| 30 | hdldlll | 50 | 30 | 10 | 50000 | 50000 |
| 31 | lldlll | 20 | 10 | 10 | 50000 | 50000 |
| 32 | hdldlll | 50 | 10 | 10 | 50000 | 50000 |

Table 13: The number of simulations where the estimated production was less than the true production out of 1000 simulations for alternatives AS1 and AS2. The levels of the variables are given for reference. Refer to Table 12 for an explanation of the levels.

| Expt | Levels | AS1 | | | AS2 | | |
|------|--------|------|------|------|------|------|------|
| | | WS 1 | WS 2 | WS 3 | WS 1 | WS 2 | WS 3 |
| 1 | lhhhh | 475 | 498 | 501 | 512 | 532 | 469 |
| 2 | hhhhh | 473 | 482 | 474 | 490 | 488 | 486 |
| 3 | llhhh | 513 | 502 | 489 | 476 | 489 | 467 |
| 4 | hlhhh | 502 | 508 | 484 | 507 | 505 | 476 |
| 5 | lhlhh | 479 | 475 | 482 | 509 | 540 | 470 |
| 6 | hhlhh | 495 | 518 | 530 | 497 | 503 | 460 |
| 7 | lllhh | 529 | 501 | 500 | 469 | 480 | 511 |
| 8 | hllhh | 508 | 499 | 507 | 460 | 481 | 478 |
| 9 | lhhlh | 517 | 492 | 475 | 510 | 493 | 471 |
| 10 | hhhlh | 502 | 512 | 508 | 497 | 517 | 516 |
| 11 | llhlh | 481 | 506 | 480 | 465 | 491 | 483 |
| 12 | hlhlh | 493 | 497 | 491 | 449 | 508 | 448 |
| 13 | lhllh | 512 | 494 | 481 | 478 | 501 | 527 |
| 14 | hhllh | 502 | 511 | 484 | 511 | 479 | 492 |
| 15 | llllh | 516 | 505 | 494 | 509 | 481 | 479 |
| 16 | hlllh | 501 | 503 | 487 | 494 | 514 | 494 |
| 17 | lhhdl | 510 | 515 | 493 | 501 | 484 | 483 |
| 18 | hhhdl | 468 | 495 | 497 | 475 | 509 | 489 |
| 19 | llhdl | 510 | 510 | 498 | 499 | 503 | 473 |
| 20 | hlhdl | 508 | 500 | 497 | 477 | 485 | 461 |
| 21 | lhldl | 490 | 498 | 541 | 467 | 506 | 477 |
| 22 | hldhl | 489 | 503 | 474 | 498 | 512 | 473 |
| 23 | llldl | 519 | 499 | 493 | 482 | 507 | 492 |
| 24 | hllldl | 499 | 513 | 512 | 464 | 489 | 476 |
| 25 | lhldl | 491 | 500 | 477 | 511 | 518 | 476 |
| 26 | hhldl | 489 | 495 | 487 | 505 | 525 | 522 |
| 27 | llldl | 480 | 506 | 511 | 470 | 480 | 464 |
| 28 | hldldl | 508 | 522 | 497 | 482 | 500 | 478 |
| 29 | lhldl | 480 | 492 | 526 | 491 | 492 | 490 |
| 30 | hhldl | 505 | 470 | 517 | 494 | 505 | 469 |
| 31 | llldl | 523 | 503 | 482 | 486 | 471 | 466 |
| 32 | hldldl | 536 | 485 | 488 | 465 | 466 | 475 |

Table 14: Means and medians of the simulated natural production *MAE*'s (in %) for the high (H) and low (L) levels of the variables in the factorial design for the three natural stocks and alternatives AS1 and AS2. Section 9.3.1 defines the variables and their acronyms.

| | Watershed 1 | | | | | | | |
|------------|-------------|------|--------|------|------|------|--------|------|
| | AS1 | | | | AS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 2.1 | 2.1 | 0.96 | 0.95 | 33.8 | 22.1 | 27.5 | 17.6 |
| <i>SR</i> | 3.5 | 0.68 | 2.1 | 0.56 | 30.1 | 25.7 | 23.9 | 19.9 |
| <i>SM</i> | 1.7 | 2.5 | 0.83 | 1.1 | 26.5 | 29.3 | 20.6 | 23.1 |
| <i>Sur</i> | 2.1 | 2.1 | 0.96 | 0.96 | 27.8 | 28.1 | 21.5 | 22.0 |
| <i>Stl</i> | 2.1 | 2.1 | 0.95 | 0.96 | 32.8 | 23.0 | 26.3 | 18.3 |

| | Watershed 2 | | | | | | | |
|------------|-------------|-----|--------|-----|------|------|--------|------|
| | AS1 | | | | AS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 2.0 | 2.0 | 1.5 | 1.5 | 79.1 | 56.6 | 74.1 | 48.9 |
| <i>SR</i> | 2.8 | 1.3 | 2.2 | 1.1 | 72.9 | 62.8 | 65.2 | 54.4 |
| <i>SM</i> | 1.8 | 2.2 | 1.3 | 1.7 | 62.9 | 72.7 | 54.2 | 65.4 |
| <i>Sur</i> | 2.0 | 2.0 | 1.5 | 1.5 | 67.0 | 68.7 | 34.1 | 34.8 |
| <i>Stl</i> | 2.0 | 2.0 | 1.5 | 1.5 | 78.1 | 57.6 | 72.6 | 49.6 |

| | Watershed 3 | | | | | | | |
|------------|-------------|-----|--------|-----|------|------|--------|------|
| | AS1 | | | | AS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 2.3 | 2.3 | 1.7 | 1.7 | 52.1 | 35.6 | 43.5 | 27.8 |
| <i>SR</i> | 3.1 | 1.5 | 2.5 | 1.3 | 53.5 | 34.1 | 45.2 | 26.6 |
| <i>SM</i> | 2.1 | 2.6 | 1.6 | 1.9 | 40.3 | 47.4 | 31.1 | 38.1 |
| <i>Sur</i> | 2.3 | 2.3 | 1.7 | 1.7 | 43.6 | 44.1 | 34.1 | 34.8 |
| <i>Stl</i> | 2.3 | 2.3 | 1.7 | 1.7 | 50.7 | 36.9 | 42.0 | 28.3 |

Table 15: Means and medians of the simulated Δ_{nt} *MAE*'s (in %) for the high (H) and low (L) levels of the variables in the factorial design for the three natural stocks and alternatives AS1 and AS2. Section 9.3.1 defines the variables and their acronyms.

| | Watershed 1 | | | | | | | |
|------------|-------------|------|--------|------|------|-----|--------|-----|
| | AS1 | | | | AS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 46.5 | 47.5 | 33.5 | 34.0 | 463 | 454 | 283 | 292 |
| <i>SR</i> | 52.4 | 41.5 | 38.4 | 29.7 | 460 | 458 | 282 | 291 |
| <i>SM</i> | 65.3 | 28.7 | 53.2 | 22.7 | 768 | 149 | 749 | 142 |
| <i>Sur</i> | 51.6 | 42.3 | 37.4 | 30.4 | 457 | 461 | 280 | 292 |
| <i>Stl</i> | 49.4 | 44.5 | 35.2 | 32.4 | 463 | 455 | 290 | 281 |

| | Watershed 2 | | | | | | | |
|------------|-------------|------|--------|------|------|-----|--------|-----|
| | AS1 | | | | AS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 41.4 | 41.1 | 30.5 | 30.2 | 422 | 368 | 192 | 202 |
| <i>SR</i> | 44.9 | 37.6 | 33.5 | 27.6 | 409 | 380 | 198 | 197 |
| <i>SM</i> | 57.3 | 25.2 | 49.4 | 20.3 | 643 | 146 | 566 | 100 |
| <i>Sur</i> | 45.4 | 37.1 | 34.4 | 26.9 | 396 | 394 | 191 | 203 |
| <i>Stl</i> | 42.6 | 39.9 | 31.2 | 29.6 | 418 | 372 | 202 | 192 |

| | Watershed 3 | | | | | | | |
|------------|-------------|------|--------|------|------|-----|--------|------|
| | AS1 | | | | AS2 | | | |
| | Mean | | Median | | Mean | | Median | |
| | L | H | L | H | L | H | L | H |
| <i>CFM</i> | 41.6 | 41.2 | 30.8 | 30.2 | 364 | 344 | 199 | 205 |
| <i>SR</i> | 45.1 | 37.7 | 33.6 | 27.6 | 369 | 339 | 202 | 202 |
| <i>SM</i> | 57.6 | 25.2 | 49.6 | 20.3 | 595 | 113 | 571 | 99.8 |
| <i>Sur</i> | 45.5 | 37.3 | 34.3 | 27.1 | 351 | 356 | 195 | 208 |
| <i>Stl</i> | 42.8 | 40.0 | 31.3 | 29.5 | 366 | 342 | 206 | 197 |

