

I. AYKSSI TITLE PAGE

Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative  
Project Final Product<sup>1</sup>

EVALUATING ASSESSMENT STRATEGIES FOR KUSKOKWIM RIVER CHINOOK SALMON

by:

**Matthew J. Catalano<sup>2</sup>, Benjamin A. Staton<sup>2</sup>, Troy Farmer<sup>2</sup>, Daniel C. Gwinn<sup>3</sup>, Steven Fleischman<sup>4</sup>**

<sup>2</sup> School of Fisheries, Aquaculture, and Aquatic Sciences, Auburn University, Alabama

<sup>3</sup> Biometric Research, LLC

<sup>4</sup> Alaska Department of Fish and Game, Sport Fish Division

June 17, 2016

<sup>1</sup> Final products of AYK Sustainable Salmon Initiative-sponsored research are made available to the Initiatives Partners and the public in the interest of rapid dissemination of information that may be useful in salmon management, research, or administration. Sponsorship of the project by the AYK SSI does not necessarily imply that the findings or conclusions are endorsed by the AYK SSI.

## II. ABSTRACT

Evaluating hypotheses for AYK Chinook salmon (*Oncorhynchus tshawytscha*) declines and managing these stocks requires that we obtain reliable estimates of stock status, trends, and productivity. These estimates have been obtained from a novel drainage-wide run reconstruction model developed specifically for this data-limited case. However, this model may have potential pitfalls that warrant further investigation. In particular, the model may be susceptible to bias stemming from incomplete temporal and spatial coverage of run enumeration projects and natural between-stock variation in population dynamics within a drainage. Although the model estimates total drainage-wide abundance, spawner recruit analyses are necessary for setting management advice and these analysis are typically completed in a second separate analysis from the abundance reconstruction, which may not allow for realistic propagation of uncertainty from observation level data to quantities of interest to managers such as Smsy. Numerous assumptions regarding sampling precision and bias are also made in these model and these assumptions warrant further investigation.

To address these issues we used a series of computer simulation-estimation analyses to test the Kuskokwim River run reconstruction and spawner recruit analyses. In Chapter 1 of this report we evaluated the influence of (1) among-stock within-drainage temporal variation in recruitment, (3) data quantity, and (4) data type on the accuracy and precision of these models. We will also considered the costs of data collection scenarios in light of model performance. In Chapter 2, we introduce and evaluate an integrated Bayesian approach to the analysis that combines the run abundance reconstruction and the spawner-recruit analysis into a single model to potentially allow for a more appropriate accounting of uncertainty. We then evaluated the sensitivity of this model to assumptions regarding data accuracy and precision. In Chapter 3, we revisit the run abundance reconstruction (specifically the 2014 version of the model) and evaluate several alternative models that represent different approaches to weighting the various data sources available to the model.

One of the most important data inputs to the run abundance reconstruction model are the telemetric mark-recapture drainage-wide escapement estimates. These estimates are the sole source of population scaling information for the run abundance reconstruction model and are therefore of critical importance. In Chapter 4, we demonstrate a new Bayesian analysis of the 2003-2007 telemetric mark recapture analysis that make full use of the spatial and temporal structure of the telemetry data to provide estimates of tributary-specific abundance and run timing at the Birch Tree Corssing tagging site.

In data limited situations, we often lack sufficient data to estimate stock status and Smsy. In these cases, alternative methods for informing these quantities are sought. One option is to develop relationships between easily measured basin scale habitat variables and salmon carrying capacity and/or productivity. In Chapter 5, we explored such relationships for Alaskan Chinook salmon via a model selection exercise.

Each of these analyses represents one component of a broader effort to evaluate methods for salmon stock assessment in data-limited situations, specifically in the context of Kuskokwim Chinook salmon, a stock of great cultural, economic and ecological importance. Below, we provide a brief overview of each of these analyses.

### *Chapter 1: The performance of a drainage-wide salmon run reconstruction model under alternative data collection schemes and population dynamics assumptions*

Kuskokwim River Chinook salmon abundance is estimated using a drainage-wide run reconstruction model. The model is fitted to aerial survey escapement indices, weir escapement counts, harvest estimates, and occasional drainage-wide mark-recapture estimates of total escapement. The performance of this model at estimating drainage-wide escapement (and abundance) has yet to be evaluated. We assessed the performance of the model, in terms of bias and precision of drainage-wide escapement estimates, under a range of scenarios that involved different data collection schemes and

varying assumptions regarding population dynamics. Different data collection schemes included a range of the number of weirs and aerial surveys sampled annually, and the frequency of drainage-wide mark-recapture total escapement estimates. We also considered the cost of data types to provide guidance on their relative efficiency. In terms of population dynamics assumptions, we assessed the sensitivity of model performance to the magnitude of variation in tributary stock contributions to total escapement, recruitment variation, and among-tributary correlations in recruitment variation. As expected, mean bias of escapement estimates decreased and precision increased with increasing numbers of aerial surveys, weirs, and mark-recapture total escapement estimates. However, we observed diminishing returns with respect to bias and precision with greater than approximately 10 aerial surveys and 6 weirs per year, and more than two sets of three-year mark recapture drainage-wide escapement estimates per 40 years. No optimal data configuration was found; for example, similar estimates of bias and precision could be achieved by increasing the number of aerial surveys while decreasing the number of weirs, and vice versa. Overall, average bias (mean relative error) was  $<0.06$  (6%) and the average CV of escapement estimates was  $<0.1$  (10%) across all data scenarios we evaluated. These values are low enough to suggest the model performed reasonably well at recovering the true escapement estimates that were simulated, at least under the range of scenarios we evaluated. Bias and precision of escapement estimates were sensitive to the degree of correlation in recruitment variability among tributary stocks, the magnitude of recruitment variation, and precision of drainage wide mark-recapture escapement estimates. An analysis of the cost of different data types (aerial survey, weir, mark-recapture), indicated that aerial surveys provide a substantially greater improvement in the precision of escapement estimate per unit cost than the other data types. This finding suggests that increasing the number of aerial surveys may be a relatively cost effective way to improve escapement estimates, at least in the context of this particular run reconstruction model. We caution that although aerial surveys were more cost effective than weirs, our analysis did not consider additional advantages of weirs beyond escapement estimation such as their critical importance for collecting age, sex and length data, and as a recapture point for tagging studies. Thus, weirs will likely continue to be an important component of any effective data collection scheme into the future. Although the drainage-wide run reconstruction appeared to perform well in our simulation, we recommend that the model continue to be challenged via simulation to assess how well it performs under additional assumption violations such as systematic variation (i.e., temporal trends) in the relative contribution of each tributary to the total escapement.

## *Chapter 2: Evaluation of an integrated run reconstruction and spawner-recruit model for Kuskokwim River Chinook salmon*

Stock assessment scientists are often faced with decisions regarding how to incorporate information into models. One primary decision revolves around how estimates that are summaries of raw data should be treated (e.g., abundance estimates derived from relative indices of abundance). The choice is to fit the terminal model after a sequence of models that produce estimates used as data or to integrate the raw data into the terminal analysis. Each approach has advantages and disadvantages that constitute a suite of trade-offs which are described in further detail. These trade-offs are investigated here by comparing a sequential analysis to an integrated analysis for Pacific salmon stock assessment, using the Kuskokwim River Chinook salmon stock, located in western Alaska, as a case study. The terminal analysis in both cases is an age-structured Bayesian state-space spawner-recruit model used to derive estimates of productivity and capacity for use in setting escapement targets. Primary findings showed that the two approaches returned similar estimates of population dynamics parameters and management reference points, both in terms of point estimates and uncertainty, showing that the treatment of measurement uncertainty in the sequential analysis was an adequate treatment of the information content of the data. The primary advantage of the integrated analysis was the added

realism of sharing calendar year abundance data between brood years, but came at the cost of slow run times. This exercise showed that while there is a trade-off between sequential and integrated analyses in terms of model complexity and realism, it may not be large enough to warrant an integrated analysis in all cases. We also investigated the sensitivity of estimates of total abundance from the integrated model to changes in the assumed variances of harvest observations, and the assumed precision and bias of the telemetric mark-recapture estimates of drainage-wide escapement. We found that drainage-wide estimates of escapement were insensitive to changes in the assumed sampling variance of harvest data and telemetric mark-recapture escapement estimates. However, the escapement estimates were very sensitive to bias in the telemetric mark-recapture data as the estimates rescaled proportionally to accomodate the change in the mark-recapture data. These findings highlight the common-sense notion that obtaining accurate estimates of drainage-wide escapement from telemetric mark-recapture studies is very important in the pursuit of unbiased estimates from this run reconstruction model.

*Chapter 3: Description of the Kuskokwim River Chinook salmon run reconstruction and an investigation of data weighting: a report to the Kuskokwim River Salmon Management Working Group*

We conducted a suite of independent analyses investigating the consequences of different ways of weighting weir and aerial survey data in the 2014 run reconstruction model for Kuskokwim River Chinook salmon. The purpose of the analysis was to provide the Kuskokwim Salmon management Working Group with an independent review of a few plausible modeling approaches in preparation for the 2014 salmon run. We concluded that the model structure was generally consistent with broadly applied and accepted stock assessment modeling approaches. We investigated the results of 5 different models. Four of the models were different approaches to weighting weirs and aerial surveys, and the fifth incorporated the 2014 mark-recapture abundance estimate: (1) ADF&G base model: individual weights are estimated for each escapement monitoring project, (2) Pooled weights model: a common weight is estimated for all weir projects and one common weight is estimated for all aerial survey projects, (3) No 2014 weirs model: 2014 weir counts are excluded from the ADF&G base model, (4) No 2014 aerial surveys model: 2014 aerial surveys are excluded from the ADF&G base model, and (5) With MRC model: the ADF&G model including the 2014 mark-recapture estimate. All of the models resulted in the same major conclusion that the 2014 run was lower than average and the drainage-wide escapement goal was met. The estimated weights for weirs were higher than for the aerial surveys, which means that, all else being equal, the model's predictions will more closely match the weir data overall. Retrospective analyses showed that the ADF&G base model run estimates varied more than the pooled weights model as more data were added, but neither model exhibited worrisome retrospective patterns. The ADF&G base model exhibited an undesirable tendency to perfectly fit to one of the weir time series in some years. This behavior is extreme and warrants further investigation and consideration of alternate models because it suggests that in those cases the model treated one of the weirs as a perfect indicator of escapement. This extreme behavior occurred in two out of eight years in the retrospective analysis: 2007 and 2010. Both of these years were not years in which the model was used to make management decisions. Future modeling efforts should continue to investigate alternate weighting schemes and model structures. The pooled weights model we constructed represents a sensible first step in this direction, but there are likely many other potential model structures that could be considered.

*Chapter 4: A bayesian model to estimate Kuskokwim River Chinook tributary run timing and abundance from telemetric mark-recapture data*

We conducted a new analysis of the telemetric mark-recapture data from Kuskokwim River Chinook salmon from 2003-2007 to demonstrate a new analytical approach that has the potential to take advantage of the full range of spatial and temporal information contained in these data.

Specifically we sought to quantify estimates of abundance and run timing for tributary-specific Chinook salmon stocks in the Kuskokwim drainage while simultaneously estimating total escapement upstream of the tagging site. To accomplish this goal, we used a Bayesian modification of the maximum likelihood methods described in Bromaghin et al. (2010). Across all years, five of the twelve tributary groups (Mainstem [mean proportion = 0.15], Aniak & Middle Tributaries [0.19], Holitna River [0.09], Hoholitna River [0.18], and Kogruklu River [0.11]) were consistently the largest contributing groups to the run across all time strata. On average these groups composed 0.72 of the total run. Tributary-specific temporal patterns also existed. Chinook salmon from upstream tributary groups tended to complete their migrations earlier than those from middle and Holitna tributary groups. Two upstream tributary groups, Upper Kuskokwim Tributaries and Swift River, contributed large proportions to the run during the first two time strata (22 June & earlier), but their contributions declined during the final three time strata (23 June & later). While this trend was clearly evident from the run timing proportions from each tributary group, the proportional composition of the run indicated that even though these upper tributary groups had higher run timing proportions early in the run they were less abundant than middle or Holitna tributary groups at the tagging site at this time due to the high abundance of the middle tributary stocks. Annual trends in total abundance were similar to those found in Schaberg et al. (2012) using the Lincoln-Petersen estimator. While annual trends were similar, the Bayesian model tended to predict lower total abundance than Schaberg et al. (2012) in most years (2003-2006; range: 15 – 23% lower), with the difference between the Bayesian model and Schaberg et al. (2012) model estimates increasing as population abundance increased (Figure 1). In conclusion, the finer-scale information on run timing and abundance provided by the Bayesian model should be useful to ongoing management efforts aimed at understanding the diversity of Chinook salmon spawning stocks within the Kuskokwim River Drainage.

#### *Chapter 5: Associations between Alaskan Chinook salmon population characteristics and drainage-wide habitat variables*

Management strategies for Alaska Chinook salmon rely on biological reference points (e.g.,  $S_{MSY}$ ) to set harvest targets which are traditionally derived using spawner-recruit analysis. However, not all stocks in Alaska are sampled intensively enough to allow for derivation of reference points through spawner-recruit analyses, yet their management must still be based on stock-specific reference points. Habitat-based methods have been developed to predict  $S_{MSY}$  for stocks without adequate spawner-recruit information based on a subset of stocks that do have this information. These models form a linear predictive relationship between  $S_{MSY}$  and certain habitat characteristics (e.g., drainage area) that can be used for more data-limited stocks. Here I build on this work by developing a hierarchical modeling approach for Alaskan Chinook salmon stocks that incorporates the uncertainty in stock-specific  $S_{MSY}$  and applying a suite of model selection techniques to select habitat variables that have predictive credibility. Findings showed that hierarchical inclusion of uncertainty in  $S_{MSY}$  did not remove the relationship between drainage area and  $S_{MSY}$ , showing that this relationship may still be used even when including this additional source of variation. Three of the four variable selection techniques consistently agreed on the best variables. The deviance information criterion performed poorly at distinguishing models and potential explanations are discussed. Drainage area as the sole predictor was unanimously chosen by the best model under the three remaining model selection approaches.

### III. PRESS RELEASE

Having reliable estimates of stock status, carrying capacity, and trends in abundance are critical for effective management of Pacific salmon. In Western Alaska, data on these salmon stocks is often difficult to collect due to extreme weather, the vast expanse of the territory, and poor access. In these situations it is important to develop sound methods for stock assessment that can deal with limitations in the quantity and quality of the data. Fortunately, such methods exist, particularly for the Kuskokwim River Chinook salmon stock. However, these methods have not yet been fully evaluated to see how well they characterize the status of this stock. A team of researchers from Auburn University took on this task by evaluating the performance of stock assessment for Kuskokwim Chinook in a project funded by the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative. A trick that scientists use to test how well a model is working is to simulate plausible fake data on a computer then run these data through the stock assessment model to see how well it can recover the true status of the fish stock that was simulated. The Auburn team used this approach to test the Kuskokwim Chinook assessment model and to evaluate what types of data make the model perform best and to see which ones are most cost-effective. They found that the model generally performs well and provides reliable estimates of spawning abundance and that adding more aerial surveys (that's counting fish in the river from above in an airplane) provided the most bang for the buck in terms of giving us the most reliable estimates at the least cost. That's not to say that other types of data such as weirs (imagine a fence across the stream so that each salmon can be individually counted as it passes upstream to spawn) are not needed. Weirs are important because they allow biologists to get their hands on fish to measure and weigh them. But the team found that at least when it comes to estimating the total spawning abundance, air surveys did an adequate job but were so much cheaper than weirs that they ended up being the most cost-effective. The Auburn team also conducted a sensitivity analysis to evaluate if the Kuskokwim Chinook model results would change if the model was set up differently. In general they found that the model results were consistent unless major changes were made to the data inputs. For example, the estimates of spawning abundance from the model increased substantially if they assumed that the fish tagging data fed into the model gave an inaccurate picture of the abundance of the stock. Ultimately, the Auburn team hoped to provide fishery managers and stakeholders with better tools to help them assess the status of the stock and make carefully-considered management decisions.

#### IV. PROJECT EVALUATION

The proposed project had four objectives as follows:

*Objective 1: Assess the influence of population dynamics, data quantity, data quality and type on the accuracy, precision and cost of estimates from drainage wide (i.e., stock-aggregated) Bayesian state-space run reconstruction models.* We completed this objective as originally planned. We report on our findings related to this objective in Chapter 1 of this report. The only deviation from the original proposed plan was that instead of evaluating the performance of the integrated Bayesian run reconstruction/stock-recruitment, we focused instead on just a run reconstruction model. The integrated model was far too complex and slow-running to realistically evaluate its performance relative to population dynamics assumptions, data quantity, data quality and type. We decided it would be better to do a thorough job evaluating the run reconstruction model rather than a pared down evaluation of the integrated model.

*Objective 2: Evaluate the sensitivity of estimates of productivity, abundance, and management quantities (e.g., Smsy) for the Kuskokwim River Chinook salmon stock to adoption of a Bayesian state-space approach.* We completed this analysis as planned. The demonstration and sensitivity of the integrated model is in Chapter 2 of this report. We submitted a version of Chapter 2 for peer review in the journal Fisheries Research. We also conducted two additional analyses that were not mentioned in the original proposal but fall under this objective. We investigated the sensitivity of the run reconstruction model to assumption regarding data weighting. This analysis was requested by the Kuskokwim Salmon Management Working Group, we include it here because of its close relationship with this project. We report on this analysis in Chapter 3.

*Objective 3: Conduct exploratory analysis of new Bayesian modeling approaches for telemetric mark-recapture using the Kuskokwim River Chinook stock as a test case.* We completed this analysis as originally proposed and report on it in Chapter 4.

*Objective 4: Analyze associations between Alaskan Chinook salmon abundance (and Smsy) and drainage basin area.* We completed this analysis as originally proposed and report on it in Chapter 5.

## **V. DELIVERABLES**

The findings of our project have been and will continue to be disseminated via conference and management meeting presentations and peer-reviewed manuscripts. We have completed six presentations, attended three meetings, and submitted one manuscript for peer review publication. We anticipate submitting an additional three to four manuscripts by the end of 2016. We have also submitted with this report all model code for each analysis as appendices at the end of each chapter.

### **Reports:**

Semiannual progress reports July 2014, January and July 2015, January 2016.

Staton, B. A., M. J. Catalano, and S. Fleischman. 2014. Overview of the Kuskokwim River Chinook salmon stock assessment and the application of a Bayesian state-space run reconstruction with integrated stock-recruitment productivity. An preliminary report to the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative. Anchorage, Alaska.

### **Presentations:**

Staton, B. A. and M. J. Catalano. 2015. Evaluation of a Bayesian state-space stock assessment model for Kuskokwim River Chinook Salmon. American Fisheries Society Conference. Portland, Oregon.

Staton, B. A. and M. J. Catalano. 2015. Development and evaluation of a habitat-based assessment model for Alaskan Chinook salmon. Dauphin Island Sea Lab Graduate Student Symposium. Dauphin Island, Alabama.

Staton, B. A., M. J. Catalano, and S. Fleischman. March 2015. An integrated Bayesian state-space run reconstruction for Kuskokwim River Chinook salmon. Kuskokwim River Interagency Meeting. Bethel, Alaska.

Staton, B. A., M. J. Catalano, and S. Fleischman. 2014. Stock assessment in data-limited situations: using run reconstruction with integrated stock-recruitment analysis on Kuskokwim River Chinook salmon. Alaska Chapter of the American Fisheries Society Conference. Juneau, Alaska.

Catalano, M. J. September 2014. An overview of uncertainty analysis in salmon management. Kuskokwim River Salmon Management Working Group: Escapement Goal Work Session. Bethel, Alaska.

Catalano, M. J. and S. J. Fleischman. April 2014. Evaluating data limited run reconstruction models for Kuskokwim River Chinook salmon. Kuskokwim River Interagency Meeting. Anchorage, Alaska.

### **Manuscripts submitted:**

Staton, B. A., M. J. Catalano, and S. J. Fleischman. Submitted May 2016. From sequential to integrated Bayesian analyses: exploring the continuum with a Pacific salmon spawner-recruit model. Fisheries Research

### **Meetings Participated:**

2015 Kuskokwim River Interagency Meeting. Bethel, Alaska.

2014 Kuskokwim River Salmon Management Working Group: Escapement Goal Work Session. Bethel, Alaska.

2014 Kuskokwim River Interagency Meeting. Anchorage, Alaska.



## **VI. PROJECT DATA SUMMARY**

Our analysis produced simulated data sets and parameter estimates from Bayesian and maximum likelihood assessment models. All model outputs are available upon request from the PI.

## **VII. APPENDIX: SUBMITTED OR DRAFT MANUSCRIPTS**

### **CHAPTER 1**

#### **THE PERFORMANCE OF A DRAINAGE-WIDE SALMON RUN RECONSTRUCTION MODEL UNDER ALTERNATIVE DATA COLLECTION SCHEMES AND POPULATION DYNAMICS ASSUMPTIONS**

##### **ABSTRACT**

Kuskokwim River Chinook salmon abundance is estimated using a drainage-wide run reconstruction model. The model is fitted to aerial survey escapement indices, weir escapement counts, harvest estimates, and occasional drainage-wide mark-recapture estimates of total escapement. The performance of this model at estimating drainage-wide escapement (and abundance) has yet to be evaluated. We assessed the performance of the model, in terms of bias and precision of drainage-wide escapement estimates, under a range of scenarios that involved different data collection schemes and varying assumptions regarding population dynamics. Different data collection schemes included a range of the number of weirs and aerial surveys sampled annually, and the frequency of drainage-wide mark-recapture total escapement estimates. We also considered the cost of data types to provide guidance on their relative efficiency. In terms of population dynamics assumptions, we assessed the sensitivity of model performance to the magnitude of variation in tributary stock contributions to total escapement, recruitment variation, and among-tributary correlations in recruitment variation. As expected, mean bias of escapement estimates decreased and precision increased with increasing numbers of aerial surveys, weirs, and mark-recapture total escapement estimates. However, we observed diminishing returns with respect to bias and precision with greater than approximately 10 aerial surveys and 6 weirs per year, and more than two sets of three-year mark recapture drainage-wide escapement estimates per 40 years. No optimal data configuration was found; for example, similar estimates of bias and precision could be achieved by increasing the number of aerial surveys while decreasing the number of weirs, and vice versa. Overall, average bias (mean relative error) was  $<0.06$  (6%) and the average CV of escapement estimates was  $<0.1$  (10%) across all data scenarios we evaluated. These values are low enough to suggest the the model performed reasonably well at recovering the true escapement estimates that were simulated, at least under the range of scanarios we evaluated. Bias and precision of escapement estimates were sensitive to the degree of correlation in recruitment variability among tributary stocks, the magnitude of recruitment variation, and precision of drainage wide mark-recapture escapement esitmates. An analysis of the cost of different data types (aerial survey, weir, mark-recapture), indicated that aerial surveys provide a substaitially greater improvement in the precision of escapement esimtate per unit cost than the other data types. This finding suggests that increasing the number of aerial surveys may be a realatively cost effective way to improve escapement estimates, at least in the context of this particular run reconstruction model. We caution that although aerial surveys were more cost effective than weirs, our analysis did not consider additional advantages of weirs beyond escapement estimation such as their critical importance for collecting age, sex and length data, and as a recapture point for tagging studies. Thus, weirs will likely continue to be an important component of any effective data collection scheme into the future. Although the drainage-wide run reconstruction appeared to perform well in our sinmulation, we recommend that the model continue to be challenged via simulation to assess how well it performs under additional assumption violations such as systematic variation (i.e., temporal trends) in the relative contribution of each tributary to the total escapement.

## INTRODUCTION

Monitoring is important for the effective management of exploited fish populations, such as Pacific salmon. For example, monitoring provides information on the state of the stock which is necessary for state-dependent management (Shea et al. 1998), and when undertaken in an adaptive framework can reveal the best management practices (Walters 1986). Furthermore, annual monitoring data can provide estimates of salmon spawner abundance and recruitment, which can be used to describe population dynamics and set management reference points to guide management goals (e.g. Fleischman et al. 2013). Unfortunately, monitoring data are typically expensive and may often be inadequate to appropriately inform management (Yoccoz et al. 2002). Thus, many salmon fisheries must be managed under information-limited situations (Shotwell and Adkison 2004).

When information for management is limited, run-reconstruction models that assimilate multiple data types can be useful. These models take advantage of the many data sources collected on salmon fisheries and are particularly useful when fisheries are new and developing, or when the type of data collected varies through time. For example, Shotwell and Adkison (2004) developed a run-reconstruction model that integrated weir-based escapement estimates, aerial surveys, commercial catch, commercial-fishery effort, test-fishery catch per unit effort, and sonar counts to estimate indices of total escapement of chum salmon (*Oncorhynchus keta*) in the Kuskokwim and Yukon Rivers. Bue et al. (2012) extended this model for Kuskokwim River Chinook salmon and incorporated periodic total escapement estimates (based on tagging experiments) to scale the model to estimate total run size and total escapement for the entire time series of the Kuskokwim Chinook stock complex.

Tributary-escapement estimates (e.g. weir projects and/or aerial surveys) and periodic estimates of total drainage-wide escapement (e.g. mark-recapture estimates) are the major data sources that stock-reconstruction models rely on for many salmon fisheries, such as Kuskokwim River Chinook and Yukon River Chinook and chum salmon (Bue et al. 2008, Bue et al. 2012, Schaberg et al. 2012). All of these data types may be required for effective salmon management when management objectives include both tributary and drainage-wide escapement goals. However, each data type contributes different information to the drainage-wide run reconstruction and incurs different collection costs. For example, tributary escapement estimates based on weir counts or aerial surveys can provide information on specific sub-stocks, but can also represent indices that describe trends in drainage-wide escapement (e.g. Shotwell and Adkison 2004). The degree to which tributary escapement estimates can index drainage-wide escapement patterns for the entire stock complex will depend in part on the proportion of the total escapement counted by the weirs or aerial surveys, the inherent similarity in temporal patterns in abundance among tributary escapements (i.e. among-stock correlation) and the magnitude of observation error associated with the counts. For example, if temporal patterns in abundance among sub-stocks are perfectly correlated, then counts from one tributary will act as a perfect index of all other tributaries and, thus, the entire stock complex. Alternatively, if temporal patterns among sub-stock demonstrate little to no correlation and/or the observation error is high, a greater number of tributaries will require weir counts or aerial surveys before the cumulative counts will index the drainage-wide escapement effectively.

In contrast to tributary escapement estimates, drainage-wide escapement estimates such as those derived from tagging studies (e.g. Schaberg et al. 2012) provide information about the total escapement on the years that these data are available and provide a scaling factor for tributary escapement estimates when jointly incorporated into a run-reconstruction model. The ability of the drainage-wide abundance estimates to support inference about total run size and escapement will depend on the magnitude of sampling error; however, unlike tributary escapement estimates, drainage-wide abundance estimates are insensitive to the proportion of the total escapement that each tributary represents and insensitive to among-stock correlation. Because of these data complexities, determining the most efficient sampling design for run reconstructions can be unclear. Moreover, the influence of

among-stock correlations in temporal patterns in productivity is important to consider in assessing the model performance in the context of different sampling designs.

In this chapter we investigate the relative value of weir-based escapement estimates, aerial-survey escapement indices and mark-recapture total abundance estimates for informing a run reconstruction model used for management of the Kuskokwim River Chinook salmon (*Oncorhynchus tshawytscha*) fishery in Alaska, USA. Our objectives were to describe the relative influence of each data type on measures of bias and precision of abundance estimates generated from the stock-reconstruction model. Thus, the analysis will assess the performance of the drainage-wide reconstruction model while at the same time providing guidance on efficient sampling designs. We are unaware of any previous studies that have evaluated these sampling design issues. Understanding the interplay of these data types for reconstructing total run size for salmonids will allow management bodies to make better decisions about allocating limited resources to inform the management of these highly valued and often contentious fisheries. This study was motivated by uncertainty over appropriate monitoring of Chinook salmon on the Kuskokwim River, but the rationale and methodological approach can be generalized to any fishery that reconstructs management reference points from combinations of tributary escapement estimates and total run size estimates of semelparous populations.

## METHODS

We used a Monte Carlo simulation frame work to understand the performance of a drainage wide run reconstruction for Kuskokwim River Chinook Salmon and the relative value of weir-based escapement estimates, aerial-survey escapement indices and mark-recapture total escapement estimates to the assesement. The simulation contained three main components: 1) a model describing patterns in abundance of a Chinook salmon multi-stock complex that is parameterized to mimic Kuskokwim River Chinook salmon (termed process model), 2) a model describing the data generating process for weir projects, aerial surveys and mark-recapture total escapement estimation (termed observation model), and 3) a statistical stock-reconstruction model that combines the multiple data types into a unified analysis that estimates the total annual run size and escapement (termed statistical model). The first two components (process and observation models) were used to generate simulated data sets that we then analyzed with the stock-reconstruction model (statistical model). Total escapement estimates were then compared to the true values that they were simulated under to evaluate the bias and precision of annual escapement estimates. We repeated this procedure under different sampling designs and stock characteristics to reveal potential underlying relationships between sampling designs, sub-stock characteristics, and reconstruction model performance.

We parameterized our simulation to mimic the monitoring of Chinook salmon on the Kuskokwim River. Monitoring of the Chinook salmon run on the Kuskokwim River has occurred since 1976. The amount of data and type of data collected has varied through time and has included harvest information, age information, and run size and escapement information. For this study, we were interested mainly in weir-based tributary-escapement estimates, aerial-survey tributary escapement indices and total escapement mark-recapture abundance estimates because these are the data types most likely to be effected by changes in budgetary limitations. Each of these data types have been collected sporadically since 1976, with a general increase in weir-based escapement estimates through time. Over the time period of 2004-2013 approximately 6 tributaries per year have received weirs (range = 0 to 6) and 10 tributaries per year have received aerial surveys (range = 2 to 14). Total escapement mark-recapture abundance estimates have only been generated for the years of 2003-2007 and 2014. All of these data have been used annually to reconstruct the total drainage-wide Chinook salmon run, forecast future run size, and set harvest goals with the reconstruction model developed by Bue et al. (2012).

### Simulation process model

We simulated multiple Chinook salmon sub-stock sampling units (i.e. section of river containing a semi-independent group of returning spawning fish) of a multi-stock complex by first specifying a total number of recruits ( $R_0$ ) at the first time step for the entire complex and apportioning them to each sub-stock sampling unit as  $R_{0,j} = R_0 P_j$ . The parameter  $P_j$  was a vector of proportions (that summed to one) drawn from a Dirichlet distribution. The values of  $R_{0,j}$  represented the carrying capacity of each sub-stock unit. The Dirichlet distribution was parameterized with a concentration parameter ( $\alpha_D$ ) that controlled the variation in  $P_j$  and, thus controlled the variation in the carrying capacity among sub-stock units. We then described the population dynamics of sub-stocks using an age-structured Ricker stock-recruitment model similar in structure to Hamazaki et al. (2012) and Fleischman et al. (2013). Our model described the recruitment of each brood year ( $i$ ) of each sub-stock ( $j$ ) returning to spawn as Ricker stock-recruitment model modified to account for correlation in recruitment variability among sub-stocks through time. The model was specified as:

$$R_{i,j} = \alpha S_{i,j} e^{-\beta_j (S_{i,j}) + \varphi v_{i-1,j} + \varepsilon_{i,j}} \quad (1)$$

where  $S_{i,j}$  is the escapement of spawners at time  $i$  for sub-stock  $j$ . The parameters  $\alpha$  and  $\beta_j$  are the standard Ricker curve parameters describing a dome-shaped relationship between recruitment and spawning-stock numbers. The parameter  $\varepsilon_{i,j}$  represents an error structure that is correlated among sub-stocks to model non-independence in recruitment variation among sub-stocks. The error structure  $\varepsilon_{i,j}$  was modeled as random draws from a multivariate Normal distribution with a mean of zero and a variance-covariance matrix generated by a specified correlation coefficient  $\rho$  and standard deviation  $\sigma_\varepsilon$ . The parameter  $\varphi$  represents the within-sub-stock time series correlation in stock-recruitment residuals (i.e., log scale residual deviation between the realized and expected recruitment) to represent environmental influences that persist from year to year. The recruitment residual  $v_{i-1,j}$  was calculated as:

$$v_{i,j} = \log(R_{i,j}) - \log(\alpha S_{i,j} e^{-\beta_j (S_{i,j})}) \quad (2)$$

where the first term  $\log(R_{i,j})$  represents the observed recruitment and the second term  $\log(\alpha S_{i,j} e^{-\beta_j (S_{i,j})})$  represents the predicted recruitment from the deterministic Ricker model.

Because the rate of maturation among individuals is variable, the number of adults returning to spawn each calendar year is composed of a proportion of fish from different brood years (i.e. ages). We accounted for this variation by predicting the number of returns-at-age as:

$$N_{i,j,k} = P_k R_{i-k,j} \quad (3)$$

where  $k$  indexes the age of returning fish and  $P_k$  represents the expected proportion of each stock returning to spawn at age  $k$ . Thus,  $N_{i,j} = \sum_k N_{i,j,k}$  represents the number of returning spawners of each sub-stock  $j$  at each calendar year  $i$ . Finally, the escapement for each calendar year was predicted as:

$$S_{i,j} = N_{i,j} (1 - U_i) \quad (4)$$

where the parameter  $U_i$  is the finite annual exploitation rate. The input parameters of the process model and definitions are reported in Table 1.

### ***Simulation observation models***

Our primary interest was weir-based escapement estimates, aerial-survey escapement indices and mark-recapture abundance estimates. We simulated tributary escapement estimates for weir-based surveys as random draws from negative binomial distributions with mean equal to  $S_{i,j}$  and over dispersion parameter ( $od_w$ ) equal to 10,000 (approximately Poisson variation) because we expected little bias and variation in true weir counts (Jones et al. 1998, Bue et al. 1998). Simulated weir projects were limited to sub-stock sampling units with an average annual escapement greater than the lowest value observed in the Chinook salmon escapement data for weirs (i.e. ~400). We simulated tributary escapement indices for aerial-based surveys as random draws from negative binomial distributions with a mean and specified over dispersion parameter ( $od_a$ ) to account for additional sampling error. To account for bias in aerial-based counts, we specified the negative binomial mean as  $\lambda_{i,j} = 0.27S_{i,j}$ .

This formula describes systematic negative bias in the salmon counts that is constant across values of  $S_{i,j}$ . We determined this relationship from paired weir and aerial-survey counts of Chinook salmon in the Kuskokwim drainage conducted by Alaska Department of Game and Fish (ADFG unpublished data, Appendix A). We limited the range of sub-stock sampling units available for aerial surveys to those with an average annual escapement between 180 and 18,000. This range approximated a range of average annual aerial-based escapement estimates from 150 to 2,800, which is the range observed in the Chinook salmon escapement data for aerial surveys. Total escapement mark-recapture abundance estimates were specified by drawing random numbers from log-Normal distributions with mean equal to the total escapement (i.e.  $S_i = \sum_j S_{i,j}$ ) for year  $i$  and specified coefficient of variation. Age data

were simulated as random draws from a multinomial distribution with trials equal to the number of fish aged and a specified proportion of each age in the sample.

### ***Assessment model***

We analyzed the simulated datasets with a simplified Bayesian formulation of a stock reconstruction model developed by Bue et al. (2012). The model utilizes multiple data types to estimate the total escapement for the Kuskokwim drainage. For this study, we limited the input data for the model to only include weir-based escapement estimates, aerial-based escapement indices and total escapement mark-recapture abundance estimates. These are the primary data informing the model and the data types of interest for this study. Details on the model and likelihood formulation can be found in Bue et al. (2010). We formulated the model in a Bayesian context to facilitate the natural estimation of parameter uncertainty. Posterior probability distributions of model parameters were estimate using a Monte Carlo-Markov chain algorithm implemented in JAGS (Plummer 2003). We called JAGS from within program R (R Development Core Team 2010) with the library RJAGS (<http://mcmc-jags.sourceforge.net>). Model code and details on the estimation procedure can be found in Appendix B.

### ***Simulation inputs***

Parameter input values were taken from the literature and specified to mimic the observed variation in the available data and capture the relevant biological and sampling characteristics of the stocks such as among-stock correlation, variation in the relative abundance across stocks, and the mean percent of the total escapement counted with one sample unit (i.e. aerial survey or weir project). All parameter inputs, definitions, and sources of their values are reported in Table 1. Some parameter

inputs values were not easily determined from the literature or existing Kuskokwim River Chinook salmon data. These parameters include the number sub-stock sampling units available for escapement counts ( $J$ ), the Dirichlet concentration parameter ( $\alpha_D$ ) controlling the variation in carrying capacity among sub-stocks sampling units, and the correlation in the recruitment among sub-stocks sampling units through time ( $\rho$ ). To determine the most appropriate values of these inputs, we devised a systematic model calibration procedure to determine the combinations of input parameter values that produced simulated data sets of similar character to the observed Kuskokwim Chinook salmon escapement data and were within biologically realistic bounds. Further details on the model calibration procedure is reported in Appendix C.

### ***The simulation***

We simulated data sets from a range of designs that included all combinations of different annual numbers of weir projects and aerial surveys in combination with different periodicities of total escapement mark-recapture abundance estimates. We evaluated a range of 1-10 annual weir projects, 2-20 aerial surveys, and 3 consecutive total escapement mark-recapture abundance estimates with intervals of 3-17 years between them. The sample numbers and periodicities were selected to include a realistic range of designs informed by Alaska Department of Fish and Game personnel. For each design, sub-stocks to be sampled with a weir or aerial survey were selected randomly from those within the defined appropriate abundance range. For each combination of design elements, we simulated 500 data sets, estimated the total escapement (Appendix B), and estimated the mean relative error, absolute relative error, the average precision (CV) of the annual total escapement estimates, and cost of the sampling program.