Table 16: Number of negative estimates for each scenario out of 1000 simulations using alternative AS2. The levels of the variables are given for reference. Refer to Table 12 for an explanation of the levels.

| | | Watershed 1 | Watershed 2 | Watershed 3 |
|------|--------|-------------|-------------|-------------|
| Expt | Levels | Escapement | Escapement | Escapement |
| 1 | lhhhh | 4 | 166 | 23 |
| 2 | hhhhh | 0 | 59 | 2 |
| 3 | llhhh | 13 | 197 | 116 |
| 4 | hlhhh | 1 | 85 | 45 |
| 5 | lhllh | 3 | 95 | 12 |
| 6 | hllhh | 0 | 25 | 0 |
| 7 | llllh | 7 | 145 | 89 |
| 8 | hlllh | 0 | 59 | 16 |
| 9 | lhllh | 3 | 139 | 24 |
| 10 | hhllh | 0 | 56 | 0 |
| 11 | llhll | 14 | 196 | 115 |
| 12 | hlhll | 2 | 87 | 34 |
| 13 | lhllh | 1 | 115 | 11 |
| 14 | hhllh | 0 | 22 | 1 |
| 15 | llllh | 10 | 145 | 69 |
| 16 | hlllh | 0 | 53 | 17 |
| 17 | lhhhl | 14 | 198 | 77 |
| 18 | hhhhl | 16 | 223 | 61 |
| 19 | llhhl | 44 | 278 | 193 |
| 20 | hlhhl | 6 | 156 | 82 |
| 21 | lhhlh | 26 | 189 | 29 |
| 22 | hhhlh | 1 | 90 | 2 |
| 23 | lllhl | 34 | 263 | 147 |
| 24 | hllhl | 10 | 124 | 51 |
| 25 | lhlll | 19 | 222 | 88 |
| 26 | hhlll | 1 | 119 | 11 |
| 27 | llhll | 54 | 266 | 200 |
| 28 | hlhll | 12 | 174 | 83 |
| 29 | lhlll | 10 | 177 | 38 |
| 30 | hhlll | 1 | 93 | 4 |
| 31 | lllll | 39 | 216 | 135 |
| 32 | hllll | 1 | 111 | 54 |



Figure 1: Approximate locations of the five hatcheries on the Sacramento–San Joaquin River system in California.

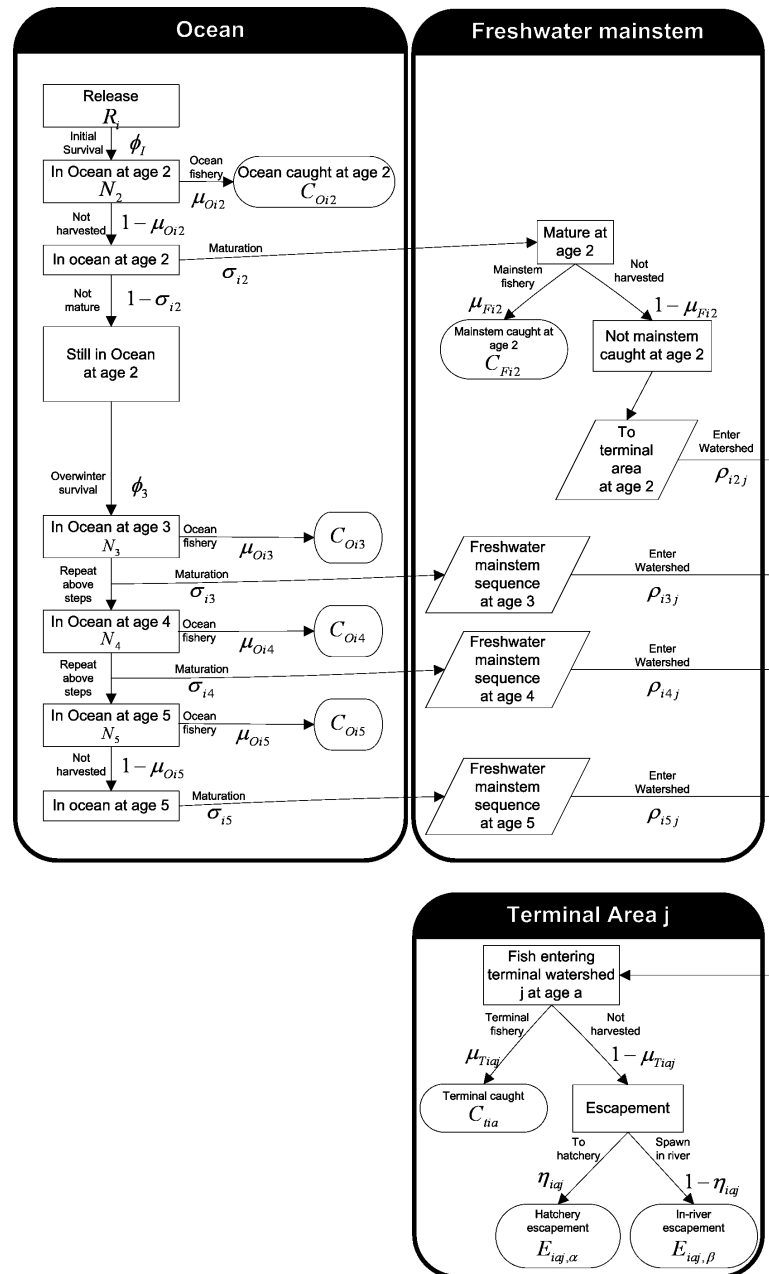


Figure 2: The sequence of fates that a fish can have. See Table 3 for an explanation of the variables. Probabilities at each step are indicated by Greek letters.

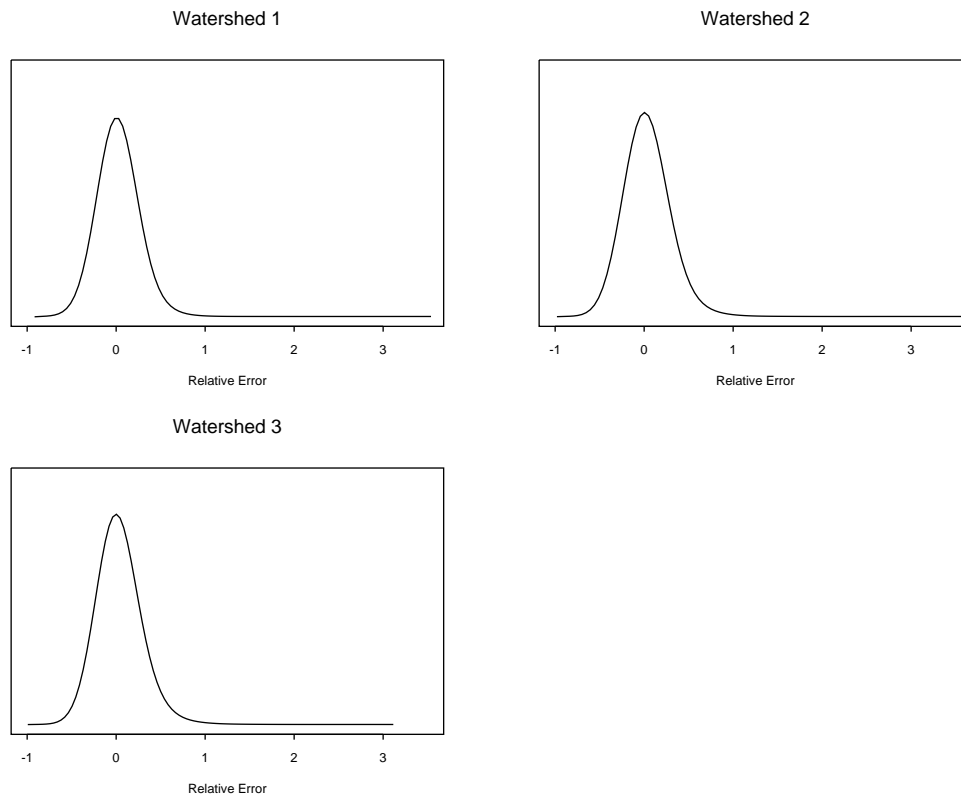


Figure 3: Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative NS1.