The mean relative error (MRE) of the escapement estimates was calculated as:

$$\text{MRE} = \frac{\sum_i \frac{(\hat{S}_i - S_i)}{S_i}}{I}, \quad (5)$$

and the absolute relative error (ARE) of the escapement estimates was calculated as:

$$\text{ARE} = \frac{\sum_i \text{abs}\left(\frac{\hat{S}_i - S_i}{S_i}\right)}{I} \quad (5)$$

where  $\hat{S}_i$  is the estimated total escapement at time  $i$ ,  $S_i$  is the true total escapement at time  $i$ , and  $I$  is the number of years of the time series. The average precision of the escapement estimates was calculated as the average coefficient of variation (CV) of the estimates as:

$$\text{CV} = \frac{\sum_i \sigma_{\hat{S}_i} / \hat{S}_i}{I} \quad (6)$$

where  $\sigma_{\hat{S}_i}$  is the standard deviation of the total escapement estimate ( $\hat{S}_i$ ). We approximated costs of each sampling design as the sum of the costs of each data type. The per-unit cost of weirs was approximated as \$150,000 USD, the per-unit cost of aerial surveys was approximated as \$1,500 USD, and the per-unit cost (3-year project) of an total escapement mark-recapture abundance estimate was approximated as \$400,000 USD (Liller per com). For simplicity, we did not include the additional cost of initiating new projects. Thus, these costs assume that the infrastructure necessary to collect the data is in place.

### ***Sensitivity analyses***

For a subset of the input parameters, including those that could not be informed directly from the analysis of available data (Appendix B, this report) or from the literature, we conducted two sensitivity analyses. The first sensitivity analysis was designed to reveal which parameters most influenced the ARE and precision (CV) of the simulation outputs (e.g. Allen and Gwinn 2012). The second sensitivity analysis was designed to reveal the input parameter values that would result in a change in our general conclusion; thus, shedding light on the robustness of our general results. We performed these analyses on the standard deviation describing the random variation in the stock-recruitment relationship ( $\sigma_\varepsilon$ ), the Dirichlet concentration parameter ( $\alpha_D$ ) controlling the variation in carrying capacity among sub-stock sampling units, the correlation in recruitment among sub-stock sampling units ( $\rho$ ), the correlation in the recruitment residual through time ( $\phi$ ), the negative binomial over-dispersion parameter describing sampling error for weir projects ( $od_w$ ), the negative binomial over-dispersion parameter describing sampling error for aerial surveys ( $od_a$ ), and the assumed coefficient of variation of the total escapement mark-recapture abundance estimates ( $MRC_{cv}$ ), describing mark-recapture sampling error.

The first sensitivity analysis was designed to elucidate the relative influence of different parameter inputs on our simulation outputs (i.e. MPE, ARE and CV). We performed the analysis with a fixed reference sampling design of six annual weirs, 10 annual aerial surveys, and three consecutive years of total escapement mark-recapture abundance estimates every two decades to represent a reference sampling design similar to sampling over the terminal decade in the Kuskokwim River data set. First we evaluated the MPE, ARE and CV resulting from the reference design with input parameter values specified in Table 1. We then changed each input parameter by 20% and evaluated any changes in the mean MPE, ARE and CV of the resultant total escapement estimates generated from the assessment model. A total of 500 data sets were generated for each scenario to determine the average influence of each input parameter.

The second sensitivity analysis was designed to reveal the dependency of the value of each data type (i.e. influence on CV per cost) on the input parameters. We performed this analysis by comparing the change in CV relative to the change in cost between three sampling designs and a default design across a range of values of each input parameter. The default design was specified to approximate a minimal design, which consisted of one weir per year, two aerial surveys per year, and a total of three consecutive annual total escapement mark-recapture estimates for the entire 40-year time trajectory. We then specified three alternative designs to compare the change in CV per change in cost of data collection. For the first alternative design, we increased the number of weirs from one per year to 10 per year; for the second, we increased the number of aerial surveys from two per year to 20 per year; and for the third, we increased the number of 3-year total escapement mark-recapture projects from one to four over the 40-year time trajectory. We performed this design comparison for values of  $od_a$  from 1 to 500, for values of  $od_w$  from 1 to 10,000, for values of  $\rho$  from 0.5 to 0.9, for values of  $\phi$  from 0.1 to 1, and for values of  $\alpha_D$  from 0.01 to 100. We chose each range of input parameters such that realistic



values were contained within the range. For example, the correlation among sub-stock sampling units ( $\rho$ ) is 0.52 for the observed escapement indices indicating that the true correlation must be greater because of error in the observation process. Alternatively, correlations higher than 0.7 are rarely reported in the literature. Thus, the range of 0.5 to 0.9 likely spans the true correlation of the sub-stock units. The range of  $\alpha_D$  was specified to include conditions where nearly all of the total carrying capacity of the stock complex was represented by a few sub-stock sampling units (i.e. 0.01) to unrealistically similar carrying capacity across stocks (i.e. 100). The range of  $\varphi$  represents a correlation in the recruitment residuals from year to year of nil to perfect correlation. We evaluated a range of the standard deviation of the recruitment residual  $\sigma_\varepsilon$  from 0.1 to 1.5. When the structural assumption is made that the sub-stocks sampling units behave as a single stock, the estimated value of  $\sigma_\varepsilon$  is  $\sim 0.28$  (Appendix B, this report). This represents a minimum expected value because estimating  $\sigma_\varepsilon$  under the false assumption of perfectly correlated population dynamics will serve to underestimate the parameter. Alternatively, a value higher than 1.5 is unexpected based on observations of stock recruitment relationships of Pacific Salmonids (Myers et al 1999, Goodwin et al 2006).

## RESULTS

### **General simulation results**

Our simulations indicated that bias (MRE and ARE) and precision of escapement estimates were related to the amount of each data type in the sampling design. The MRE and ARE decreased and CV increased with increases in mark-recapture years, weir projects, or aerial surveys (Figure 1). The average parameter bias (MRE) was less than 0.06 (6%) for all scenarios and was 0.01 (positive or negative) so long as at least 10 air surveys and 6 weir projects were operating, and at least two sets of three-year mark recapture total escapement estimates were done over the forty-year time horizon. Similarly, the absolute relative error did not exceed 0.05 (5%) under these same scenarios. The average CV of escapement estimates was less than 0.1 (10%) across all scenarios and was less than 0.05 so long as at least 10 air surveys and/or 6 weir projects were operating and at least two sets of three-year mark recapture total abundance estimates were done over the forty-year time horizon. Similar CVs could be achieved by increasing the number of air surveys and decreasing the number of weirs, or vice versa. For example, the average CV was 0.05 under 16 air surveys and one weir or one air survey and six weirs, if two sets of three-year mark recapture escapement estimates were conducted.

The impact of each data type on the MPE, ARE, and coefficient of variation (CV) of escapement estimates demonstrated a saturating effect, where the change in these quantities declined as the number of sample units of each data type increased. This is most easily observed on Figure 1 where the isopleths become further apart as the number of weir projects or aerial surveys increase. Although harder to observe on Figure 1, this pattern was also true for mark-recapture (MRC) projects. This saturation effect suggests that there is diminishing returns in terms of reducing parameter uncertainty and bias as more and more escapement indices are added.

Figure 1 also indicates that there was no optimal ratio of weir to aerial survey projects where the bias and precision per total sampling units would be optimized. This is indicated by the nearly straight and parallel isopleths on the contour plots of Figure 1. Similarly, the parallel nature of isopleths of Figure 1 also indicates that the relative value of weir projects and aerial surveys is constant across sampling designs. For example, a similar change in the CV (approximately 0.042 to 0.039) can be achieved by adding two weir project per year, five aerial surveys per year, or one 3-year MRC project to the 40-year time series. However, from the perspective of cost, these design alterations were drastically different. Increasing the sampling design by an additional 3-year mark-recapture experiment, two additional weir project per year, or five additional aerial survey per year increased the 40-year cost of

the data collection by \$1.2 million, \$12 million, or \$0.3 million, respectively. Thus, on a per-cost basis, aerial surveys had the greatest influence on the bias and precision of the abundance estimates and weir projects had the least influence within the range of designs explored.

### ***Sensitivity analysis results***

Our first sensitivity analysis was designed to investigate the relative influence of our input parameter values on the ARE and CV outputs. This analysis demonstrated that there was more variation around the ARE than CV outputs, indicating that CV was more sensitive metric. Otherwise, we found that the relative influence of the input parameters on the CV was similar in pattern to their influence on the ARE. We found that the ARE and CV were most sensitive to the input values of  $\rho$ ,  $\sigma_\varepsilon$ , and  $MRC_{cv}$  (Figure 2). We observed a small influence of  $od_a$  and no observable influence of  $\phi$ ,  $\alpha_D$ , and  $od_w$  (Figure 2).

Our second sensitivity analysis was designed to investigate the robustness of our conclusions about the relative influence-per-cost of each data type to input parameter values. The results of this analysis corroborated fairly well with the results of the first sensitivity analysis. This analysis clearly indicated that the relative value of each data type did not depend on our assumption of the values of  $\phi$  and  $\alpha_D$  (Figure 3). This was also true for  $od_w$  values between ~200-10,000; however, below values of 200, the reduction in the CV due to changes in the number of aerial surveys became stronger (Figure 3). The value of the input parameters,  $\rho$ ,  $\sigma_\varepsilon$ , and  $MRC_{cv}$  had a clear influence on the relative value of the data types. When the correlation among sub-stock sampling units ( $\rho$ ) was increased the relative value of both weirs and aerial surveys decreased (Figure 3). This was an especially strong response for aerial surveys. When we increased the random variation in recruitment,  $\sigma_\varepsilon$ , the opposite response occurred, i.e., the relative value of weirs and aerial surveys increased. Increasing the assumed precision of the total-escapement MRC abundance estimates (i.e. a decrease in  $MRC_{cv}$ ) decreased the relative value of aerial surveys and increased the relative value of weirs (Figure 3). Although, the assumed values of some of the input parameters influenced the relative values of each data type, they were never influenced to a degree that would change the general conclusion that aerial surveys are the most cost-effective sampling method. This was indicated in Figure 3 by the red line, representing the change in CV per change in cost for aerial surveys, never crossing either the black or green lines. Thus the aerial surveys were the most cost effective data type across all parameter input values evaluated, indicating this as a robust result.

## **DISCUSSION**

Here we demonstrate the high relative value of aerial surveys for informing the run abundance reconstruction model for Chinook salmon on the Kuskokwim River. We found that, per cost, aerial surveys outperformed both weir projects and total-escapement MRC abundance estimates for increasing the precision of total escapement estimates. We also found that weir projects demonstrated the lowest information value per cost. These are important findings because they can help managers decide how to allocate funds to achieve the best information within fiscal limitations. Furthermore, our results can help to guide the addition or subtraction of data collection efforts as budgets fluctuate, ensuring the best use of limited funding.

Our general findings are in contrast to commonly held perceptions about the relative data quality of aerial surveys and weir counts. The efficacy of aerial surveys is typically thought to be low, because the technique is affected by many factors, such as observer experience, water clarity, water depth, surface wind, cloud cover, angle and intensity of sun (Chapman 1986, Swan 1989, Bue et al. 1998). Aerial surveys are also known to be systematically biased low because of an imperfect ability to

observe all fish present (Jones 1998, Bue et al. 1998). Alternatively, weir counts, which are less sensitive to environmental variables (Bue et al. 1998), are typically considered a more valuable data type for informing management.

Our results suggest that this general perception may not hold for the specific case of the Kuskokwim Chinook salmon abundance reconstruction model. Our finding is primarily a consequence of the large difference in monetary cost of aerial surveys compared to alternative data types. For example, the cost of a weir project per year is approximately 100 times the cost of an aerial survey; thus, the influence of the weir counts on the CV would need to be approximately 100 times the influence of the aerial surveys on the CV for the data types to be approximately equivalent. By contrast, we found that approximately one weir project demonstrates only twice the influence on the CV than one aerial survey on average. The relatively high precision of weirs at estimating escapement abundance may have been lost due to how they are used in the model. In this case, the weirs were used in a similar manner as the air surveys in the model, that is, as an index of escapement abundance rather than a measure of the absolute magnitude of escapement. The weirs are used as an index because of the large amount of uncounted escapement in the drainage. The scaling of escapement in the model comes from the periodic drainage-wide escapement estimates via mark-recapture studies. Thus a given weir counts some unknown fraction of the total escapement and therefore no matter how precise, the estimate will only inform relative changes in escapement over time rather than the absolute magnitude of escapement.

Our simulation revealed that the aforementioned result was robust to the value of input parameters despite model sensitivity to some input parameters. For example, we found that model outputs were most sensitive to the assumed level of random variation in recruitment ( $\sigma_\epsilon$ ) and correlation in recruitment among sub-stocks ( $\rho$ ). These findings are consistent with our *a priori* expectation that the value of additional tributary escapement estimates would decline as the correlation among stocks increased and the random variation in the recruitment process decreased. This occurs mainly because the ability of one tributary to index the total escapement increases when the correlation is high and noise is low; thus, sampling additional tributaries becomes less informative. Additionally, the assessment model assumes that the sub-stocks conform to a single population dynamics model (in other words, perfect correlation among sub-stocks). Thus, generating data from highly correlated sub-stocks with little variation would better meet to this model assumption, likely increasing both the accuracy and precision of the model estimates. Alternatively, that the assumed CV of the drainage-wide MRC escapement estimates influenced the value of tributary escapement estimates was unexpected. However, this outcome also makes sense given the role of this information as a scaler in the assessment model, i.e. the more accurate and precise the information on the drainage-wide escapement the more accurate and confident the reconstruction.

Models in ecology are not intended to be exact replicas of reality, but simplifications that allow us to understand and predict complex processes. While model complexity can be limitless, the most useful models are those that incorporate only enough complexity to meet research and management objectives. In this study, we modelled only the aspects of population dynamics and sampling process important for understanding the relative value of different data types on estimates of total salmon escapement; thus, our results have several limitations worth discussing. First, our simulation assumes that error in the tributary escapement estimates and MRC estimates is random and does not vary in any systematic way through time. However, if observer efficiency systematically changes through time because of changing weather patterns, river characteristics, or salmon abundance, we would expect a systematic trend in bias in the aerial-survey based escapement estimates. This would introduce bias into the assessment model and reduce the relative value of the aerial-survey information. We would expect an analogous effect for weir projects and MRC projects. Second, our analysis did not consider the value

of the data types for other uses. For example, we did not consider the value of each data type for determining if tributary-based escapement goals are met. Clearly, the more precise weirs would serve this purpose better than air surveys. Nor did we consider the value of the information derived by handling fish during weir projects such as age and length information, as recapture points for mark-recapture studies. Managers will need to carefully consider these advantages of the weirs when making decisions on allocation of resources for surveys. Because of these additional advantages that we did not consider, weirs will likely continue to be an important type of survey to inform management of Kuskokwim Chinook and other species. A third important caveat is that several important population dynamics parameters of the operating model were not well informed by data. For example, we did not have good estimates of  $\rho$  (among stock correlation in recruitment variation over time). Thus we used an intensive model tuning exercise to evaluate which values of  $\rho$  could have plausibly produced the existing time series of data from the Kuskokwim Chinook salmon stock. We believe that our tuning exercise was a reasonable approach in the absence of better information, but we also acknowledge that future analyses would be improved if better estimates of these parameters could be obtained. The current AYKSSI-funded project looking at harvest diversity (Connors, Coggins, Walters, and Fleischman) may in fact produce estimate of at least some of these important population dynamics parameters, particularly  $\rho$ .

Another important assumption of the model is that each tributary contributes the same proportion of the total escapement, on average, over time. This assumption would be violated if some tributary stocks vary systematically over time due to natural causes thereby causing a shift in their contribution to the total escapement. This type of assumption violation was not evaluated in our analysis and should receive attention in future analyses. This assumption may also be violated if the relative harvest rates across tributary stocks changes over time. For example, a reduction in harvest rates on early running up-river stocks may increase the relative contributions of these stocks to the total escapement. This scenario may be playing out in recent years as severe subsistence harvest restrictions have apparently increased escapement to up-river tributaries. Thus, although our analysis suggests that parameter bias is low, we caution that these findings may not hold up as we further challenge the model with plausible additional assumption violations.

The sampling program for Kuskokwim Chinook has generally resulted in an average of 10 aerial surveys and 6 weirs being sampled annually, and three to six mark recapture estimates of total drainage-wide escapement per 40 year time series. Our analysis suggests that adding more of these sampling projects is unlikely to appreciably increase parameter precision and decrease bias. However, our findings do suggest that similar levels of bias and precision could be achieved at less cost by increasing the number of air surveys while decreasing the number of weirs. However, doing so would come at the cost of loss of information on age, sex, and length data, not to mention loss of recapture sites for tagging studies. We caution that our analysis did not quantify the value of these additional weir-based data streams on the assessment and management activities for this stock.

## REFERENCES

- Allen, M.S. and D. Gwinn. 2013. Population models for assessment and management of inland striped bass fisheries. Pages 351-364 in J.S. Bulak, C.C. Coutant, and J.A. Rice, editors. Biology and management of inland striped bass and hybrid striped bass. American Fisheries Society, Symposium 80, Bethesda, Maryland.
- Bolker B.M. 2008. Ecological Models and Data. Princeton University Press: Princeton, NJ.
- Bue, B.G., S.M. Fried, S. Sharr, O.G. Sharp, J.A. Wilcock, and H.J. Geiger. 1998. Estimating salmon escapement using area-under-the-curve, aerial observer efficiency, and stream-life estimates: The Prince William Sound pink salmon example. N. Pac. Anadr. Fish Comm. Bull. No.1: 240-250
- Bue, B.G., K.L. Schaberg, Z.W. Liller, and D.B. Molyneaux. 2012. Estimates of the historic run and escapement for the Chinook salmon stock returning to the Kuskokwim River, 1976-2011. Alaska Department of Fish and Game, Fishery Data Series No. 12-49, Anchorage.
- Bue, B.G., D.B. Molyneaux, and K. L. Schaberg. 2008. Kuskokwim River chum salmon run reconstruction. Alaska Department of Fish and Game, Fishery Data Series No. 08-64, Anchorage.
- Chapman, D.W., D.E. Weitkamp, T.L. Welsh, and T.H. Schadt (1986) Effects of river flow on the distribution of chinook salmon redds. Transactions of the American Fisheries Society, 115, 537-547.
- Fleischman, S.J., M.J. Catalano, R.A. Clark, and D.R. Bernard. 2013. An age-structured state-space stock-recruit model for Pacific salmon (*Oncorhynchus* spp.). Canadian Journal of Fisheries and Aquatic Sciences. 70: 401-414.
- Hamazaki T., M. J. Evenson, S. J. Fleischman, and K. L. Schaberg. 2012. Spawner-recruit analysis and escapement goal recommendation for Chinook salmon in the Kuskokwim River Drainage. Alaska Department of Fish and Game, Fishery Manuscript Series No. 12-08, Anchorage.
- Shea, K., P. Amarasekare, M. Mangel, J. Moore, W.W. Murdoch, E. Noonburg, A. Parma, M.A. Pascual, H.P. Possingham, C. Wilcox, and D. Yu. 1998. Management of populations in conservation, harvesting, and control. Trends in Ecology and Evolution 13: 371-375.
- Shotwell, S.K. and M.D. Adkison. 2004. Estimating indices of abundance and escapement of Pacific salmon for data-limited situations. Transactions of the American Fisheries Society 133:538-558.
- Maunder, M.N. and R.B. Deriso. 2003. Estimation of recruitment in catch-at-age models. Canadian Journal of Fisheries and Aquatic Science 60: 1204-1216.
- Mertz, G. and R.A. Myers. 1996. Influence of fecundity on recruitment variability of marine fish. Canadian Journal of Fisheries and Aquatic Science 53: 1618-1625.
- Peterman, R.M. and B. Dorner. 2012. A widespread decrease in productivity of sockeye salmon (*Oncorhynchus nerka*) populations in the Western North America. Canadian Journal of Fisheries and Aquatic Science, 69, 1255-1260.
- Plummer, M. (2003) JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing* (eds F. Leisch & A. Zeileis). Vienna, Austria.
- R Development Core Team 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org/>.
- Swan, G.A. (1998) Chinook salmon spawning surveys in deep waters of a large, regulated river. Regulated Rivers: Research & Management, 4, 355-370.
- Schaberg, K.L., Z.W. Liller, D.B. Molyneaux, B.G. Bue, and L. Stuby. 2012. Estimates of the total annual return of Chinook salmon to the Kuskokwim River, 2002-2007. Alaska Department of Fish and Game, Fishery Data Series No. 12-36, Anchorage.

Table 1. Parameter input values, definitions and sources for the simulation model.

Symbol	Value	Definition
<b><u>Process Model</u></b>		
$R_0$	210,000 <sup>2</sup>	The total recruitment across sub-stocks at time zero
$J$	70 <sup>1</sup>	The total number of sub-stock units, i.e. tributaries or tributary segments that could potentially be sampled with a weir or aerial survey.
$\alpha$	5.49 <sup>2</sup>	parameter of Ricker stock-recruit model
$\beta_j$	$\log(\alpha)/R_{0,j}$	parameter of Ricker stock-recruit model
$k$	4, 5, 6, 7	The ages of fish returning to spawn
$P_k$	0.20, 0.39, 0.38, 0.03 <sup>4</sup>	The expected proportion of fish of different ages returning to spawn ( $k$ )
$\rho$	0.80 <sup>1,5,6</sup>	Expected correlation among sub-stock recruitment through time
$\phi$	0.85 <sup>2</sup>	Expected correlation in recruitment among consecutive time steps for each sub-stock
$\sigma_\varepsilon$	0.60 <sup>8,9</sup>	Process error around the stock-recruitment model
$\alpha_D$	0.25 <sup>1</sup>	The Dirichlet concentration parameter that controls how $R_0$ is apportioned among sub-stocks
$U$	0.4 <sup>2</sup>	Finite annual exploitation rate of fishery
<b><u>Aerial Survey Model</u></b>		
$od_a$	40 <sup>2,3,4</sup>	Negative binomial over-dispersion parameter of aerial surveys
$od_w$	10,000 <sup>3,4</sup>	Negative binomial over-dispersion parameter of weir counts (set to mimic Poisson variation)
<b><u>In-River Mark-Recapture Model</u></b>		
$MRC_{cv}$	15% <sup>7</sup>	Coefficient of variation of the mark-recapture abundance estimates

<sup>1</sup> Appendix C

<sup>2</sup> Chapter 2 (this report)

<sup>3</sup> Jones et al. (1998)

<sup>4</sup> Bue et al. (1998)

<sup>5</sup> Mueter et al. (2002)

<sup>6</sup> Peterman and Dorner (2012)

<sup>7</sup> Bue et al. (2000)

<sup>8</sup> Mertz and Myers (1996)

<sup>9</sup> Maunder and Deriso (2003)

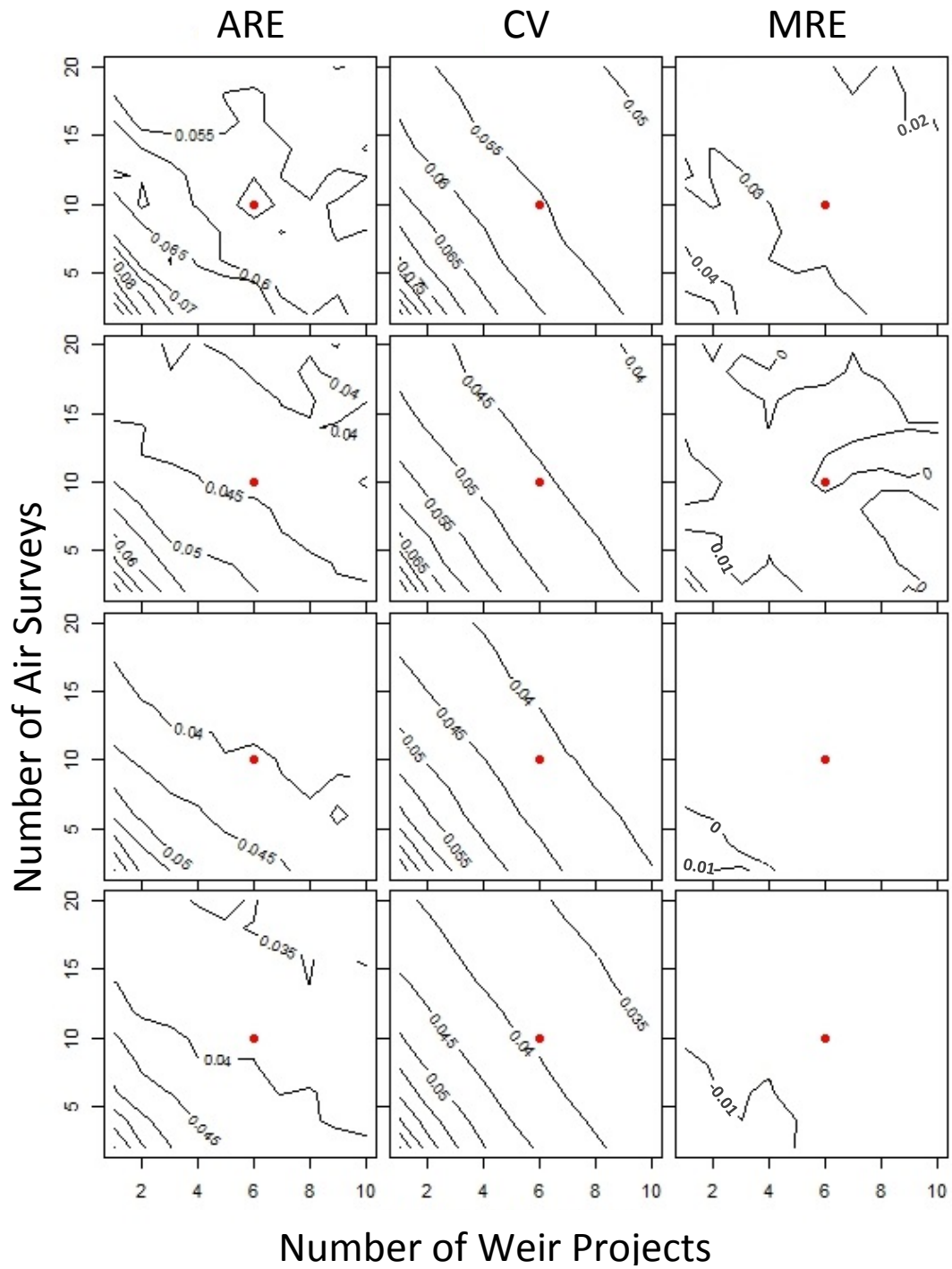


Figure 1. The mean absolute relative error (ARE), coefficient of variation (CV), and mean relative error (MRE) of total-escapement estimates as a function of the number of weirs (x axis) aerial surveys (y axis) and three-year mark-recapture drainage-wide escapement projects (per 40 year time horizon). The red dots indicate recent sampling aerial survey and wier sampling intensity.

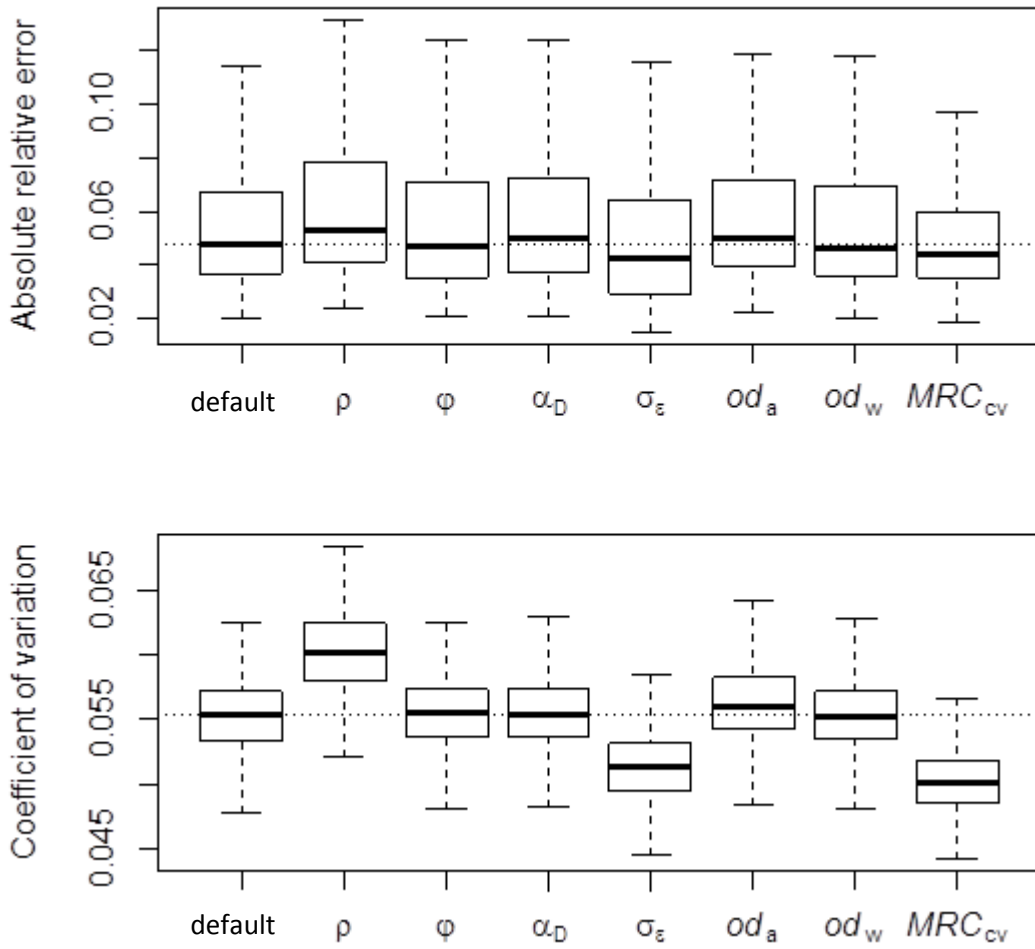


Figure 2. Mean absolute relative error (ARE; top panel) and coefficient of variation (CV; bottom panel) of annual escapement estimates with default model input parameter values ('default') and after a 20% decrease in model input parameters.



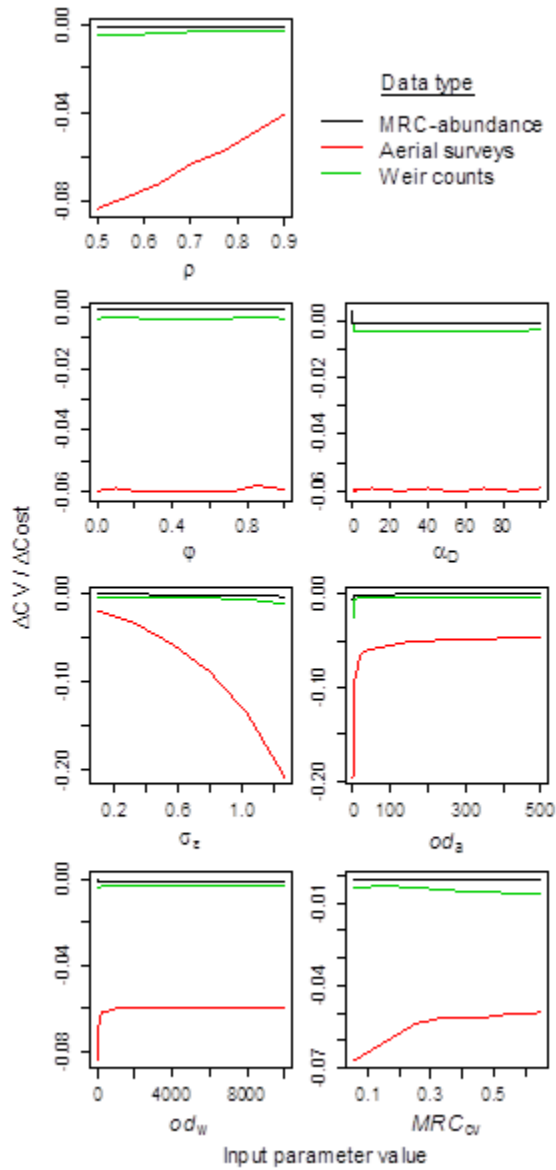


Figure 3. The change in the average coefficient of variation of annual escapement estimates per cost ( $\Delta CV / \Delta Cost$ ) when increasing a sampling design from 2 to 20 annual aerial surveys (red line), 1 to 10 annual weir projects (green line) and 1 to 4 3-year mark-recapture projects (black line) with a range of model parameter input values. More negative values on the y-axis represent greater reductions in the CV per cost and thus a greater value.

## Appendix A. Methods for determining systematic bias in Chinook salmon aerial survey counts.

Research conducted on Pacific salmon species, such as pink salmon (*Oncorhynchus gorbuscha*), indicate that aerial surveys can be negatively biased and the bias can be related to fish abundance (e.g. Jones et al. 1998); however, we are unaware of any published research determining the expected bias and variation of Chinook salmon counts conducted from aircraft. The purpose of the following analysis was to determine the expected bias and describe its relationship to the true abundance of Chinook salmon in the Kuskokwim River drainage.

### Data source

To determine the expected bias of aerial survey counts and its relationship to true abundance, we utilize available salmon counts from weirs and aerial surveys that were conducted on the same tributaries. We obtained weir-count and aerial-count data from the ADFG data portal (website). A total of 34 paired weir/aerial counts from five different tributaries were available (Table A1).

### Analysis

We assumed that weir-based salmon counts closely represented the true tributary escapement. Thus we evaluated the bias of aerial surveys relative to the weir counts. We calculated the bias of each aerial survey as,  $Bias = (C_{aerial} - C_{weir}) / C_{weir}$ , where  $C$  represents the count derived from either weir or aerial surveys. We assumed that the Bias across surveys could be explained by a normal distribution where the mean was modeled as:

$$\mu = \omega(1 - e^{-\psi C_{weir}})$$

Where  $\omega$  represents the asymptotic proportional bias and  $\psi$  models its relationship to  $C_{weir}$ . We described the relationship between the bias of aerial counts and abundance

Table A1. List of paired weir and aerial counts for five tributaries of the Kuskokwim River.

Table 1. Annual weir and aerial survey counts. Aerial surveys only include those with rated Fair or better.															
Year	Kwethluk			Tuluksak			Salmon (A)			Kogrukuk			Holitna		
	Weir	Aerial Survey	AS%	Weir	Aerial Survey	AS%	Weir	Aerial Survey	AS%	Weir	Aerial Survey	AS%	Weir (Kogrukuk)	Aerial Survey	AS%
2014	3,187			320			1,757	478	27%	3,732			3,732	3,086	83%
2013	650	1,165*	179%	193	83	43%	625	154	25%	1,819			1,819	670	37%
2012				555				49							
2011	4,079			284				79		6,731			6,731		
2010	1,668			201						5,812			5,812		
2009	5,744			362			1,656			9,528	1,649	17%	9,528		
2008	5,276	487	9%	701			2,376	589	25%	9,750	1,641	17%	9,750	468	5%
2007	12,927			374	173	46%	6,255	1,458	23%		1,879				
2006	17,619			1,043			7,075			20,205	4,159	21%	20,205	1,866	9%
2005		5,059		2,653	672	25%	4,097			21,819	4,954	23%	21,819	1,760	8%
2004	28,605	6,801	24%	1,475	1,196	81%	217			19,812	7,252	37%	19,812	4,051	20%
2003	14,474	2,661	18%	1,064	94	9%	1,242			12,008	3,058	25%	12,008		
2002	8,502	1,795	21%	1,346			1,236			10,025	1,204	12%	10,025	1,578	16%
2001				997			598			7,475			7,475	1,130	15%
2000	3,547						238			3,242			3,242	301	9%
1999										5,543			5,543	741	13%
1998							557								
1997							980			13,190			13,190	2,093	16%
1996							985			13,771			13,771		
1995							1,442			20,662			20,662	1,887	9%
1994				2,918											
1993				2,218						12,377			12,377	1,573	13%
1992	9,675			1,083						6,563			6,563	2,022	31%
1991				697						6,835			6,835		
1990										10,093			10,093		
1989															
1988										8,028	1,987	25%	8,028		
1987										715					
1986														650	
1985										4,442			4,442		
1984										4,922			4,922		
1983														1,069	
1982										13,126			13,126	521	4%
1981										16,089			16,089		
1980															
1979										11,393			11,393		
1978										14,533			14,533	2,766	19%
1977											1,344				
1976										5,638	702	12%	5,638	2,571	46%
n			5			5			4			9			17

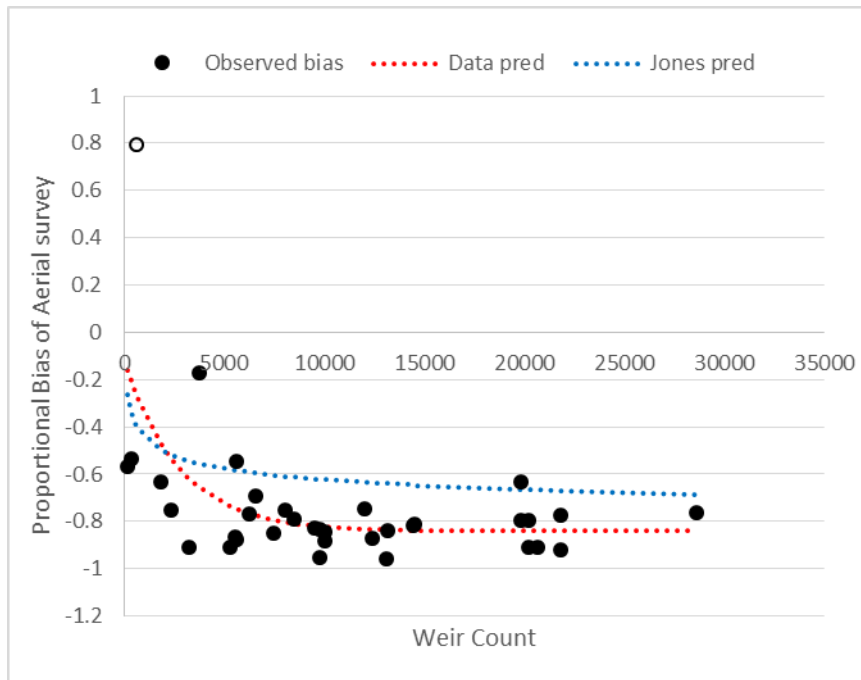


Figure A1. Proportional bias of aerial survey Chinook salmon counts as a function of weir counts from paired aerial survey-weir counted tributaries of the Kuskokwim River.