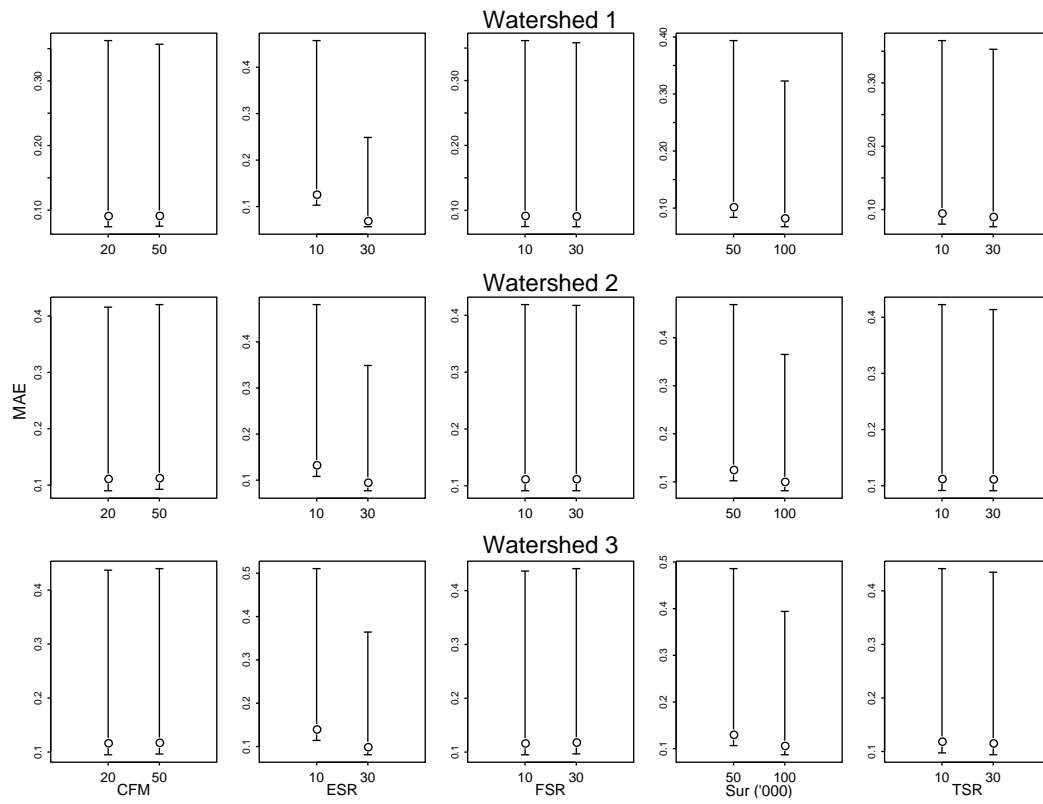


Figure 4: Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the *MAE* of estimated production for each treatment level and natural stock simulated under alternative NS1.

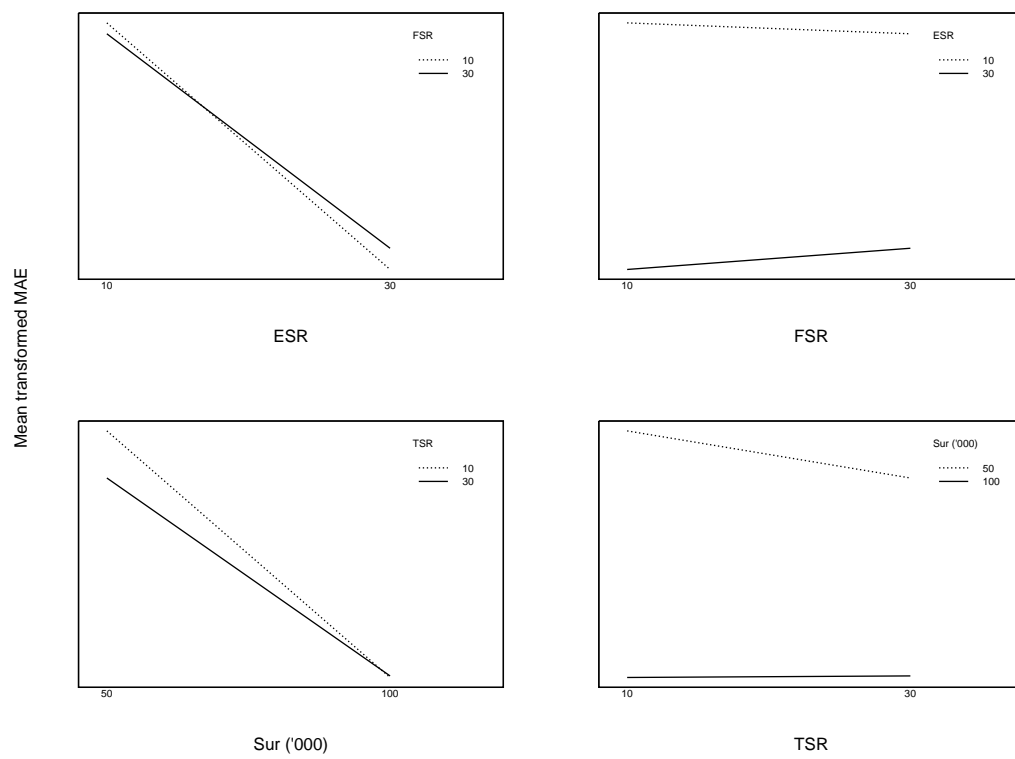


Figure 5: Significant $ESR-FSR$ and $TSR-Sur$ interactions on the transformed MAE for Watershed 3 simulated under alternative NS1.

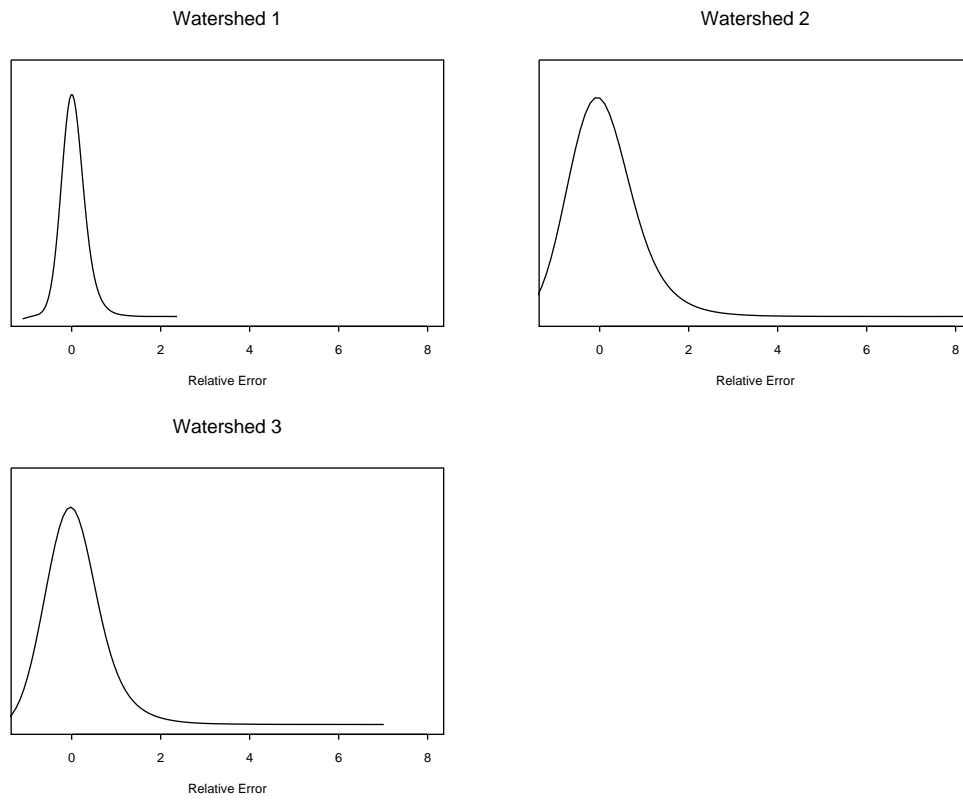


Figure 6: Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative NS2.

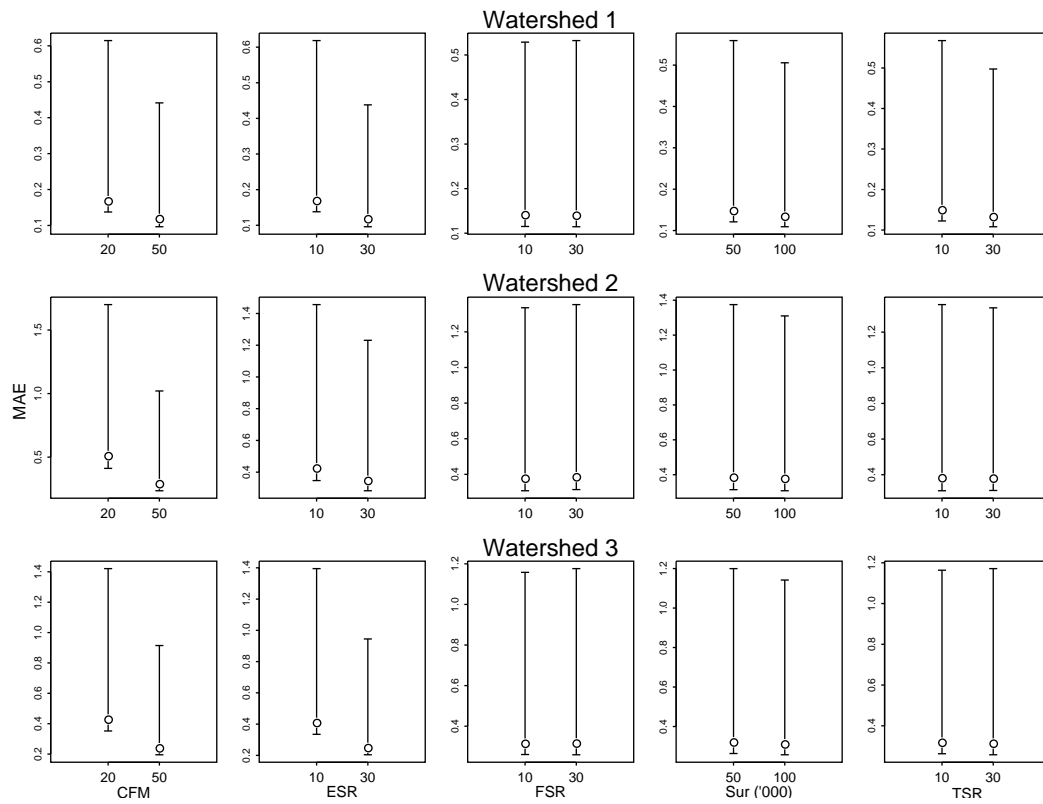


Figure 7: Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the *MAE* of estimated production for each treatment level and natural stock simulated under alternative NS2.

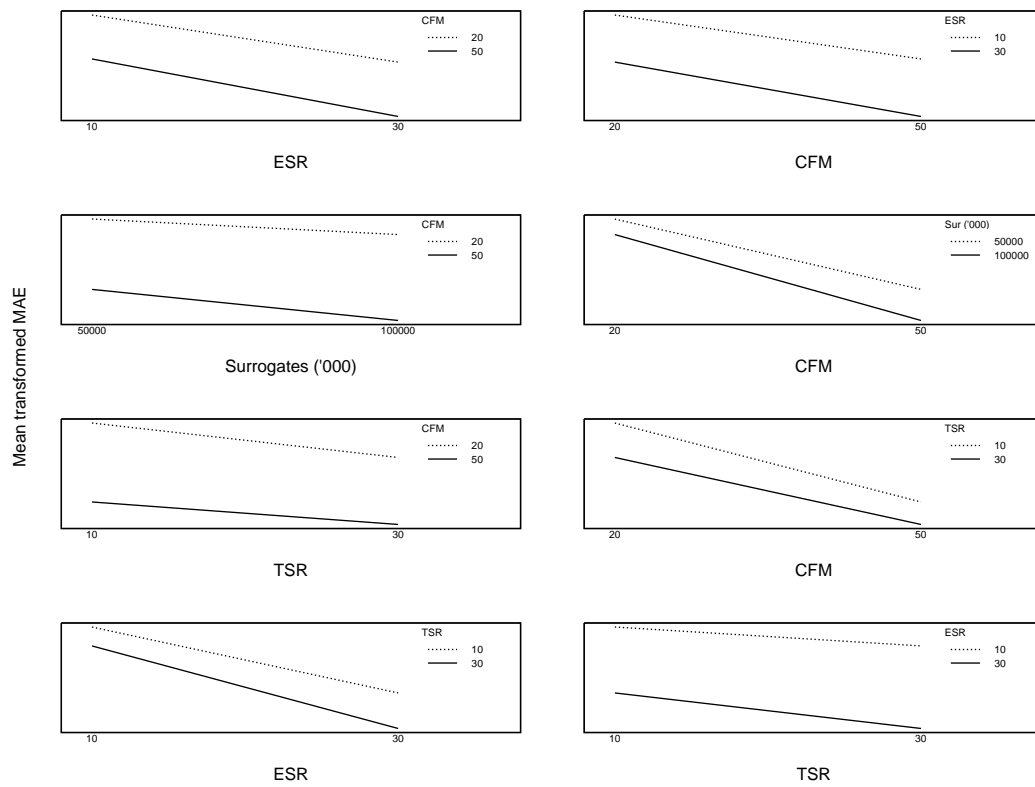


Figure 8: Significant *CFM-ESR*, *CFM-TSR*, and *ESR-TSR* interactions on the transformed *MAE* for Watershed 1 simulated under alternative NS2.

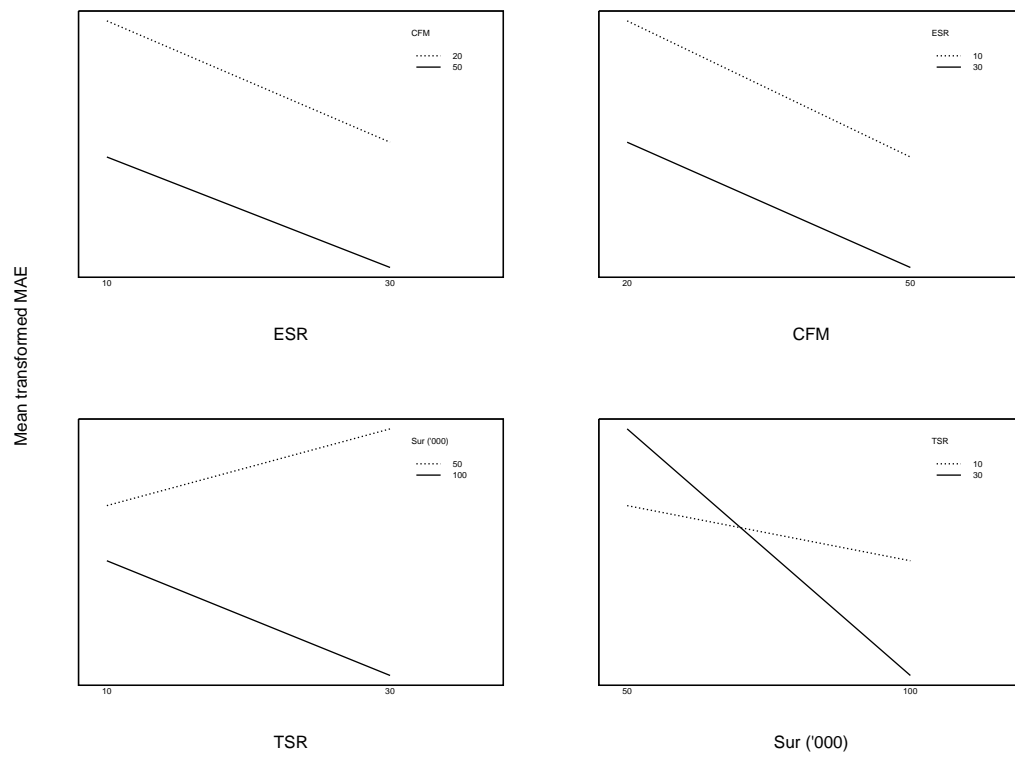


Figure 9: Significant *CFM*–*ESR* and *TSR*–*Sur* interactions on the transformed *MAE* for Watershed 3 simulated under alternative NS2.

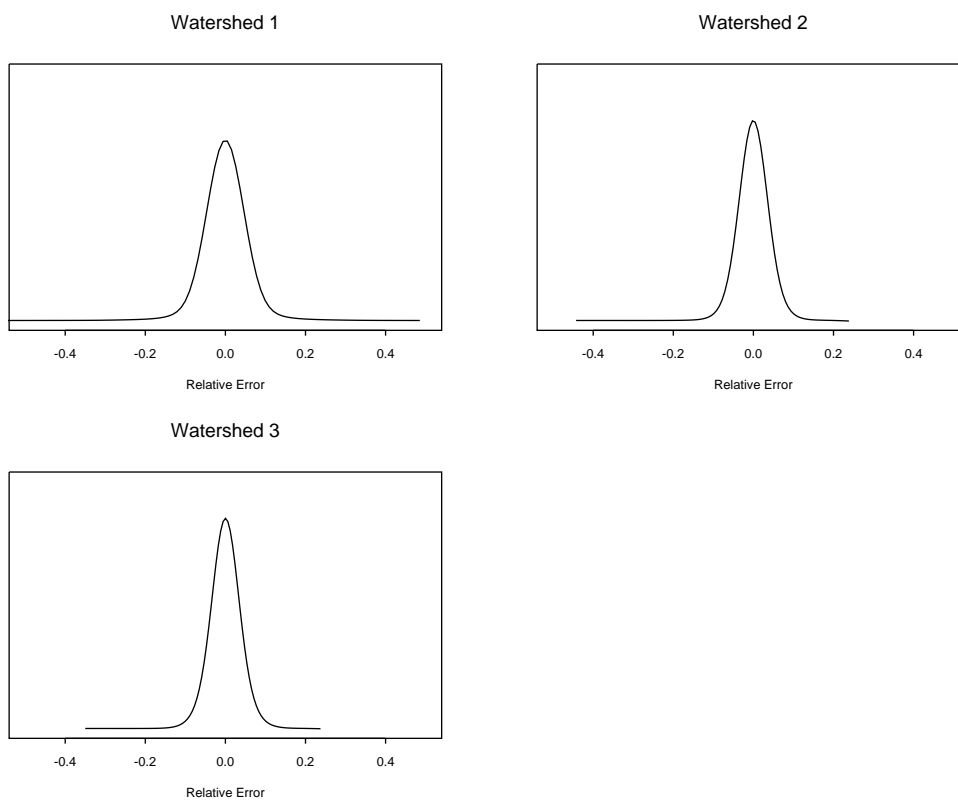


Figure 10: Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative AS1.