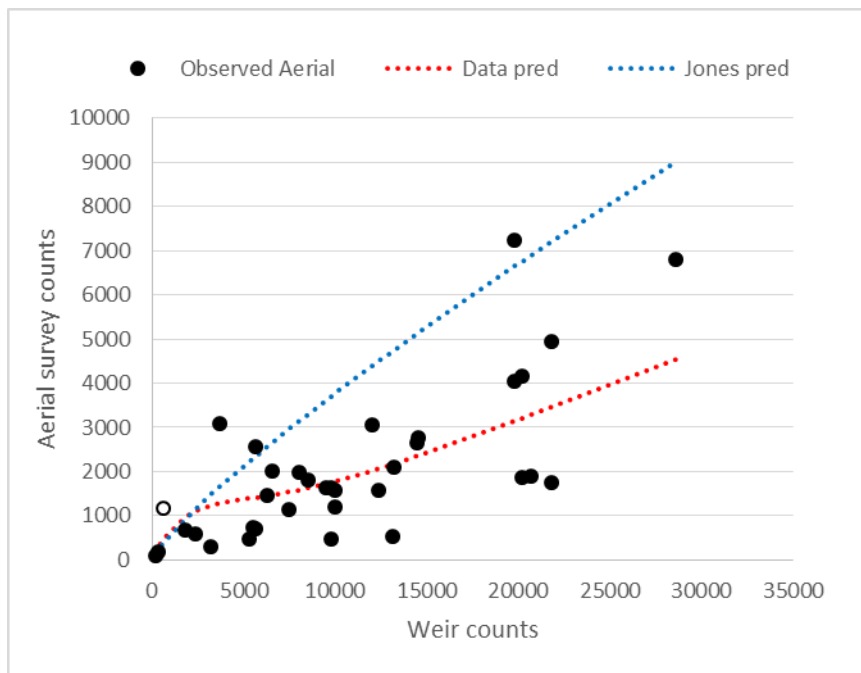


Figure A2. Aerial survey counts as a function of weir counts.

## Appendix B. Estimation model and fitting procedure.

Posterior probability distributions of model parameters were estimate using a Monte Carlo-Markov chain algorithm implemented in JAGS (Plummer 2003). We called JAGS from within program R (R Development Core Team 2010) with the library RJAGS (<http://mcmc-jags.sourceforge.net>). All prior distributions were standard uninformative distributions specified to have no influence on the posterior probability distributions. Priors for log-scale parameters (i.e., returning spawners N and tributary expansion parameters k) were set as flat normal distributions with mean equal to zero and standard deviation equal to one hundred. Priors of inverse standard deviation parameters were specified as flat gamma distributions with shape parameters set to 0.01 and were verified to not influence the range of the posterior distributions. Inference was drawn from 4,000 posterior samples taken from 2 chains of 20,000 samples,each, thinned to every 10. We allowed a burn in of 20,000 samples to remove the effects of initial values. Convergence was diagnosed by using the Gelman-Rubin statistic (with values < 1.1 indicating convergence; Kéry 2010, Gelmin *et al.* 2004).

Gelman, A., Carlin, J.B., Sterns, H.S. & Rubin, D.B. (2004) *Bayesian Data Analysis*. Chapman and Hall, Boca Raton.

Kery, M. (2010) *Introduction to WinBUGS for Ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses*. Academic Press.

Plummer, M. (2003) JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing* (eds F. Leisch & A. Zeileis). Vienna, Austria.

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

### JAGS model code for simplified version of stock reconstruction model developed by Bue et al. (2012)

```
model{
##set priors; looped priors are in
#other loops for efficiency##
  mushape ~ dnorm(0,.01)|(-2,6)
  taushape ~ dgamma(.01,.01)
  for(i in 1:ns){
    logshape[i] ~ dnorm(mushape,taushape)
    shape.esc[i] <- exp(logshape[i])
  }

#estimating total escapement
  for(y in 1:ny){
    log.N[y]~dnorm(0, 0.01)
    N[y]<-exp(log.N[y])
    S[y]<-N[y]-catch.tot[y]
    U[y]<-catch.tot[y]/N[y]
  }

#estimated escapement for each trib
  mulogk~dnorm(0,0.01)
```

```

taulogk~dgamma(.01,.01)
for(j in 1:ns){
  log.k[j]~dnorm(mulogk, taulogk)
  for(y in 1:ny){
    k[y,j]<-exp(log.k[j])
    est.esc[y,j]<-(S[y]/k[y,j])
    rate.esc[y,j]<-shape.esc[j]/est.esc[y,j]
    lambda[y,j]~dgamma(shape.esc[j], rate.esc[y,j])
    esc[y,j]~dpois(lambda[y,j])
    qweir[y,j] <- lambda[y,j]/N[y]
  }
}
#in river likelihood
for(i in 1:nmrc){inr.n[mrcvec[i]]~dlnorm(S[mrcvec[i]], tau.inr[mrcvec[i]])}
}

```

## Appendix C. Parameter input values and model calibration.

### Model calibration methods

Because the exact values for three of the parameter inputs could not easily be determined from the literature or available data, we devised a systematic model calibration procedure to determine the combinations of values that were both biologically realistic and produced simulated data sets of similar character to the observed Kuskokwim Chinook salmon escapement data. The procedure involved simulating 500 data sets from each of all possible combinations of a biologically realistic range of the three inputs parameters, calculating three metrics describing the simulated data, comparing these metrics of the simulated data to the metrics of the observed Kuskokwim Chinook salmon escapement indices. The three parameters evaluated were the number sub-stock sampling units available for escapement counts ( $J$ ), the Dirichlet concentration parameter ( $\alpha_D$ ) controlling the variation in carrying capacity among sub-stocks, and the correlation in the recruitment among sub-stocks through time ( $\rho$ ). We simulated data from a range of  $\rho$  from 0.65 to 0.80. The lower bound of this range represents the average among-stock correlation observed in the data for weirs and, thus, the simulated correlation input must be at least this value. The upper bound of  $\rho$  represents a value slightly higher than is reported in the literature for Pacific salmon (~0.7, Mueter et al. 2002). There was no available guidance on  $\alpha_D$  or  $J$ , so we simply simulated data from an arbitrary range of values. The input range for  $\alpha_D$  was 0.25 – 0.75 and for  $J$  was 70 – 130.

We attempted to mimic three observable metrics for weir-based and areal-based Chinook salmon escapement data with our simulated data. These metrics included the mean annual coefficient of variation among tributary escapement indices (CV), the mean correlation among tributary escapement indices through time (Cor), and the mean percent of the total escapement sampled with a single weir or aerial survey (i.e., coverage, Cov). These characteristics of the escapement data are key drivers influencing the relative value of an escapement index for informing total escapement. These metrics were calculated for all simulated data sets and for the observed data. We then sought the parameter input values that minimized the proportional deviation (PD) of the metrics of the simulated data from the metrics of the observed data as

$$PD_{i,g} = (x_{sim,i,g} - x_{obs,i,g}) / x_{obs,i,g}$$

where  $PD_i$  is the proportional deviation of metric  $i$  (i.e. CV, Cor, Cov) and gear  $g$  (i.e. weirs projects or aerial surveys),  $x_{sim,i,g}$  represents the value of metric  $i$  calculated from the simulated data and  $x_{obs,i,g}$  is the value of metric  $i$  calculated from the observed data. Note that when  $x_{sim,i,g} = x_{obs,i,g}$ , the  $PD_{i,g}$  approaches zero, indicating a perfect match of the simulated data to the observed data.

The optimal parameter input values are those that both minimize the  $\sum PD_{i,g}$  and the  $CV(PD_{i,g})$ . Thus, we ranked each parameter set by an objective function defined as the sum of these two calculations normalized across parameter sets, i.e.  $Obj = \sum PD_{i,g}^* + CV(PD_{i,g})^*$ . The \* indicates that the term is normalized, i.e. centered on zero and scaled to one standard deviation. This ranking assumes that minimizing the sum of the normalized proportional deviations of the metrics is equally important as minimizing the normalized variation of the proportional deviations among the three metrics.

### Model calibration results

We identified the top four parameter sets using our model calibration procedure. All parameter values differed among these input sets except for  $\alpha_D$ , which was a value of 0.25 for all four sets (Table A1). Parameter input Set-1 distinctly stood out as the best parameter value set with the lowest  $Obj$ ,  $\sum PD_i$ , and  $CV(PD_i)$  values (Table C1). The average data metrics (i.e. CV, Cor, Cov) produced by parameter Set-1 closely predicted the metrics calculated from the observed data (Figure C1). Observed metric values were always with in the 50 percentile of the simulated variation (Figure C1); however, differences among parameter sets were subtle. The model calibration procedure produced clear evidence that Set-1 replicated the observed data the best, and thus, we chose this set of inputs for this study.

Table C1. The four top parameter sets and the resultant average mean squared error (MSE) and coefficient of variation (CV) across parameters.

Parameter set	J	$\alpha_D$	$\rho$	$\sum PD_{i,g}$	$CV(PD_{i,g})$	$Obj$
Set - 1	70	0.283	0.75	0.55	0.34	-3.84
Set - 2	90	0.283	0.75	0.66	0.35	-3.36
Set - 3	70	0.400	0.70	0.53	0.49	-3.02
Set - 4	100	0.283	0.75	0.79	0.36	-2.85



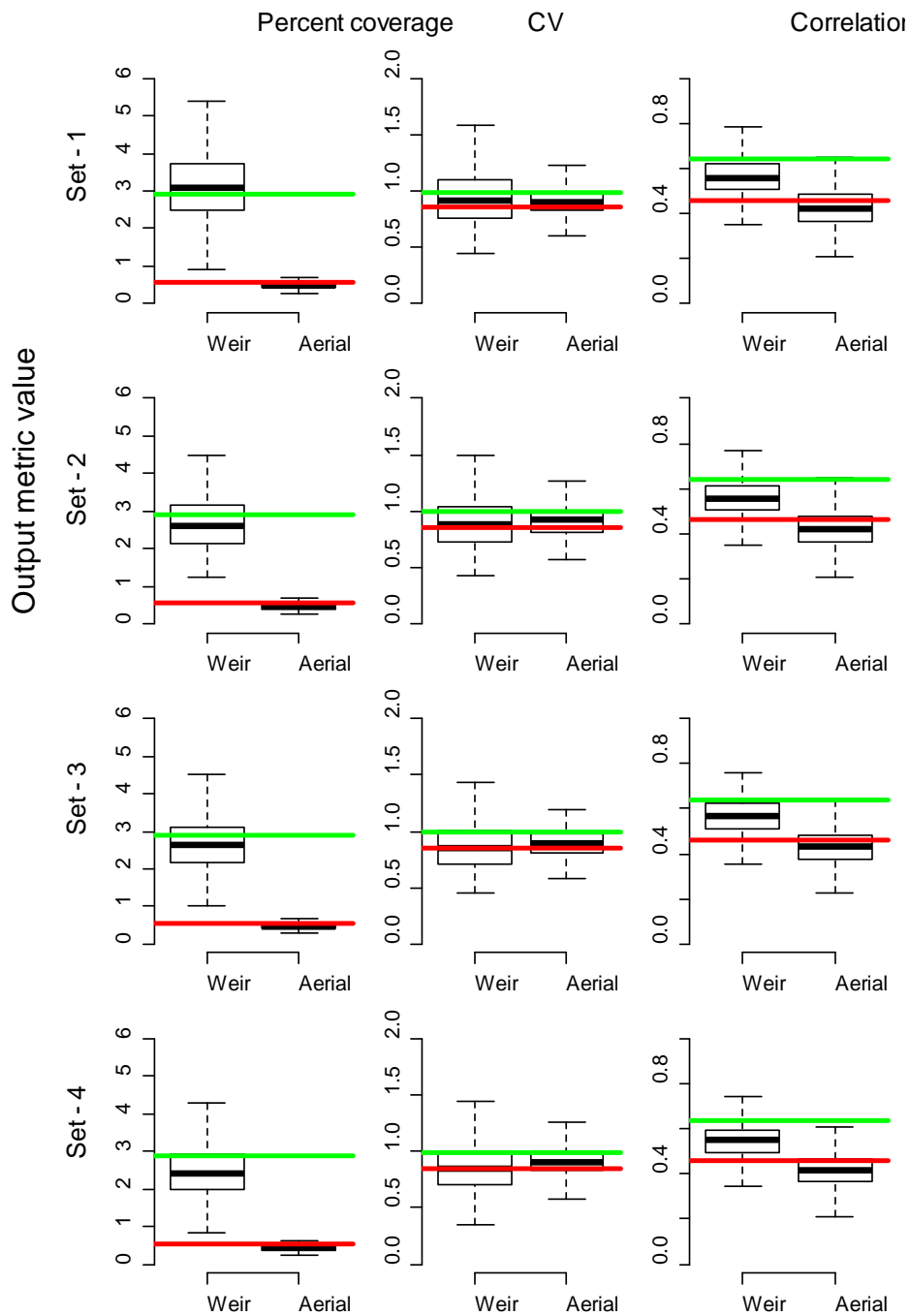


Figure C1. Box and whisker plots of the simulated values of three metrics describing simulated data sets (% coverage, Cov; coefficient of variation, CV; and correlation, Cor) for wier and aerial survey indices. Set-1, 2, 3, and 4 represent the top four parameter input sets evaluated. The green and red lines depict the observed values for weirs and aerial sveys, respectively, from 1976-2013 for Kuskokwim Chinook salmon .

## CHAPTER 2

### EVALUATION OF AN INTEGRATED RUN RECONSTRUCTION AND SPAWNER-RECRUIT MODEL FOR KUSKOKWIM RIVER CHINOOK SALMON

#### ABSTRACT

Stock assessment scientists are often faced with decisions regarding how to incorporate information into models. One primary decision revolves around how estimates that are summaries of raw data should be treated (e.g., abundance estimates derived from relative indices of abundance). The choice is to fit the terminal model after a sequence of models that produce estimates used as data or to integrate the raw data into the terminal analysis. Each approach has advantages and disadvantages that constitute a suite of trade-offs which are described in further detail. These trade-offs are investigated here by comparing a sequential analysis to an integrated analysis for Pacific salmon stock assessment, using the Kuskokwim River Chinook salmon stock, located in western Alaska, as a case study. The terminal analysis in both cases is an age-structured Bayesian state-space spawner-recruit model used to derive estimates of productivity and capacity for use in setting escapement targets. Primary findings showed that the two approaches returned similar estimates of population dynamics parameters and management reference points, both in terms of point estimates and uncertainty, showing that the treatment of measurement uncertainty in the sequential analysis was an adequate treatment of the information content of the data. The primary advantage of the integrated analysis was the added realism of sharing calendar year abundance data between brood years, but came at the cost of slow run times. This exercise showed that while there is a trade-off between sequential and integrated analyses in terms of model complexity and realism, it may not be large enough to warrant an integrated analysis in all cases. We also investigated the sensitivity of estimates of total abundance from the integrated model to changes in the assumed variances of harvest observations, and the assumed precision and bias of the telemetric mark-recapture estimates of drainage-wide escapement. We found that drainage-wide estimates of escapement were insensitive to changes in the assumed sampling variance of harvest data and telemetric mark-recapture escapement estimates. However, the escapement estimates were very sensitive to bias in the telemetric mark-recapture data as the estimates rescaled proportionally to accomodate the change in the mark-recapture data. These findings highlight the common-sense notion that obtaining accurate estimates of drainage-wide escapement from telemetric mark-recapture studies is very important in the pursuit of unbiased estimates from this run reconstruction model.

## INTRODUCTION

Stock assessment scientists are often faced with decisions regarding how to incorporate information into models. This is particularly true when there are various types of information, some of which are raw data and some are estimates from previous models (e.g., catch-at-age data versus estimates of abundance; Maunder and Punt 2013). Additionally, these cases arise when two processes give rise to the observed dataset, but only one is of interest, such as when data are produced by a biological population process and a measurement process, but inference is desired only for the population process. In these cases, the practitioner has the option to either separate the information into a sequence of analyses or integrate the raw information into one analysis. This choice may at first seem arbitrary and inconsequential, but may prove to have substantial ramifications on the results and interpretation of the assessment model (Brooks and Deroba 2015).

Under the sequential analysis approach (often termed “two-step”), raw data are used in one model to produce estimates of interest that are then passed to another model as either data or informative prior distributions (Michielsens et al. 2008). A common example of this practice in stock assessment modeling is the estimation of spawning stock biomass and a recruitment time series from a Virtual Population Analysis-type model that are then passed to a spawner-recruit model to estimate parameters that govern the population dynamics (Maunder and Punt 2013). This sequential approach allows for the inclusion of information in a summarized form into a more simplified terminal model that requires it in that form. However, as noted by Maunder (2001), the sequential approach has some potential disadvantages including loss of information in the raw data that could be exploited by the terminal analysis if not already summarized, inadequate treatment of uncertainty in the observation level-data, and reduced diagnostic ability.

By contrast, an integrated analysis attempts to incorporate the raw data (as fully as appropriate) into a single model (Fournier and Archibald 1982). An integrated analysis is often conducted with the goal of preserving the complete information content of the data and partitioning uncertainty in the analysis due to observation error and process variation. These types of integrated population models are gaining popularity in ecological assessments in both applied and academic settings (Schaub and Abadi 2011; Royale and Dorazio 2008), however, due to the inherent complexity of integrated analyses, they present unique problems, such as how to formulate the joint likelihood, convergence issues, and computational complexities (Maunder and Punt 2013). However, with recent advances in computing power and statistical approaches to fitting complex models, particularly Bayesian estimation using Markov chain Monte Carlo (MCMC) methods, these barriers are becoming less troublesome for the applied stock assessment practitioner.

Even with integrated analyses, however, the investigator is tasked with deciding which data to include in its raw form versus which information should be summarized before inclusion in the model. As such, it is more appropriate to view the contrast between sequential and integrated analyses as a continuum rather than a dichotomy. On the extreme sequential side of the continuum, all raw data are summarized or combined into estimates prior to being included in the final model. On the extreme integrated side of the continuum, every raw data point is included as an observation of the system and is used for inference. There exists intermediate scenarios between these extremes where some data are summarized pre-analysis and some data are included in raw form. As one moves the analysis in either direction along this continuum, the relative advantages and disadvantages of either approach should become apparent in the output of the whole analysis.

Assessment approaches for Pacific salmon typically involve collecting observations of annual spawner abundance and harvest then combining this information into a brood table to obtain brood year recruits based on the year and age at which the fish returned (e.g., Clark et al. 2009; Bue et al. 2012). Note that only the calendar year processes (escapement and harvest) can be observed, which are then used to obtain brood year recruitment that cannot be directly observed. Oftentimes, spawner

abundance and harvest are not complete counts and must be estimated (e.g., run reconstruction; Shotwell and Adkison 2004). These estimates are then passed to a spawner-recruit analysis to estimate productivity, carrying capacity, and variation in recruitment for the derivation of biological reference points used in setting escapement goals (Clark et al. 2009). This approach constitutes a sequential analysis. However, in some cases it is possible to integrate the run reconstruction model directly into the spawner-recruit analysis whereby brood year spawner and recruit abundances are reconstructed from observational-level data and a spawner-recruit relationship is simultaneously fit to these estimates. This approach may have the benefits of more fully addressing uncertainty due to the interaction between the raw data (i.e., observation model) with the spawner-recruit analysis (i.e., process model) and the sharing of information between calendar years due to the intrinsic link between calendar years via the spawner-recruit analysis (Maunder 2001).