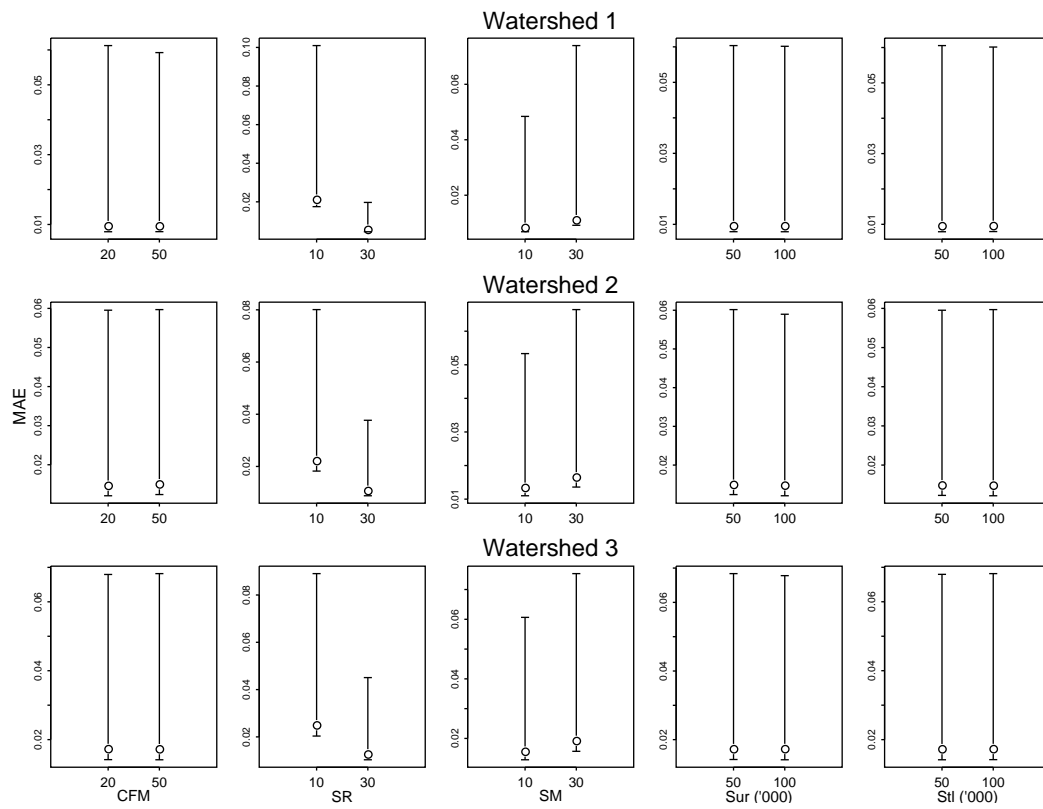


Figure 11: Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the *MAE* of estimated production for each treatment level and natural stock simulated under alternative AS1.

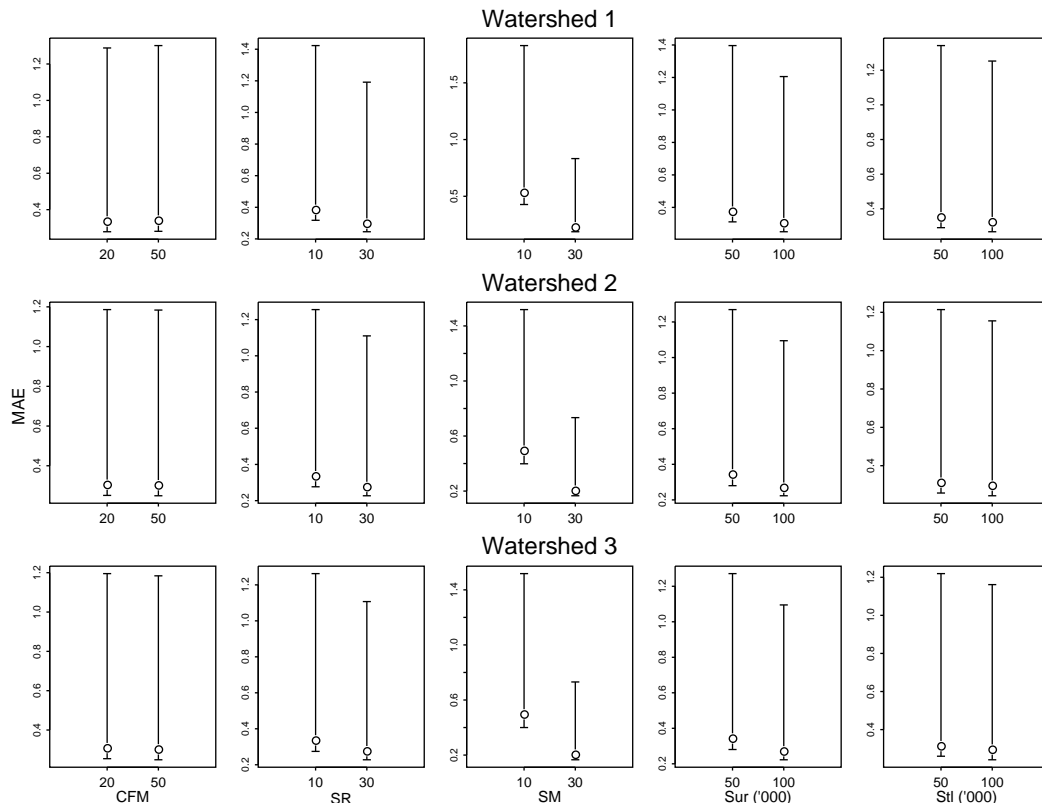


Figure 12: Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the MAE of $\hat{\Delta}_{nt}$ for each treatment level and natural stock simulated under alternative AS1.

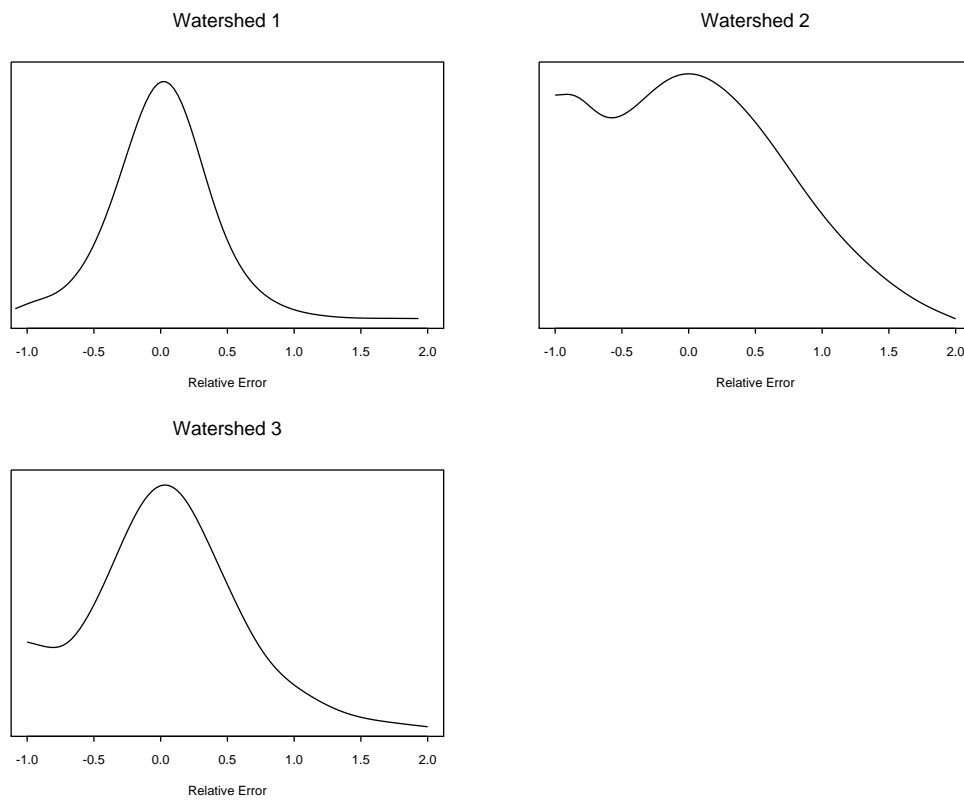


Figure 13: Watershed specific, smoothed, non-parametric densities of the relative estimation error for all treatment combinations in alternative AS2.

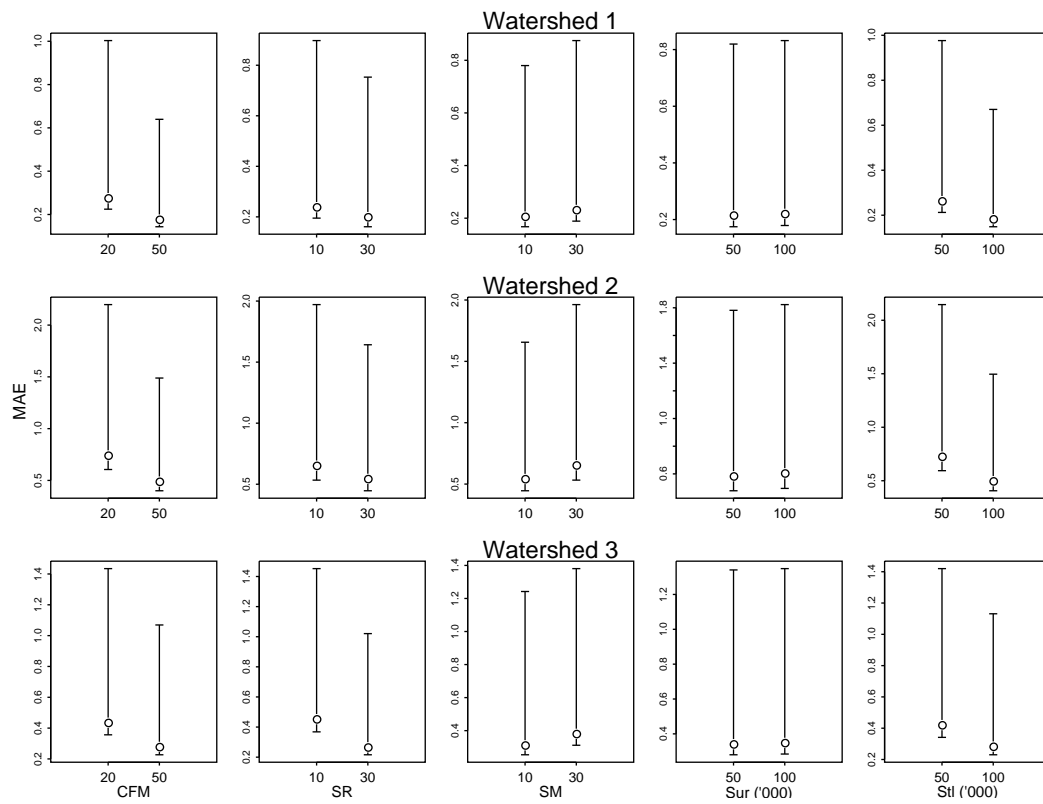


Figure 14: Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the *MAE* of estimated production for each treatment level and natural stock simulated under alternative AS2.

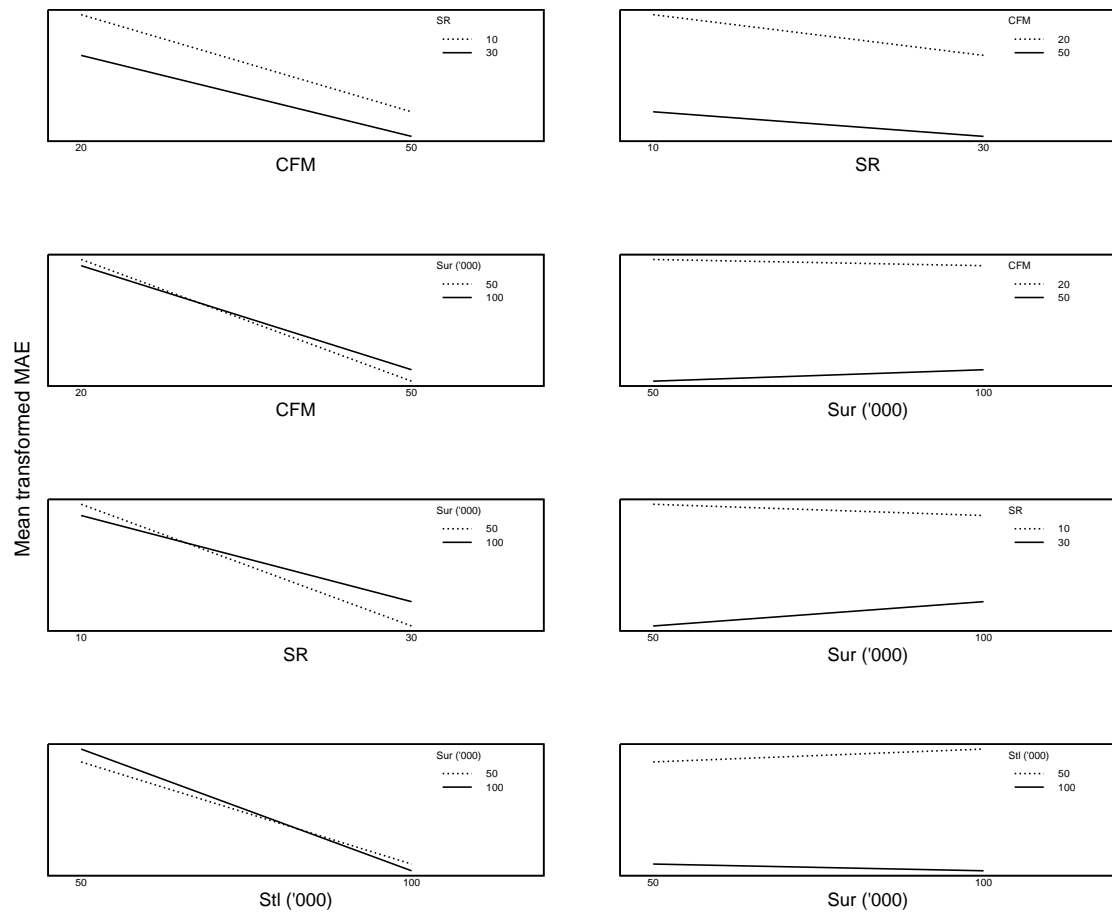


Figure 15: Significant *CFM–SR*, *CFM–Sur*, *SR–Sur*, and *Sur–Stl* interactions on the transformed *MAE* for Watershed 1 simulated under alternative AS2.

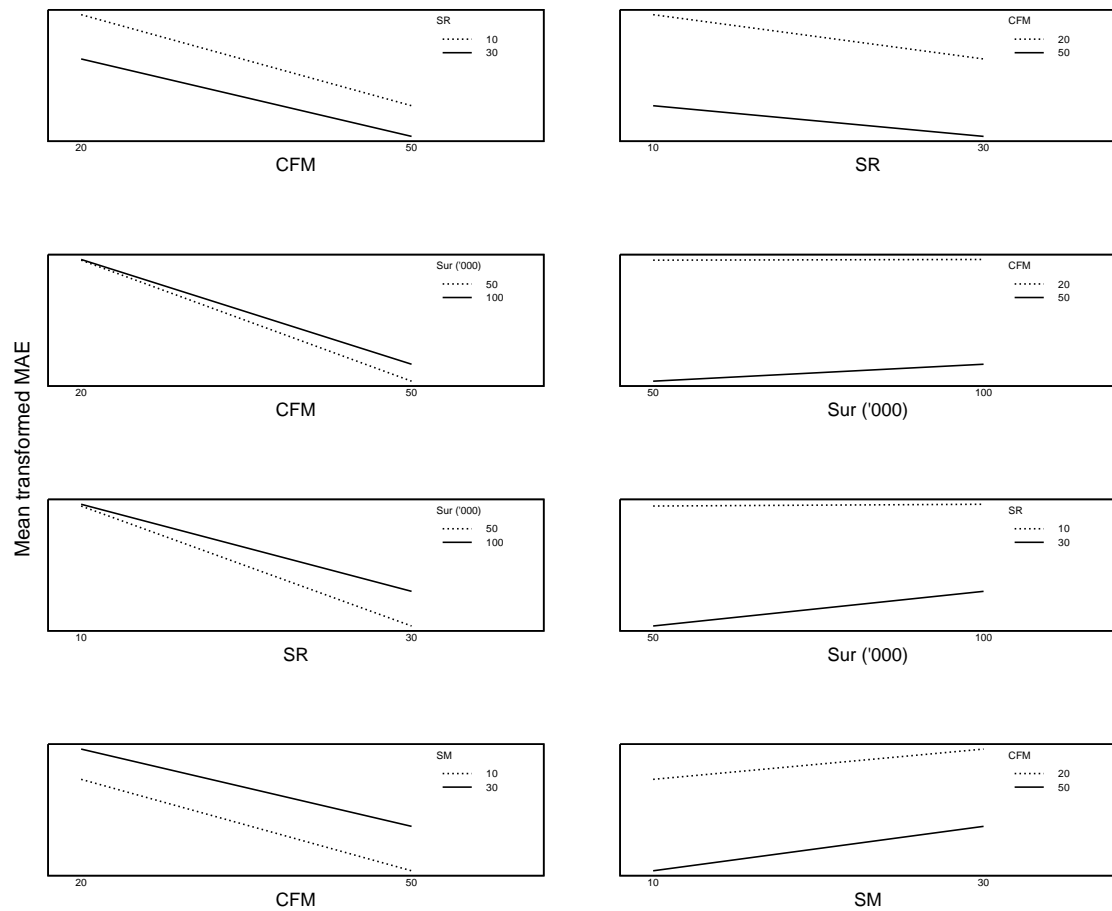


Figure 16: Significant *CFM-SR*, *CFM-Sur*, *SR-Sur*, and *CFM-SM* interactions on the transformed *MAE* for Watershed 2 simulated under alternative AS2.