In this chapter, I present the application of an analysis that moves a stock assessment further along the continuum from a sequential two-step analysis to a more fully integrated analysis. Potential trade-offs of integrating the analysis will be investigated by independently running both a sequential and an integrated analysis and comparing relevant quantities of interest. The terminal model in both cases is fit using a Bayesian state-space framework to allow for simultaneous incorporation of both measurement error in the input information and process variation in the spawner-recruit relationship (Fleischman et al. 2013). This analysis is conducted using the Kuskokwim River Chinook salmon stock, located in western Alaska, as a case study. The objectives of this exercise are to (1) confirm that the integrated analysis is computationally feasible in this case given the information content of the data and the model complexity, (2) compare the findings from the integrated model to a sequential assessment to determine if the proposed advantages of integrated analyses are evident in this case and (3) assess the sensitivity of the integrated analysis to parametric uncertainty in terms of assumed variance components and biases in scaling information. The primary *a priori* expectation is that there should be greater posterior uncertainty in quantities from the integrated analysis than the sequential analysis because the measurement error in the raw data is allowed to interact more directly with the spawner-recruit model in the integrated analysis.

## METHODS

### ***Study System***

The Kuskokwim River is the second largest drainage system in the state of Alaska (~130,000 km<sup>2</sup>), with the main stem traveling approximately 1,500 km from its headwaters in the Kuskokwim Mountains to the southwestern coast where it empties into the Kuskokwim Bay of the Bering Sea (Hamazaki et al. 2012). Historically, the Kuskokwim River has provided the largest subsistence fishery for Chinook salmon in the state (39 year average ~70,000 fish/year), but has seen low runs in recent years leading to substantial conservation measures including fishery closures for Chinook salmon and other species. While the directed commercial fishery for Chinook salmon was closed in 1987, incidental harvest (28 year average ~14,000 fish/year since closure) of Chinook salmon still occurs in the chum salmon *O. keta* and sockeye salmon *O. nerka* commercial fisheries (Linderman and Bergstrom 2009).

Assessment data for the Kuskokwim River Chinook salmon stock date back to 1976 and consist of escapement counts made at weirs ( $n = 6$ ), aerial surveys flown at peak spawning ( $n = 14$ ), total annual harvest estimates from the commercial and subsistence fisheries, historical catch-per-unit-effort data from the commercial fishery, age composition data from various areas in the drainage, and large-scale mark-recapture estimates (Table 1; Bue et al. 2012).

### ***Integrated Model Structure***

The integrated model combines the estimation of abundance and population dynamics parameters into one Bayesian state-space model (as opposed to separate models, described in

*Sequential Analysis Structure* below). This framework moves the assessment further on the continuum towards an integrated analysis. This model is hereafter referred to as the “integrated model”. The integrated model structure has been applied to other salmon assessments across the state of Alaska including the Kenai early and late Chinook runs (McKinley and Fleischman 2013; Fleischman and McKinley 2013) and shows promise in terms of dealing with measurement error and other statistical problems inherent to traditional spawner-recruit analyses (Chapter 1, this Thesis; Fleischman et al. 2013). The general model formulation follows that of Fleischman et al. (2013) who presented an integrated analysis for the Karluk River Chinook salmon population on Kodiak Island, Alaska, which is monitored annually using one weir and harvest sampling. The spawner-recruit analysis served as the unobserved state process submodel which was informed by escapement (weir) and harvest observations in the observation process submodel. The integrated model presented in this chapter extends this formulation to include all of the assessment data from the Kuskokwim River Chinook salmon stock collected over the past 39 years (Table 1)

The drainage-wide integrated model was built specifically for the Kuskokwim River drainage Chinook salmon stock. Within the state-space model there are two submodels which are fully described below. The process submodel defines the true unobserved state of the system: the population dynamics of the salmon stock including Ricker spawner-recruit productivity and brood year-specific maturation schedules. Based on these components, the process model produces an expected calendar year run abundance. The observation submodel then links these true states to the observed states by incorporating all of the observations that have been made of the system for the past 39 years (1976-2014). The integrated model is cast in the Bayesian mode of inference and uses MCMC methods to estimate the joint posterior probability distribution for all non-fixed quantities in the model. The process and observation submodels emulated the approaches followed by Fleischman et al. (2013) and Bue et al. (2012), respectively.

#### *Process Submodel*

Hereafter, the term “run” refers to the total number of fish coming to the river mouth in a single calendar year which is made of several brood years (cohorts) and the “return” is all of the fish that return in multiple years but were all spawned in the same brood year ( $R_y$ ) and can thus be used interchangeably with the term “recruitment” or “recruits”. Returns of Chinook salmon to the Kuskokwim River were treated as unobserved states modeled using the linearized Ricker (1954) spawner-recruit model, but adapted for lag-1 auto-regressive [AR(1)] lognormal process error:

$$\ln(R_y) = \ln(S_y) + \ln(\alpha) - \beta S_y + \phi \omega_{y-1} + \varepsilon_y \quad (1)$$

where  $S_y$  is the escapement in year  $y$  that produced brood year returns  $R_y$ .  $\alpha$  is the productivity parameter,  $\beta$  is the capacity parameter, and  $\phi$  is the AR(1) coefficient that specifies the strength of serial autocorrelation in the recruitment time series. AR(1) process errors were used to account for trends in juvenile survival, wherein positive recruitment residuals in brood year  $y-1$  lead to higher-than-expected residuals in brood year  $y$ , and vice versa (i.e., autocorrelation). Lognormal errors were used to allow residual variance to increase as the mean increases, a phenomenon commonly observed in spawner-recruit residuals and ecological data in general (Hilborn and Walters 1992).  $\omega_y$  is the model residual:

$$\omega_y = \omega_{y-1} + \varepsilon_y \quad (2)$$

where  $\varepsilon_y$  are independent [non-AR(1)] normally-distributed process errors with standard deviation  $\sigma_R$ . The first seven brood year returns, which could not be linked to monitored escapement data through the spawner-recruit relationship, were modeled as random effects drawn from a shared lognormal distribution with parameters  $\ln(R_0)$  and  $\sigma_{R_0}$ .

The model was age-structured to account for recruits from a given brood year being able to return at age four, five, six, or seven. Furthermore, the model allowed for annual variation the

proportion of fish mature-at-age ( $p_{y,a}$ ) from a given brood year (maturation dictates the age-at-return). Brood year maturation schedules were modeled as Dirichlet random effects drawn from a common Dirichlet distribution following Fleischman et al. (2013). The Dirichlet is the conjugate distribution for the multinomial distribution, just as the beta distribution is for the binomial (i.e., it models the probability of success for each possible outcome; here the probability of maturing and returning at a given age; McCarthy 2007). Rather than utilizing the built-in Dirichlet distribution in JAGS, the common (i.e., shared) distribution was implemented hierarchically using the distributions that make up the Dirichlet. Formulating the Dirichlet in this fashion allowed for the hyperparameters to be estimated and monitored including the concentration parameter  $D$ , which controls the variation in maturation-at-age between brood years (smaller values, more variation) and the expected frequencies of maturation at age ( $\gamma_a$ ). The Dirichlet distribution was constructed by using a series of independent Gamma distributions. The hyperparameter vector of expected frequencies for maturing at each age ( $\gamma$ ) were independent gamma random variables. These can be used to return the central tendencies of the proportion maturing-at-age from any given brood year from the common (i.e., shared) Dirichlet:

$$\pi_a = \frac{\gamma_a}{\sum_a \gamma_a} \quad (3)$$

Brood year-specific maturation schedules were then drawn from this common distribution:

$$p_{y,a} = \frac{g_{y,a}}{\sum_a g_{y,a}} \quad (4)$$

$$g_{y,a} \sim \text{Gamma}(\text{shape} = \gamma_a, \text{rate} = 1.0) \quad (5)$$

The number of fish of age  $a$  in the run occurring in year  $y$  ( $N_{y,a}$ ) was:

$$N_{y,a} = R_{y-a} p_{y-a,a} \quad (6)$$

And the total run abundance in calendar year  $y$  was:

$$N_y = \sum_{a=4}^7 N_{y,a} \quad (7)$$

The relevant biological reference points for the calculation of escapement goals were calculated within the model which allowed for their marginal posterior distributions to be summarized like any other quantity in the model. The three most relevant reference points are (1)  $S_{eq}$ , the spawner abundance that produces replacement recruitment to keep the population at equilibrium, (2)  $S_{MSY}$ , the spawner abundance that produces maximum sustained yield (MSY), and (3)  $S_{MAX}$ , the spawner abundance that produces maximum recruitment. Before calculating the reference points,  $\alpha$  was corrected for lognormal process error and autocorrelated recruitment residuals (Hilborn 1985)

$$\ln(\alpha_c) = \ln(\alpha) + \frac{\sigma_R^2}{2(1-\phi^2)} \quad (8)$$

Where  $\sigma_R$  is the lognormal process error around the expected spawner-recruit curve. The relevant biological reference points were then calculated:

$$S_{MAX} = \frac{1}{\beta} \quad (9)$$

$$S_{eq} = \ln(\alpha_c) * S_{MAX} \quad (10)$$

$$S_{MSY} = S_{eq} * (0.5 - 0.07 * \ln(\alpha_c)) \quad (11)$$

### Observation Submodel

The role of the observation submodel is to synthesize all of the data from the past 39 years from Kuskokwim Chinook salmon sampling and to fit it to the true states given by the process submodel. For simplicity of presentation, the observation submodel can be partitioned into five main components representing the primary data sources: (1) weir and aerial survey counts, (2) commercial catch-per-unit-effort (CPUE), (3) annual escapement estimates from mark-recapture studies, (4) age-composition data in the form of scale counts, and (5) total annual harvest components for the commercial and subsistence fisheries. These data were supplied by ADF&G biologists.

(1) *Weirs and Aerial Surveys* – Weir projects ( $n = 6$ ) were operated by placing a barrier in the stream channel with a narrow passage opened periodically through which fish are counted as they pass (Blain et al. 2014). Aerial surveys ( $n = 14$ ) were flown during the expected peak of the run to survey a subset of streams that could not be sampled by weirs due to logistical or monetary constraints (Hansen and Blain 2014). Weirs and aerial surveys could not count all escaping fish to a tributary and therefore counts made by each project should be interpreted as indices of escapement rather than censuses. These counts can be proportionally related to total escapement such that the expected number of fish counted in project  $j$  in year  $y$  is:

$$I_{jy} = \frac{\hat{S}_y}{k_j} \quad (12)$$

where  $\hat{S}_y$  is the total annual escapement in the drainage predicted by the model,  $I_{jy}$  is the predicted count from weir or aerial survey project  $j$  in year  $y$  and  $k_j$  is a proportionality coefficient that scales the index to the total drainage-wide escapement. The ‘hat’ over  $S_y$  in equation 2.12 denotes that this value is the predicted drainage-wide escapement in a given year, to make the distinction between observed (estimated) escapements from mark-recapture studies. Note that  $k_j$  was constant over time (has no  $y$  index), which made the assumption that the proportion of total escapement counted at each tributary was constant every year. These proportionality coefficients represented a mean inverse proportion of the total escapement that each project counted each year.

These index counts were assumed to have a negative binomial sampling distribution to account for over-dispersion (Hilborn and Mangel 1997). The negative binomial distribution was formulated to include an over-dispersion parameter ( $r_j$ ) which represents the amount of extra-Poisson variation present in the dataset. The tributary counts were over-dispersed because they included both process error (annual variation in the proportion of the total escapement going to each tributary) and measurement error associated with annual sampling.

(2) *Commercial CPUE* – The commercial CPUE component used weekly ( $w$ ) catch ( $C_{yw}$ ) and effort ( $B_{yw}$ ) data from the commercial fishery that occurs in district W1 (Lower River) to serve as a relative index of abundance. Effort data were expressed in units of permit hours per week  $w$ . Expected catch was calculated using the Baranov catch equation where weekly catch was a function of the fishing mortality and the number of fish available for harvest. The number of fish reaching district W1 by week was calculated using run timing estimates from a gill net test fishery operated daily in-season in Bethel, AK. The proportion of the annual run  $P_{yw}$  present in week  $w$  of year  $y$  was calculated using the test fishery CPUE:

$$P_{yw} = \frac{CPUE_{yw}}{\sum CPUE_{yw}} \quad (13)$$

and the number of fish in district W1 in week  $w$  of year  $y$  ( $W_{yw}$ ) was:

$$W_{yw} = N_y P_{yw} \quad (14)$$

where  $N_y$  is the expected run size in year  $y$  given by the process submodel. The expected catch in week  $w$  of year  $y$  was:

$$C_{yw} = W_{yw}(1 - e^{-qB_{yw}}) \quad (15)$$

where  $q$  is the estimated catchability of the gear used. Due to historical changes in gill net restrictions and technologies, separate catchabilities must be estimated for different time periods in equation 2.15. During the first few weeks in the early years of data collection there were no mesh size restrictions were in place so  $q_{unr}$  was used. From 1976-1984 a 6-inch stretched mesh size restriction was put into place ( $q_{res}$ ). In the early 1980s a new gill net material was introduced that increased the efficiency of gill nets (Bue 1986), and it was assumed that this gear was used from 1985 to the present ( $q_{mono}$ ).  $C_{yw}$  was fit to the observed catch in year  $y$  and week  $w$  using a lognormal likelihood density function.

(3) *Mark-Recapture Scaling* – In order to calibrate the relative abundance indices to drainage-wide abundances, a large-scale mark-recapture study was conducted from 2002 to 2007 using the Chapman modification to the Lincoln-Petersen estimator (Schaberg et al. 2012). The 2002 estimate violated the assumption of equal probability of recapture, and thus was excluded from this analysis (Schaberg et al. 2012). Tagging occurred in Kalskag, AK and recaptures were made by observing tagged fish as they passed weirs, so the mark-recapture estimates only accounted for in-river abundance upstream of the tagging site ( $S_u$ ). Downstream abundance ( $S_d$ ) was estimated using the Parken et al. (2006) habitat-based model (Schaberg et al. 2012). Total observed (estimated) escapement was calculated:

$$S_y = S_u + S_d - H_u \quad (16)$$

where  $H_u$  is the harvest that occurred upstream of the tagging site. Uncertainty in  $S_u$ ,  $S_d$ , and  $H_u$  were estimated using bootstrapping, predictive error from the Parken et al. (2006) regression model, and survey/estimation variation, respectively. Because the variance of a sum is the sum of the component variances:

$$\text{var}(S_y) = \text{var}(S_u) + \text{var}(S_d) - \text{var}(H_u) \quad (17)$$

The expected escapement  $\hat{S}_y$  was fit to the observed escapement  $S_y$  given in equation 2.16 using a lognormal likelihood.

(4) *Age-Composition* – Age-composition proportions for this model were the same as those used by Hamazaki et al. (2012), with additional years (2012-2014) obtained directly from ADF&G biologists. Those proportions were calculated by taking a weighted average of all age sampling projects including commercial and subsistence fisheries, Bethel test fishery, and weir projects (Bue et al. 2012). The result of this weighted average are vectors of calendar year proportions-at-age  $q_{obs,y}$  that sum to one. These proportions were then used to inform the expected age proportions using a multinomial likelihood by multiplying the  $q_{obs,y}$  vector by the effective sample size ( $n_y$ ) for year  $y$ :

$$X_y = q_{obs,y} n_y \quad (18)$$

$$X_y \sim \text{multi}(q_y, n_y) \quad (19)$$

where

$$q_{y,a} = \frac{N_{y,a}}{N_y} \quad (20)$$

from the process submodel. In the multinomial distribution, the  $n_y$  specifies how closely the modeled expectation must fit to the observed data  $X_y$ , similar to the variance component in a normal likelihood (larger  $n$  results in stronger information content about age proportions and a closer model fit; Maunder 2011). The values of  $n_y$  were chosen by ADF&G such that age-composition data collected early in the data time series were more flexible in the model than more recently collected age composition data when the scale sampling efforts were more intensive and presumably more representative. For the time periods 1976-1999 and 2000-2014,  $n_y$  was set at 25 and 100, respectively following Hamazaki et al. (2012).

(5) *Harvest Observations (Total Annual)* – The original Bue et al. (2012) run reconstruction model assumed harvest was known perfectly. However, in the state-space framework, it is possible to model the harvest and observation processes that gives rise to the observed harvest estimates in order to include measurement error (i.e., uncertainty in harvest estimates). This component was divided into commercial and subsistence harvests, with each having different assumed observation variances. Since commercial harvest reporting is mandatory, there was assumed to be less observation error than for subsistence harvest. These harvests were modeled using a total finite harvest rate  $U_y$  which was further divided:

$$U_{com,y} = U_y p_{com,y} \quad (21)$$

$$U_{sub,y} = U_y (1 - p_{com,y}) \quad (22)$$



$$H_{com,y} = U_{com,y}N_y \quad (23)$$

$$H_{sub,y} = U_{sub,y}N_y \quad (24)$$

$$\hat{S}_y = N_y(1 - U_y) \quad (25)$$

where  $p_{com,y}$  is the annual proportion of the total harvest rate that was made up of commercial harvest. These harvest components were fit using a lognormal likelihood. It was the  $\hat{S}_y$  derived in equation 2.25 that was used throughout the rest of the observation submodel (e.g., equation 12).

### ***Sequential Analysis Structure***

In contrast to the integrated analysis, the sequential analysis functioned by estimating abundance and the population dynamics (i.e., spawner-recruit) parameters in two separate and sequential models. The first model was a drainage-wide run reconstruction model (hereafter “run reconstruction”) which was developed by Bue et al. (2012) following methods first proposed by Shotwell and Adkison (2004). The run reconstruction served to synthesize the escapement and harvest observations into calendar year run abundance estimates and uncertainty, expressed as a coefficient of variation on each run estimate (CV). Equations 12-17, 25 from the integrated model described above were based on the original run reconstruction formulae presented in Bue et al. (2012) and the two components are equivalent in nearly every regard. Primary differences between the observation submodel of the integrated analysis and the run reconstruction from the sequential analysis is that the latter is fit using maximum likelihood methods, it scales by fitting to a drainage-wide total abundance estimate based on mark-recapture rather than escapement, and assumes harvest is known without error.

In order to estimate the population dynamics parameters, the terminal model in the sequential analysis was a post-hoc Bayesian state-space spawner-recruit model (hereafter “post-hoc model”). This model was originally developed by Hamazaki et al. (2012) and was based on the aforementioned Fleischman et al. (2013) age-structured Bayesian state-space spawner-recruit model. The post-hoc model treated the run abundance estimates from the run reconstruction model as independently observed data with fixed uncertainty captured in the CVs on each run estimate. The post-hoc model was made up of equations 2.01-2.11 and 2.18-2.25 from the integrated analysis. Note that measurement error and process variation in fish abundance were estimated in separate models under the sequential assessment.

### ***Comparisons between the Integrated and Sequential Kuskokwim Assessments***

In order to investigate the trade-offs of integrating the analysis, the most relevant model output from the sequential assessment (run reconstruction then post-hoc spawner-recruit analysis) was compared to the corresponding output of the integrated analysis. Important quantities for comparison to the integrated model included: (1) total run abundance from the run reconstruction and the expected run abundance from the post-hoc model, (2) key spawner-recruit parameters from the post-hoc model, (3) scaling and over-dispersion parameters in the run reconstruction components, and (4) biological reference points. In addition to point estimates (posterior means), uncertainty was expressed in terms of 95% Wald confidence intervals derived from the Hessian matrix (inverse variance-covariance matrix) for the run reconstruction and Bayesian 95% credible intervals for the integrated and post-hoc models. It is important to note that the equations for all model predictions and likelihoods were the same between the integrated model and the sequential analysis. The key difference between the two approaches was the component in the sequential analysis that linked the post-hoc model to the run reconstructed estimates via a likelihood. In the integrated model, the run abundance estimates were simply a required derived quantity, not information that the model was fit to.