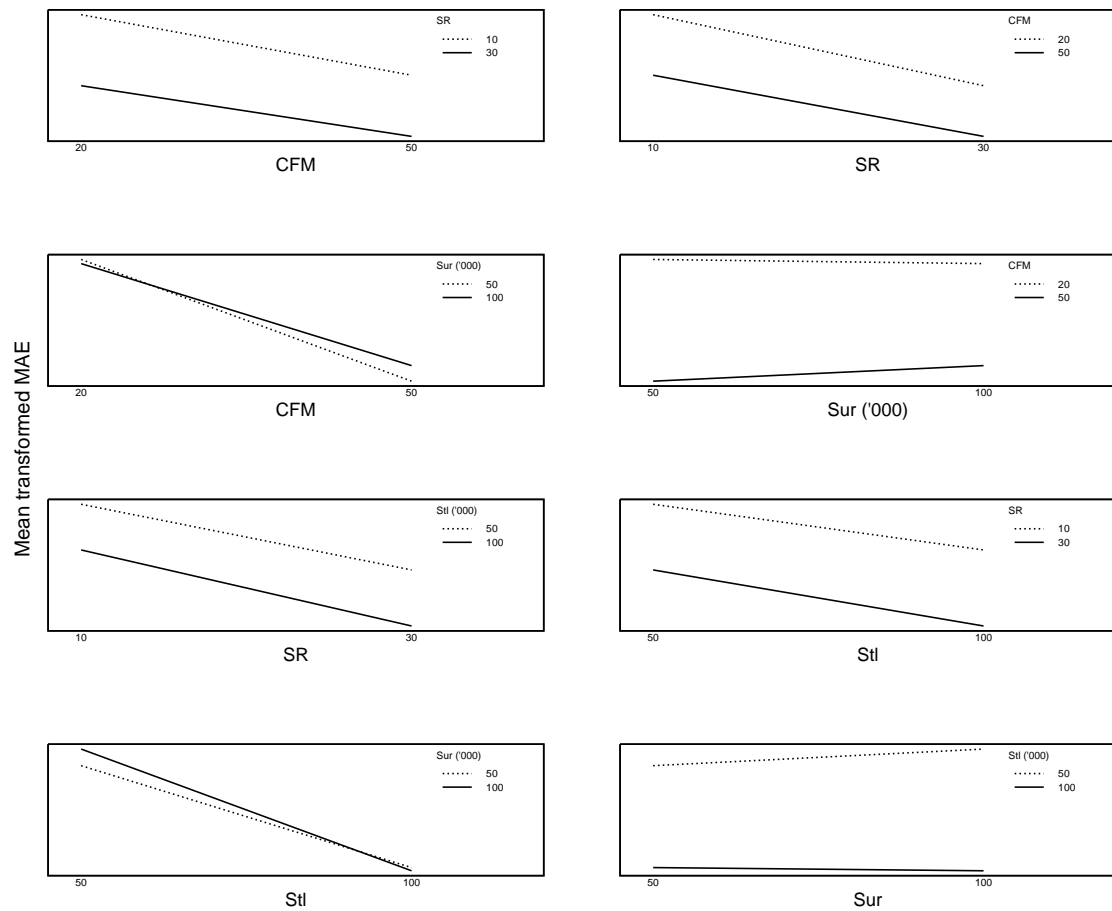


Figure 17: Significant *CFM-SR*, *CFM-Sur*, *SR-Stl*, and *Sur-Stl* interactions on the transformed *MAE* for Watershed 2 simulated under alternative AS2.

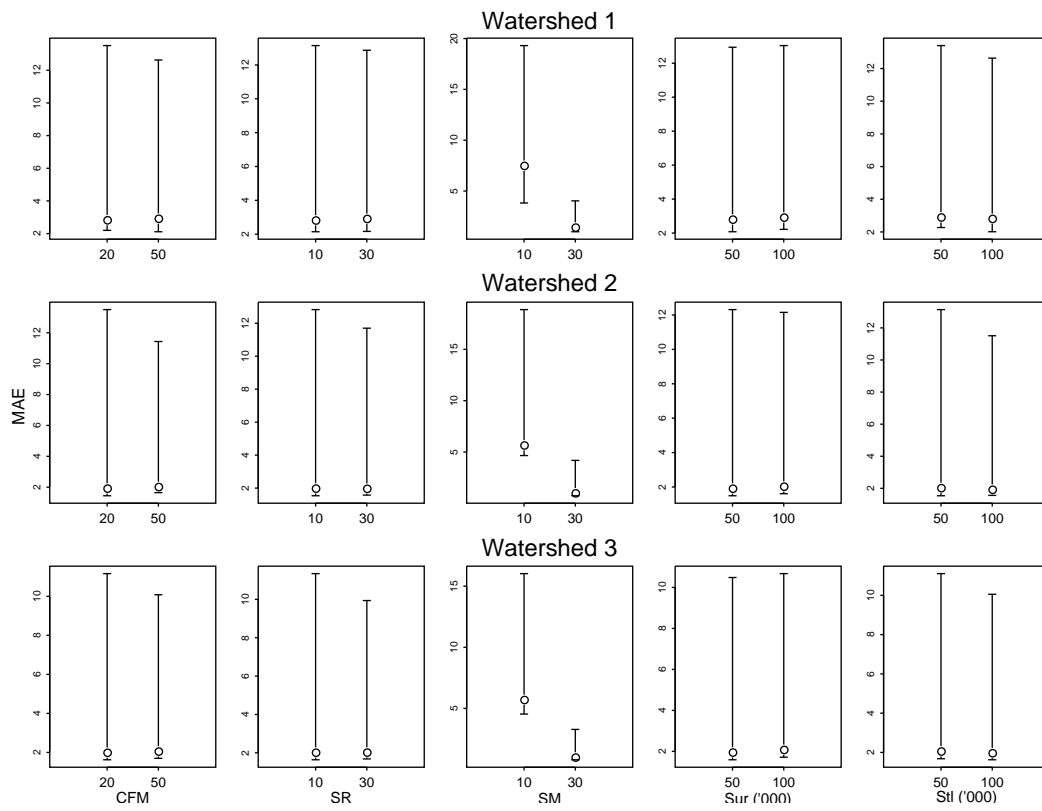


Figure 18: Medians (circles) and error bars defined by the 10th and 90th percentiles plotted for the MAE of $\hat{\Delta}_{nt}$ for each treatment level and natural stock simulated under alternative AS2.

A Estimating Δ_{it} in selective fisheries

To estimate Δ_{it} the age three through five survival rates, ϕ_3 , ϕ_4 , and ϕ_5 , are assumed known and constant, while initial survival and maturation are estimated using the recoveries of *Surrogate* fish. The approach can be partitioned into three parts: (a) estimating life history parameters using the *Surrogate* group; (b) estimating Δ_{et} for the *Stealth* group using the estimated parameters; (c) estimating Δ_{nt} for the natural stock using Δ_{et} and the escapement estimates for the *Stealth* and natural stocks.

A.1 Estimating life history parameters using the *Surrogate* group

The calculations are lengthy and the notation is quite complicated because multiple years of parameters are needed. The year of interest is denoted by a subscript t . Catches (C), escapements (E), and harvest rates (μ_O , μ_F , and μ_T) have three subscripts. The first is either b for the *Surrogate* group, e for the *Stealth* group, or n for natural fish. The second is age of the fish, $a=2, 3, 4$, or 5 , and the third subscript is the year of interest for the caught or escaping fish, t . For example, $C_{b4,t-1}$ is the catch of surrogate age 4 fish caught in the year prior to t . Release numbers (R) have two subscripts; b or e for the first, and brood year (not the year of interest) relative to the year of interest for the second, $t-a$. Survival rates (ϕ) are subscripted first by I , 3, 4, or 5 and then the year of interest. Similarly, maturation rates (σ) are subscripted by age of maturation, a , and the current year. The symbol P_F , with the same three subscripts as catches and escapement, denotes the sum of all of the freshwater observations for a given stock, including freshwater mainstem catch, terminal area catch, and escapement. For example, in year t , the age a freshwater return,

$$P_{Fbat} = C_{Fbat} + C_{Tbat} + E_{bat}.$$

The initial survival and maturation parameters can be estimated using the following equations.

$$\widehat{S}_{0,t-g} = \frac{\widehat{C}_{Ob2,t-g} + \widehat{P}_{Fb2,t-g} + \frac{\widehat{C}_{Ob3,t-g} + \widehat{P}_{Fb3,t-g} + \frac{\widehat{C}_{Ob4,t-g} + \widehat{P}_{Fb4,t-g} + \frac{\widehat{C}_{Ob5,t-g} + \widehat{P}_{Fb5,t-g}}{\phi_5}}{\phi_4}}{\phi_3}}{R_{b,t-g-2}} \quad g = 0, 1, 2, 3 \quad (54)$$

$$\widehat{\sigma}_{2,t-g} = \frac{\widehat{P}_{Fb2,t-g}}{R_{b,t-g-2}\widehat{\phi}_{I,t-g} - \widehat{C}_{Ob2,t-g}} \quad g = 0, 1, 2, 3 \quad (55)$$

$$\widehat{\sigma}_{3,t-g} = \frac{\widehat{P}_{Fb3,t-g}}{[R_{b,t-g-2}\widehat{\phi}_{I,t-g} - (\widehat{C}_{Ob2,t-g} + \widehat{P}_{Fb2,t-g})]\phi_3 - \widehat{C}_{Ob3,t-g}} \quad g = 1, 2, 3 \quad (56)$$

$$\widehat{\sigma}_{4,t-g} = \frac{\widehat{P}_{Fb4,t-g}}{[R_{b,t-g-2}\widehat{\phi}_{I,t-g} - (\widehat{C}_{Ob2,t-g} + \widehat{P}_{Fb2,t-g})]\phi_3 - (\widehat{C}_{Ob3,t-g} + \widehat{P}_{Fb3,t-g})]\phi_4 - \widehat{C}_{Ob4,t-g}} \quad g = 2, 3 \quad (57)$$

A.2 Estimating Δ_{et} for the *Stealth* group

The difference in escapement for the stealth group is calculated using the estimated initial survival and maturation parameters.

$$\widehat{\Delta}_{et} = \sum_{a=2}^5 \left[\widehat{\mathcal{E}}(E_{eat}|\theta_{at} = 0) - \widehat{E}_{eat} \right] \quad (58)$$

$$\widehat{\mathcal{E}}(E_{e2t}|\theta_{2t} = 0) = R_{e,t-2}\widehat{\phi}_{I,t}\widehat{\sigma}_{2t} \quad (59)$$

$$\widehat{\mathcal{E}}(E_{e3t}|\theta_{3t} = 0) = R_{e,t-3}\widehat{\phi}_{I,t-1}(1 - \widehat{\sigma}_{2,t-1})\phi_3\widehat{\sigma}_{3t} \quad (60)$$

$$\widehat{\mathcal{E}}(E_{e4t}|\theta_{4t} = 0) = R_{e,t-4}\widehat{\phi}_{I,t-2}(1 - \widehat{\sigma}_{2,t-2})\phi_3(1 - \widehat{\sigma}_{3,t-1})\phi_4\widehat{\sigma}_{4t} \quad (61)$$

$$\widehat{\mathcal{E}}(E_{e5t}|\theta_{5t} = 0) = R_{e,t-5}\widehat{\phi}_{I,t-3}(1 - \widehat{\sigma}_{2,t-3})\phi_3(1 - \widehat{\sigma}_{3,t-2})\phi_4(1 - \widehat{\sigma}_{4,t-1})\phi_5 \quad (62)$$

It should be noted that *Stealth* fish are assumed not to stray given this formulation. However, Appendix B explains a method for dealing with straying *Stealth* fish.

A.3 Estimating Δ_{nt} for the natural stock

To calculate $\widehat{\Delta}_{nt}$ for natural fish represented by the *Stealth* group, the expected escapement of natural fish in the absence of fishing mortality is first calculated using equations (59)-(62) and age-specific estimates of the natural stock and *Stealth* group escapements.

$$\widehat{\mathcal{E}}(E_{nat}|\theta_{at} = 0) = \widehat{\mathcal{E}}(E_{eat}|\theta_{at} = 0) \frac{\widehat{E}_{nat}}{\widehat{E}_{eat}} \quad (63)$$

The intuition behind equation (63) can be seen by substituting in expected values. For example, at age 2,

$$\begin{aligned} \widehat{\mathcal{E}}(E_{e2t}|\theta_{2t} = 0) \frac{\widehat{E}_{n2t}}{\widehat{E}_{e2t}} &\approx R_{e2,t-2}\phi_{I,t}\sigma_{2t} \frac{R_{n2,t-2}\phi_{nI,t}(1 - \mu_{O2t}^*)\sigma_{n2t}(1 - \mu_{F2t}^*)(1 - \mu_{Tj2t}^*)}{R_{e2,t-2}\phi_{I,t}(1 - \mu_{O2t}^*)\sigma_{2t}(1 - \mu_{F2t}^*)(1 - \mu_{Tj2t}^*)} \\ &= R_{n2,t-2}\phi_{nI,t}\sigma_{2t} \end{aligned}$$

The subscript n for the natural stock's survival rate and maturation parameters is included to emphasize the point that these parameters can differ between *Stealth* and natural fish.

The age-specific natural stock escapement estimates, \widehat{E}_{nat} , would be calculated as for NS1, using equation (26). Then, using the age-specific escapement estimates and the estimate of the expected escapement for natural fish,

$$\widehat{\Delta}_{nt} = \sum_{a=2}^5 \left[\widehat{\mathcal{E}}(E_{nat}|\theta_{at} = 0) - \widehat{E}_{nat} \right] \quad (64)$$

B Estimating Δ_{nt} when *Stealth* fish stray

To correctly estimate the expected escapement of natural fish (equation (63)), the *Stealth* group must not stray, as is assumed for natural fish, or the straying rate (ρ) must be accounted for. The effect of straying can be seen below, using age 2 in the AS case as an example.

$$\begin{aligned}\widehat{\mathcal{E}}(E_{n2t}|\theta_{2t} = 0) &= \widehat{\mathcal{E}}(E_{e2t}|\theta_{2t} = 0) \frac{\widehat{E}_{n2t}}{\widehat{E}_{e2t}} \\ &\approx R_{e,t-2} \phi_{I,t} \sigma_{2t} \frac{R_{n,t-2} \phi_{nI,t} (1 - \mu_{O2t}^*) \sigma_{2t} (1 - \mu_{F2t}^*) (1 - \mu_{Tj2t}^*)}{R_{e,t-2} \phi_{I,t} (1 - \mu_{O2t}^*) \sigma_{2t} (1 - \mu_{F2t}^*) \sum_{j=1}^k \rho_{2jt} (1 - \mu_{Tj2t}^*)} \\ &= \frac{R_{n,t-2} \phi_{nI,t} \sigma_{2t} (1 - \mu_{Tj2t}^*)}{\sum_{j=1}^k \rho_{2jt} (1 - \mu_{Tj2t}^*)}\end{aligned}$$

The straying rate, ρ , is now subscripted with age and brood year, while the stock subscript is assumed. This is because age and cohort variation may exist in the straying rates.

The bias may be reduced somewhat by using the *Stealth* escapement to the natal watershed j only, assuming $\rho_{2j} \approx 1$.

$$\begin{aligned}\widehat{\mathcal{E}}(E_{n2t}|\theta_{2t} = 0) &= \widehat{\mathcal{E}}(E_{e2t}|\theta_{2t} = 0) \frac{\widehat{E}_{n2t}}{\widehat{E}_{e_j2t}} \\ &\approx R_{e,t-2} \phi_{I,t} \sigma_{2t} \frac{R_{n,t-2} \phi_{nI,t} (1 - \mu_{O2t}^*) \sigma_{2t} (1 - \mu_{F2t}^*) (1 - \mu_{Tj2t}^*)}{R_{e,t-2} \phi_{I,t} (1 - \mu_{O2t}^*) \sigma_{2t} (1 - \mu_{F2t}^*) \rho_{2jt} (1 - \mu_{Tj2t}^*)} \\ &= \frac{R_{n,t-2} \phi_{nI,t} \sigma_{2t}}{\rho_{2jt}}\end{aligned}$$

Rather than assume that the straying to non-natal watersheds is insignificant, the probability of returning to the natal watershed can be estimated using the *Surrogate* group, which are also used to calculate survival, maturity, and harvest parameters for stealth fish. However, as with the other estimates, the implicit assumption that *Surrogate* and the *Stealth* fish have the same life history parameters must hold, or bias will occur.

$$\widehat{\rho}_{jat} = \frac{\widehat{C}_{bjat} + \widehat{E}_{bjat}}{\sum_{j'=1}^k [\widehat{C}_{bj'at} + \widehat{E}_{bj'at}]}$$

Using the above estimate $\widehat{\rho}_{jat}$, the estimate of the expected natural escapement without any incidental mortality is

$$\widehat{\mathcal{E}}(E_{nat}|\theta_{2t} = 0) = \widehat{\mathcal{E}}(E_{eat}|\theta_{2t} = 0) \widehat{\rho}_{jat} \frac{\widehat{E}_{nat}}{\widehat{E}_{e_jat}}$$