To investigate potential biases, the spawner-recruit parameters of interest were further compared to a traditional ordinary least-squares linear regression technique that is more commonly applied to spawner-recruit analyses. To conduct this traditional analysis, the run abundance estimates (ignoring their uncertainty) were split into calendar year abundance-at-age according to age-composition information from scale counts and the appropriate years and ages were summed to arrive at brood year returns. To obtain brood year escapement, calendar year harvest estimates were subtracted from total calendar year abundance estimates. These brood year escapement and returns were then used in a linear regression:

$$\ln\left(\frac{R_y}{S_y}\right) = \ln(\alpha) - \beta S_y + \varepsilon_y \sim N(0, \sigma) \quad (26)$$

to obtain estimates of  $\alpha$ ,  $\beta$ , and  $\sigma$ . Note that this method ignores measurement uncertainty in the quantities  $R_y$  and  $S_y$  and assumes that the recruitment residuals are independent and identically distributed through time, which ignores any time series patterns. These assumptions make the traditional spawner-recruit analysis susceptible to both time series and error-in-variables biases (Walters 1985; Walters and Ludwig 1981). Uncertainty in the parameter estimates and derived reference points was estimated using parametric bootstrapping: randomization of the regression residuals, summation with the predicted values from the original fit, and re-estimation of the model parameters (Hamazaki et al. 2012).  $\alpha$ ,  $\beta$ , and  $\sigma$  were saved and biological reference points were calculated for each bootstrapped sample using equations 2.05-2.08.

### ***Sensitivity Analysis***

There are cases where a single dataset includes both measurement and process error and the model is required to separate these sources of variation under the state-space framework. Oftentimes, the model cannot parse out what variation is due to the underlying biological or fishery process and what variation is due to measurement. In these cases, one of the sources of variation must be assumed known. One example is the harvest model presented in equations 2.21-2.25. The variation in the observed harvest arises from two processes: (1) annual variation in the true  $U_y$  where fishers harvest fish from the annual run and (2) in the measured (estimated) perception of the total harvest. When provided with an annual data point to inform this component, it is impossible for the model to partition which variation should lead to uncertainty in the harvest process and which should be assigned to measurement error. It is therefore necessary to provide an assumed variance, and for this case an assumed known observation error was provided and the model was allowed to handle process variation. This was conducted by assuming an observation CV. The CVs on harvest were combinations of output from harvest estimation models used by ADF&G and assumptions based on expert judgement (Hamazaki et al. 2012). The components of the model that included these fixed observation CVs were (1) commercial harvest, (2) subsistence harvest, and (3) the mark-recapture estimates of escapement. Model sensitivity to these assumptions was investigated by altering the assumed CV within reasonable ranges and comparing the resulting changes to the original model. All assumed variances were altered by 50% and 200% to assess the impact of being more or less certain about the corresponding estimates, respectively. Each sensitivity scenario was conducted in isolation of other scenarios so that any changes in model output could be attributed to a single source.

Considering that the model was reliant on the mark-recapture estimates of escapement for scaling information, I thought it prudent to investigate the model sensitivity to biases in these estimates. The mark-recapture estimates were altered by  $\pm 5\%$ ,  $10\%$ , and  $20\%$  and the model was fitted to the altered estimates. The relative uncertainty in these estimates was not altered (i.e., the assumed CV remained the same as the base model). Because these estimates constitute the principle scaling information, one would expect to see a proportional increase/decrease in the model expected

escapement in not only the years with mark-recapture estimates, but in the whole escapement time series as well.

### **Computation**

The integrated model was fit using Bayesian integration with MCMC methods to sample from the joint posterior probability distribution. MCMC sampling was conducted using parallel computing with the JAGS software (“Just Another Gibbs Sampler”, Plummer 2013) implemented through R (R Core Development Team 2014) using the R package “R2jags” (Su and Yajima 2015). Prior distributions on all unknown parameters were uninformative and their structures were based on recommendations from Fleischman et al. (2013) and Bolker (2008) with necessary truncations to prevent the sampler from drawing implausible parameter values (e.g.,  $\log(\alpha)$  had a diffuse normal prior truncated at zero). MCMC sampling was conducted using two chains with different initial values to verify convergence and to detect potential multiple solutions. Convergence of the chains was assessed with visual inspection of the posterior distribution sampled by each chain, trace plots, and the Gelman-Rubin statistic (Gelman et al. 2004). MCMC sampling of all Bayesian models (i.e., integrated model scenarios and post-hoc model) involved a burn-in period of 500,000 iterations, 1,000,000 post-burn-in iterations, and a thinning interval of 200 iterations, using two chains to ensure convergence. This resulted in a total of 10,000 posterior samples retained for analysis. The model converged under these specifications, as evidenced by Gelman-Rubin statistics of  $<1.05$  for every estimated parameter. Point estimates (posterior mean) and Bayesian 95% credible intervals (2.5 and 97.5 percentiles of posterior distribution) for quantities of interest were calculated from their respective marginal posterior distributions.

## **RESULTS**

### ***Comparisons to the Sequential Kuskokwim Assessment***

#### ***Run Abundance and Uncertainty***

The integrated model performed similarly with regards to trends in total run abundance point estimates from both the run reconstruction and post-hoc model (Figure 1, upper panel). The integrated model exhibited a mean 2% decrease in run abundance from the run reconstruction, and a mean 3% decrease from post-hoc model. Both the integrated model and the post-hoc model exhibited some shrinkage in run abundance, which is a statistical characteristic of hierarchical models where quantities at intermediate levels are pulled to the values suggested by higher-level distributions (i.e., hyperdistributions) since all sources of information are simultaneously considered by the model. In this case, the spawner-recruit component treated the extreme high and low run abundances as unlikely to have occurred and tended to pull them toward the mean spawner-recruit curve. This was possible because the run estimates used by the models were not fixed, but had uncertainty and flexibility and could thus be adjusted to satisfy the spawner-recruit likelihood. Both the post-hoc model and the integrated model shrank by approximately the same amount and in the same years.

All three models exhibited a slightly decreasing temporal trend in run abundance CV (Figure 1, lower panel), due to the addition of more escapement monitoring projects through time (Table 1). The run reconstruction CVs were the most variable between years, while the integrated and post-hoc models resulted in CVs that fluctuated less from year to year (Figure 1). This finding resulted from total abundance being estimated nearly independently each year in the run reconstruction (tributary proportionality scalars were shared by all years, which links them) whereas years were linked explicitly in the integrated and post-hoc models via the spawner-recruit relationship. For example, if escapement in one year was uncertain due to very few monitoring projects, it could be informed by later years based on how many recruits it produced. The CVs were relatively similar between the post-hoc model and the integrated model, although the integrated model did consistently show slightly more uncertainty in run abundance (mean 1.2 CV percentage points greater than post-hoc model).

### *Spawner-Recruit Parameters and Biological Reference Points*

Overall, the spawner-recruit point estimates between the integrated and post-hoc models were similar (Table 2). The integrated model indicated that the maximum productivity of the stock ( $\alpha$ ) was between 2.45 and 12.36 with 0.95 probability whereas the post-hoc model indicated it was between 2.39 and 11.74. Although the credible bounds on  $\alpha$  were slightly wider for the integrated model, the post-hoc model estimated slightly wider credible intervals for most other spawner-recruit parameters of interest (Table 2). It is clear that overall, both the integrated and post-hoc models treated the spawner-recruit components of the assessment similarly (Figure 2). When the traditional linear regression technique was implemented on the run reconstruction estimates, the estimate of  $\alpha$  was much higher (7.29; 95% CI 5.62-9.34) than both the integrated and post-hoc models. The linear regression technique showed approximately the same amount of residual variation in the spawner-recruit relationship ( $\sigma_R = 0.29$ ; 0.20-0.36) as the integrated ( $\sigma_R = 0.23$ ; 0.15-0.34) and post-hoc ( $\sigma_R = 0.24$ ; 0.16-0.35) models. In all model runs, the  $\gamma$  parameters (those governing the expected frequencies of fish maturing at age from all brood years) were the slowest mixing (i.e., showed the most MCMC autocorrelation), along with all parameters that were derived from them. The mean probability of maturing-at-age across brood years ( $\pi_a$ ) were essentially the same between the integrated and post-hoc models (Table 2). The amount of serial autocorrelation in the recruitment time series ( $\phi$ ), had very similar posterior distributions between both models (Table 2). Based on the posterior mean of  $\phi$  under both models, there was a large amount of serial autocorrelation in the recruitment residuals, indicating the potential for time series bias if not accounted for.

With regards to biological reference points, the two Bayesian models resulted in relatively similar point estimates, yet the integrated model estimated wider credible bounds for all three quantities (Table 2). The traditional regression approach did not return reference points that were systematically higher or lower than the two Bayesian models, however the confidence bounds were much narrower.

### *Index Scalars and Over-dispersion Parameters*

The proportionality coefficients that controlled the scaling of weir and aerial survey counts to the drainage-wide mark-recapture estimates were very similar between the run reconstruction and the integrated model (integrated scalars mean 3.7% larger than run reconstruction; Figure 3). This finding showed that the two models used the escapement index information similarly: on average, each tributary count made up the same proportion of the total escapement in both the run reconstruction and the integrated model. It is clear that both models dealt with over-dispersion similarly as well, with the integrated model estimating slightly more over-dispersion in nearly all projects (Figure 4).

### ***Sensitivity Analysis***

#### *Assumed Observation CV on Harvest Estimates*

When the assumed variances on commercial and subsistence annual harvests were halved or doubled, the model responded by fitting more closely or loosely, respectively (Figure 5). Doubling the CV down-weights those data and loosens the extent to which the model must fit to them (Francis 2011). Both commercial and subsistence harvest exhibited this behavior, although when the subsistence harvest CV was doubled, it lead to much more variation in the expected harvest than did commercial harvest. This discrepancy was due to the greater assumed baseline CV on subsistence harvest estimates (CV varied annually between 1%-10%) than commercial harvest (2% CV in all years). However, alternating the precision of these data had little effect on the point estimates of drainage-wide escapement.

#### *Assumed Observation CV on Mark-Recapture Estimates of Escapement*

Altering the precision of the 2003-2007 drainage-wide escapement estimates had a very small effect on the model point estimates of escapement in those years (Figure 6). However, when the assumed CV was halved, the escapement uncertainty in the corresponding years was reduced by 33% on average when compared to the base model. Similarly, when the assumed CV was doubled, the escapement uncertainty in those years increased by 49% on average. Not only did these changes in assumed variances affect uncertainty in the years with mark-recapture estimates, but the uncertainty the whole escapement time series was reduced or increased depending on the scenario.

#### *Sensitivity to Changing the Drainage-Wide Mark-Recapture Estimate*

When the sensitivity of the model to biases in the scaling information (i.e., drainage-wide mark-recapture estimates) was investigated, the model responded by scaling with the altered estimates (Figure 7). Changing the mark-recapture estimate by  $\pm 5\%$ ,  $10\%$ , and  $20\%$  resulted in average changes of  $\pm 4\%$ ,  $9\%$ , and  $18\%$ , respectively. These percent changes were consistent for years with mark-recapture estimates and the whole time series (Figure 8).

### DISCUSSION

With the analyses presented in this chapter, it is clear that the integrated model is a feasible model formulation for the Kuskokwim River Chinook salmon fishery assessment. The model converged well, did not show any major MCMC or residual problems, and provided very similar estimates to the current two-step assessment approach. Similar point estimates were expected, as both assessment approaches used the same data and the same equation structures (e.g. same tributary count model, same spawner-recruit function, etc.). The key difference in the two analyses is that under the sequential analysis, the run reconstruction is fit with maximum likelihood estimation that provides run estimates and uncertainty to the post-hoc model, whereas this happens simultaneously under the integrated model in a single Bayesian framework. The overall consistency in estimates between the two approaches suggests that integrating the analysis did not change the inference regarding stock abundance, productivity, carrying capacity, or management recommendations.

Somewhat contrary to the *a priori* expectations, however, was the similarity in the posterior uncertainty in quantities of interest between the integrated and post-hoc models. It was expected that the integrated model would lead to substantially greater posterior uncertainty in quantities like run abundance, spawner-recruit parameters, and biological reference points due to the direct interaction between the observation-level data (e.g., weir and aerial survey counts) and the process (i.e., spawner-recruit) submodel. As these analyses have shown, the extent of the increase in uncertainty was not as large as expected. Posterior uncertainty in run abundance under the integrated model was slightly higher than the post-hoc model, but less than expected. This difference is likely due to the assumption of statistical independence of the run abundance estimates under the post-hoc model. After investigating the correlations in estimated run abundances from the run reconstruction, it was found that the run estimates were indeed correlated in the run reconstruction model (0.30 mean correlation, 0.57 maximum correlation). The post-hoc model did not carry these correlations forward by treating them as independent observations, which ignored that there was information about other estimates contained in any one estimate and thus resulted in underestimation of the posterior uncertainty in run abundance. One advantage of the integrated model is that it was able to internally address these correlations by estimating the abundances within the same model, as stated by Maunder and Punt (2013). The integrated model estimated slightly more uncertainty in  $\alpha$  and  $D$ , but all other spawner-recruit parameters had slightly more uncertainty under the post-hoc model. The integrated model did result in more uncertainty in the biological reference points, but likely not enough to alter the management recommendations. This modeling exercise gave credence to the two-step assessment in

terms of its ability to deal with measurement uncertainty by capturing it with the CV on the run estimates from the run reconstruction model.

The integrated model proved to be robust to parametric uncertainty, particularly with regards to the assumed fixed CV on the mark-recapture scaling information. The expectation was that when the CVs were increased on the mark-recapture estimates, posterior uncertainty in escapement in that year would increase and vice-versa. Additionally, since the mark-recapture estimates of escapement were the only scaling information provided to the model, changing their uncertainty should not substantially alter the point estimates of escapement in those years. The results of this exercise confirmed these expectations. Even when the CVs on the estimates were doubled, the model did not show any systematic scaling biases or shrinking of the escapement estimates in those years, which showed that the model can perform well even with weakly informative scaling data. The only consequence was that the posterior uncertainty of abundance quantities was increased, which is what would be expected when weaker scaling information is provided to the model. This finding was critical because it proved that the model is robust to measurement error in the scaling information.

As previously mentioned, the only scaling information provided to the model were the drainage-wide estimates of escapement from mark-recapture (and lower river habitat expansion). Thus, altering these point estimates should result in a proportional change in the expected escapement. The results of this exercise confirmed this expectation, not only for escapement in the years with mark-recapture estimates but for the entire escapement time series. This issue is related more to the accuracy of the mark-recapture estimate than the performance of the model. The mark-recapture estimates represent the best-available scientific understanding of drainage-wide escapement, and so sensitivity to these estimates is necessary. This is one case where it is desirable for a model to be sensitive to the accuracy of the input information.

This being said, there are several areas in which the integrated model performed sub-optimally. One case is the slow mixing of the maturation schedule component. This is one of the most complex portions of the model and it comes as no surprise that it mixes slowly. In the early years of the data time series, the scale count frequencies were weakly informative and there were some years with missing scale counts. In these cases, the model could satisfy the data in many different ways so the parameters in this component mixed slowly and MCMC samples were highly correlated. The inclusion of the Dirichlet random effects aided this problem, as the model could use the more informative years to inform the central tendency of the proportions mature-at-age ( $\pi_a$ ). However, this formulation assumes that the central tendency does not change over time (all brood year proportions are drawn from the same multivariate distribution) which may not be realistic given size-selective fishing pressure and potentially biased age sampling. These topics will be dealt with in more detail in Chapter 3.

The slow mixing behavior of the maturation schedule required a large number of posterior samples to be drawn with a wide thinning interval to achieve convergence. This resulted in long run times for the integrated model, nearly 24 hours. The post-hoc model only required approximately eight hours to complete the same number of posterior samples and burn-in period. This extended time period is cumbersome when running a large amount of sensitivity analyses, like those presented in this chapter, but is relatively irrelevant when the model has been finalized and needs only to be run once every year. This disadvantage was expected and is consistent with the literature (e.g., Maunder and Punt 2013) as the integrated analysis was a more complex model involving more priors, data, and likelihoods than the post-hoc model in the sequential analysis.

I presented this analysis as moving further down the continuum towards fully integrated analyses because there are still estimates that are passed to the model as “data” (in the sense that they are fit to in the joint likelihood). One key example is the mark-recapture estimates which represent a summary of a large number of tagged and recaptured fish observations. It would be possible to incorporate the mark-recapture estimation model within this assessment to make it more fully

integrated, however, these analyses have shown that this is likely not necessary. It seems that fitting to their estimates and including their uncertainty in the form of a CV is an adequate expression of the information content of these data.

Integrated analyses, particularly Bayesian state-space models, are gaining popularity throughout the state of Alaska and in the stock assessment and ecological literature as a whole (Maunder and Punt 2013; Schaub and Abadi 2011). The overall theme of using a Bayesian state-space framework for addressing the problems in the traditional spawner-recruit analysis is apparent in the present analysis. The traditional spawner-recruit analysis yielded a higher  $\alpha$  value which is potentially positively biased due to the time series problems inherent to the linear regression approach, as noted by Walters (1985) and Walters and Ludwig (1981). However, this did not lead to systematically higher or lower estimates of biological reference points, as one might expect. Outside of the similar reference point estimates under the traditional approach, it was clear that the Bayesian state-space approaches increased the uncertainty in these quantities, likely due to the incorporation of measurement error into the spawner-recruit analysis. This alone is a meritorious reason to favor the state-space spawner-recruit approach. If one is to manage a stock based on biological reference points, it is critical that the uncertainty in those quantities be fully addressed and taken into account in management objectives and actions.

Other investigators have shown promise for the Bayesian state-space approaches to spawner-recruit analyses for Pacific salmon including Meyer and Millar (2001) using pink salmon data from the Fraser River in British Columbia, the aforementioned Fleischman et al. (2013) analysis on the Karluk River Chinook salmon stock, Fleischman and Borba (2009) with Yukon fall chum salmon, the work with Kenai Chinook salmon stocks by McKinley and Fleischman (2013) and Fleischman and McKinley (2013), and a meta-analysis of Alaskan Chinook stocks by Catalano (2012). All of these analyses used the Bayesian state-space spawner-recruit analysis framework presented here (although some were two-step analyses) and found reasons to suggest that this approach is superior to the traditional linear regression technique in terms of dealing with statistical problems and accounting for measurement error. Clearly, this approach is general enough for application to many different stocks and flexible enough to deal with multiple different species of anadromous Pacific salmon with various life histories and population parameters.

Although the integrated model and the two-step analysis provided very similar overall estimates, there are still advantages to the integrated analysis. The most obvious reason is that the whole process is contained in a single model, which makes for straightforward implementation by biologists and a seamless framework for sensitivity analysis. All of the sensitivity scenarios presented in this chapter were conducted by changing several lines of code and re-running the model. Under the two-step approach, one would have to alter the run reconstruction, make note of the changes, and then pass the altered estimates through to the post-hoc model. In the integrated analysis, this process was seamless, and any changes in the model were reflected in a single joint posterior probability distribution. Another advantage to the integrated analysis is the fact that separating the analysis into two stages forces the run reconstruction model to estimate each year's abundance in isolation of all other data (with the exception of the common tributary scalars across all years within a single project). There is no intrinsic link between years under the run reconstruction, however, there is in reality. In the real system, counts made in four consecutive years are all counting some proportion of fish that were all from the same brood year produced by the same number of spawners, creating a link between years. Integrating the analysis allows the run estimates to be informed by the spawner-recruit analysis (and vice-versa) which accounts for this very real time linkage which ultimately leads to a more realistic model, regardless of how similar the estimates are to the two-step approach. This sort of information sharing is one of the primary advantages to an integrated analysis (Maunder 2011; Maunder and Punt 2013). One could argue that integrated models like the one presented in this chapter are advantageous purely because of this added realism.

## REFERENCES

- Blain, B.J., T.R. Hansen, J.N. Clark, L.M. Robbins, and K.L. Schaberg. 2014. Salmon escapement monitoring on the Kuskokwim River, 2012. Alaska Department of Fish and Game, Fishery Data Series No. 14-34, Anchorage, AK.
- Bolker, B.M. 2008. Ecological models and data in R. Princeton University Press, New Jersey.
- Brooks E.N and J.J. Deroba. 2015. When “data” are not data: the pitfalls of post hoc analyses that use stock assessment model output. Canadian Journal of Fisheries and Aquatic Sciences. 72(4): 634-641.
- Bue, B.G. 1986. Effects of gill net selectivity on sockeye salmon in the Egegik and Naknek-Kvichak Districts, Bristol Bay, Alaska. University of Alaska, Fairbanks.
- Bue, B.G., K.L. Schaberg, Z.W. Liller, and D.B. Molyneaux. 2012. Estimates of the historic run and escapement for the Chinook salmon stock returning to the Kuskokwim River, 1976-2011. Alaska Department of Fish and Game, Report to the Board of Fisheries, Anchorage.
- Clark, R.A., D.R. Bernard, and S.J. Fleischman. 2009. Stock-recruitment analysis for escapement goal development: a case study of pacific salmon in Alaska. Pages 743-757 in C.C. Krueger and C.E. Zimmerman, ed. Pacific salmon: ecology and management of western Alaska’s populations. American Fisheries Society, Symposium 70, Bethesda, Maryland.
- Fleischman, S.J. and B.M. Borba. 2009. Escapement estimation, spawner-recruit analysis, and escapement goal recommendation for fall chum salmon in the Yukon River drainage. Alaska Department of Fish and Game, Fishery Manuscript Series No. 09-08, Anchorage.
- Fleischman, S.J., M.J. Catalano, R.A. Clark, and D.R. Bernard. 2013. An age-structured state-space stock-recruit model for Pacific salmon (*Oncorhynchus* spp.). Canadian Journal of Fisheries and Aquatic Sciences. 70(3): 401-414.
- Fleischman S.J., and T.R. McKinley. 2013. Run reconstruction, spawner-recruit analysis, and escapement goal recommendation for late-run Chinook salmon in the Kenai River. Alaska Department of Fish and Game, Fishery Manuscript Series No. 13-02, Anchorage.
- Fournier, D. and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Canadian Journal of Fisheries and Aquatic Sciences. 39(8): 1195-1207
- Francis, R.I.C. C., 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences. 68(6): 1124-1138.
- Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2004. Bayesian data analysis. Second ed. CRC/Chapman & Hall, Boca Raton, Florida.
- Hamazaki, T., M.J. Evenson, S.J. Fleischman, and K.L. Schaberg. 2012. Spawner-recruit analysis and escapement goal recommendation for Chinook salmon in the Kuskokwim River Drainage. Alaska Department of Fish and Game, Fishery Manuscript Series No. 12-08, Anchorage.
- Hansen, T.R. and B.J. Blain. 2014. Salmon escapement monitoring in the Kuskokwim River, 2013. Alaska Department of Fish and Game, Fishery Data Series No. 14-54, Anchorage, AK.
- Hilborn, R. 1985. Simplified calculation of optimum spawning stock size from Ricker’s stock recruitment curve. Canadian Journal of Fisheries and Aquatic Sciences. 42(11): 1833-1834.
- Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Dluwer Academic Publishers, Boston, Massachusetts.
- Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, New Jersey.
- Irwin, B.J., M.J. Wilberg, M.L. Jones, and J.R. Bence. 2011. Applying structured decision making to recreational fisheries management. Fisheries 36(3): 113-122.
- Linderman, J.C., Jr. and D.J. Bergstrom. 2009. Kuskokwim management area: salmon escapement, harvest, and management. Pages 541-599 in C. C. Krueger and C. E. Zimmerman, ed. Pacific salmon:



- ecology and management of western Alaska's populations. American Fisheries Society, Symposium 70, Bethesda, Maryland.
- Maunder, M.N. 2001. Integrated tagging and catch-at-age analysis (ITCAAN). Pages 123-146 in G.H. Kruse, N. Bez, A. Booth, M.W. Dorn, S. Hills, R.N. Lipcius, D. Pelletier, C. Roy, S.J. Smith, and D. Witherell, ed. Spatial processes and management of fish populations. Alaska Sea Grant College Program Report No. AK-SG-01-02, University of Alaska, Fairbanks.
- Maunder, M.N. 2011. Review and evaluation of likelihood functions for composition data in stock-assessment models: estimating the effective sample size. *Fisheries Research*. 109(2): 311-319.
- Maunder, M.N. and A.E. Punt. 2013. A review of integrated analysis in fisheries stock assessment. *Fisheries Research*. 142:61-74.
- McCarthy, M.A. 2007. Bayesian methods for ecology. Cambridge University Press, New York.
- McKinley, T.R., and S.J. Fleischman. 2013. Run reconstruction, spawner-recruit analysis, and escapement goal recommendation for early-run Chinook salmon in the Kenai River. Alaska Department of Fish and Game, Fishery Manuscript Series No. 13-03, Anchorage.
- Meyer, R. and R.B. Millar. 2001. Bayesian dynamic modeling of stock-recruitment relationships. Technical Report STAT 0004. Department of Statistics, University of Auckland, Auckland, New Zealand.
- Michielsens, C.G.J., M.K. McAllister, S. Kuikka, S. Mantyniemi, A. Romakkaniemi, T. Pakarinen, L. Karlsson, and L. Uusitalo. 2008. Combining multiple Bayesian data analyses in a sequential framework for quantitative fisheries stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*. 65(5): 962-974.
- Parken, C.K., R.E. McNicol, and J.R. Irvine. 2006. Habitat-based methods to estimate escapement goals for data limited Chinook salmon stocks in British Columbia, 2004. Department of Fisheries and Oceans, Canadian Scientific Advisory Secretariat Research Document 2006/083.
- Plummer, M. 2013. JAGS Version 3.4.0 user manual.
- R Development Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Version 3.1.0. Vienna, Austria. <http://www.R-project.org>.
- Ricker, W.E. 1954. Critical statistics from two reproduction curves. *Journal of the Fisheries Research Board of Canada*. 11(5):559-623.
- Royale, J.A. and R.M. Dorazio. 2008. Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations, and communities. Academic Press, San Diego, CA. 444 pp.
- Schaberg, K.L., Z.W. Liller, D.B. Molyneaux, B.G. Bue, and L. Stuby. 2012. Kuskokwim River Chinook salmon run estimates, 2002-2007. Alaska Department of Fish and Game, Fishery Data Series No. 12-36, Anchorage, AK.
- Schaub, M. and F. Abadi. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology* 152 (Suppl 1): S227-S237.
- Schindler, D., C. Krueger, P. Bisson, M. Bradford, B. Clark, J. Conitz, K. Howard, M. Jones, J. Murphy, K. Myers, M. Scheuerell, E. Volk, and J. Winton. Artic-Yukon-Kuskokwim Chinook Salmon Research Action Plan: Evidence of Decline of Chinook Salmon Populations and Recommendations for Future Research. Prepared for the AYK Sustainable Salmon Initiative (Anchorage, AK). v+70 pp.
- Shotwell, S.K., M.D. Adkison. 2004. Estimating indices of abundance and escapement of Pacific salmon for data-limited situations. *Transactions of the American Fisheries Society* 133(3):538-558.
- Su, Y. and M. Yajima. 2015. R2jags: Using R to run "JAGS". R package version 0.5.6. <http://CRAN.R-project.org/package=R2jags>.
- Walters, C.J. 1985. Bias in the estimation of functional relationships from time-series data. *Canadian Journal of Fisheries and Aquatic Sciences*. 44(1): 147-149.
- Walters, C.J., and D. Ludwig. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. *Canadian Journal of Fisheries and Aquatic Sciences*. 38(6): 704-710.

Walters, C.J. and S.J.D. Martell. 2004. Problems in the assessment of stock-recruitment relationships. Ch. 7. Pages 151-178 *in* Fisheries ecology and management. Princeton University Press, New Jersey.

Table 1. Escapement indices on the tributaries of the Kuskokwim River used by Bue et al.'s (2012) run reconstruction and the integrated model. Years operational are not necessarily consecutive.

<b>Project Type</b>	<b>Tributary</b>	<b>Years in Data</b>	<b>First Year in Data</b>
<b>Weirs</b>	Kwethluk	13	1992
	Tuluksak	18	1991
	George	16	1996
	Kogrukluksuk	29	1976
	Tatlawiksuk	15	1999
	Takotna	16	1996
<b>Aerial Surveys</b>	Kwethluk	11	1977
	Kisaralik	21	1978
	Tuluksak	12	1977
	Salmon (Aniak)	29	1978
	Kipchuk	22	1987
	Aniak	20	1981
	Holokuk	13	1993
	Oskawalik	19	1987
	Holitna	17	1976
	Cheeneetnuk	21	1977
	Gagaryah	19	1977
	Pitka	11	2001
	Bear	16	1976
	Salmon (Pitka)	26	1977

Table 2. Spawner-recruit parameters and biological reference points from three models. Traditional refers to a basic ordinary least squares linear regression model that does not allow for autocorrelated recruitment residuals. Values in parentheses are 95% credible intervals for the two Bayesian models (Post-Hoc and Integrated) and 95% bootstrapped confidence intervals for the traditional model, obtained by randomizing the regression residuals, adding them to the predicted values, and refitting the model as per Hamazaki et al. (2012).

Value	Traditional SR	Post-Hoc Model	Integrated Model
$\alpha$	7.29 (5.62-9.34)	6.03 (2.39-11.74)	6.33 (2.45-12.36)
$\beta$	9.07E-6 (7.70E-6-1.04E-5)	8.88E-6 (6.39E-6-1.14E-5)	8.91E-6 (6.33E-6-1.17E-5)
$\sigma_R$	0.29 (0.20-0.36)	0.24 (0.16-0.35)	0.23 (0.15-0.34)
$\phi$	—	0.81 (0.49-0.98)	0.81 (0.50-0.98)
$D$	—	78.59 (42.50-139.80)	75.28 (42.25-134.23)
$\pi_1$	—	0.19 (0.17-0.22)	0.19 (0.17-0.22)
$\pi_2$	—	0.39 (0.36-0.42)	0.39 (0.36-0.42)
$\pi_3$	—	0.39 (0.36-0.42)	0.39 (0.35-0.42)
$\pi_4$	—	0.03 (0.02-0.04)	0.03 (0.02-0.05)
$S_{MSY}$	80,041 (73,829-87,551)	76,985 (56,506-102,696)	77,944 (56,232-104,722)
$S_{MAX}$	110,458 (96,546-129,888)	115,167 (87,483-156,464)	115,069 (85,419-157,951)
$S_{eq}$	223,593 (210,329-237,917)	211,188 (139,156-324,362)	215,338 (139,760-332,249)

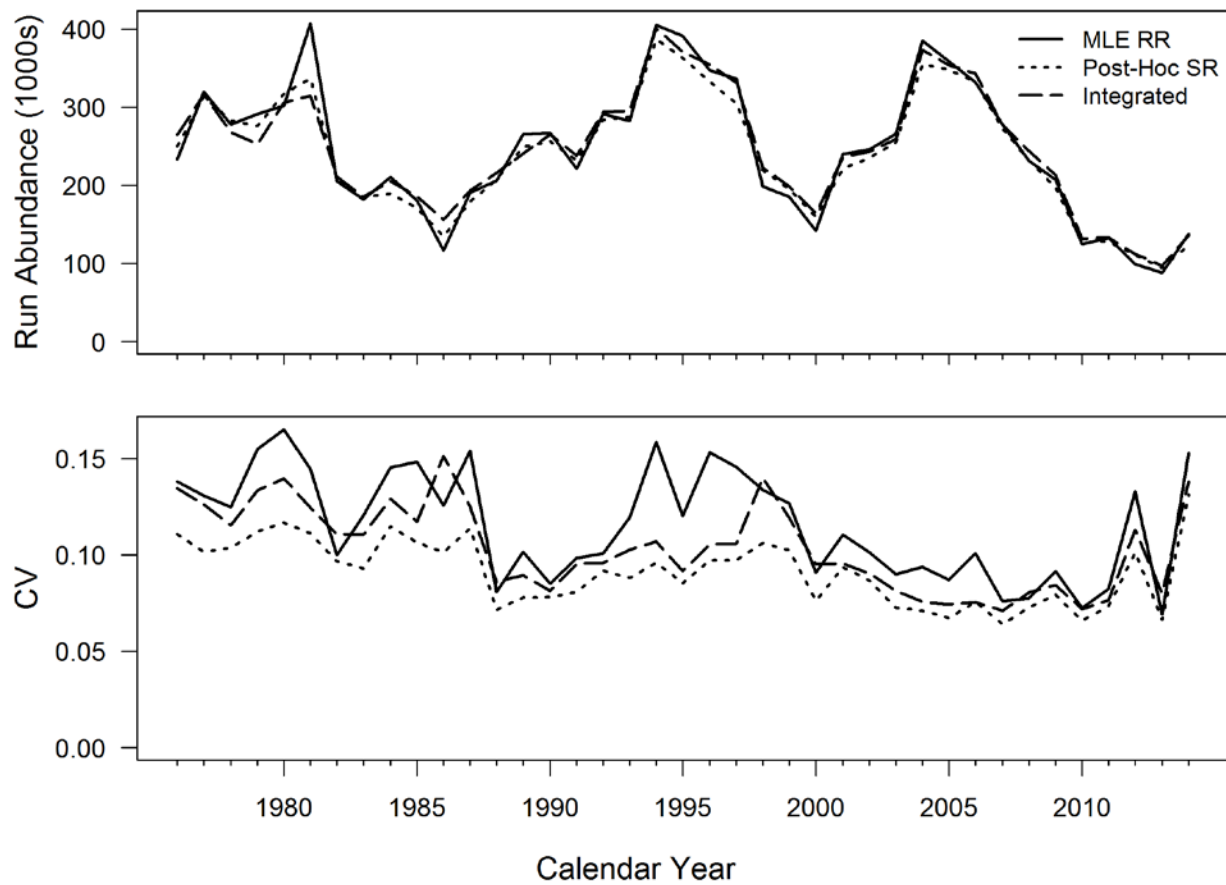


Figure 1. Estimated total run abundance (upper) and corresponding CVs (lower) from three models: the Bue et al. (2012) run reconstruction (MLE RR), the Hamazaki et al. (2012) post-hoc spawner-recruit analysis (Post-Hoc SR), and the integrated model.

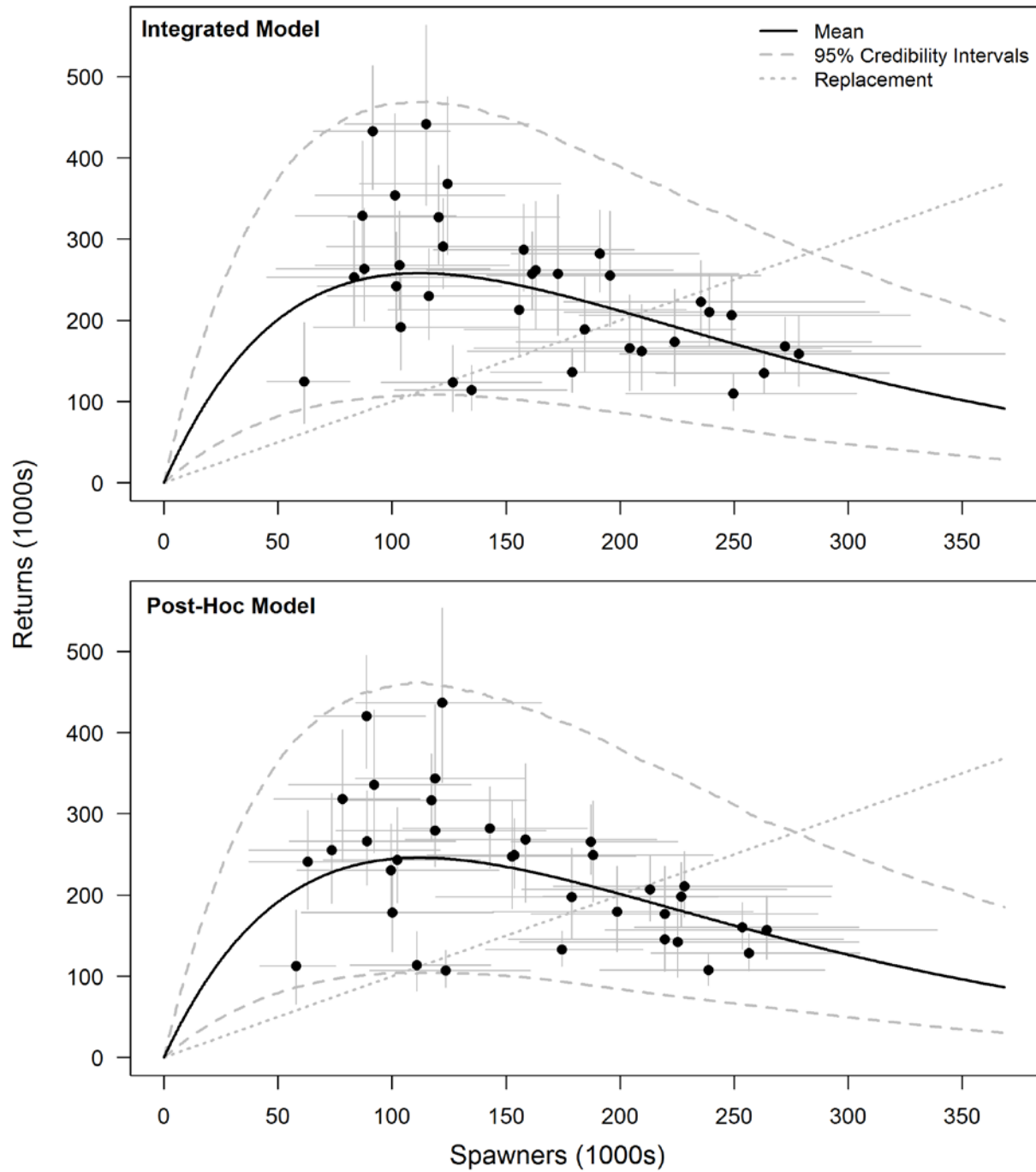


Figure 2. Spawner-recruit relationships from the integrated (upper) and post-hoc (lower) models. Error bars on points represent the 95% Bayesian credibility intervals on escapement-recruitment pairs and gray curves are 95% Bayesian credibility intervals for predicted recruitment given stock size. The dashed line represents 1:1 replacement.

# Proportionality Coefficients

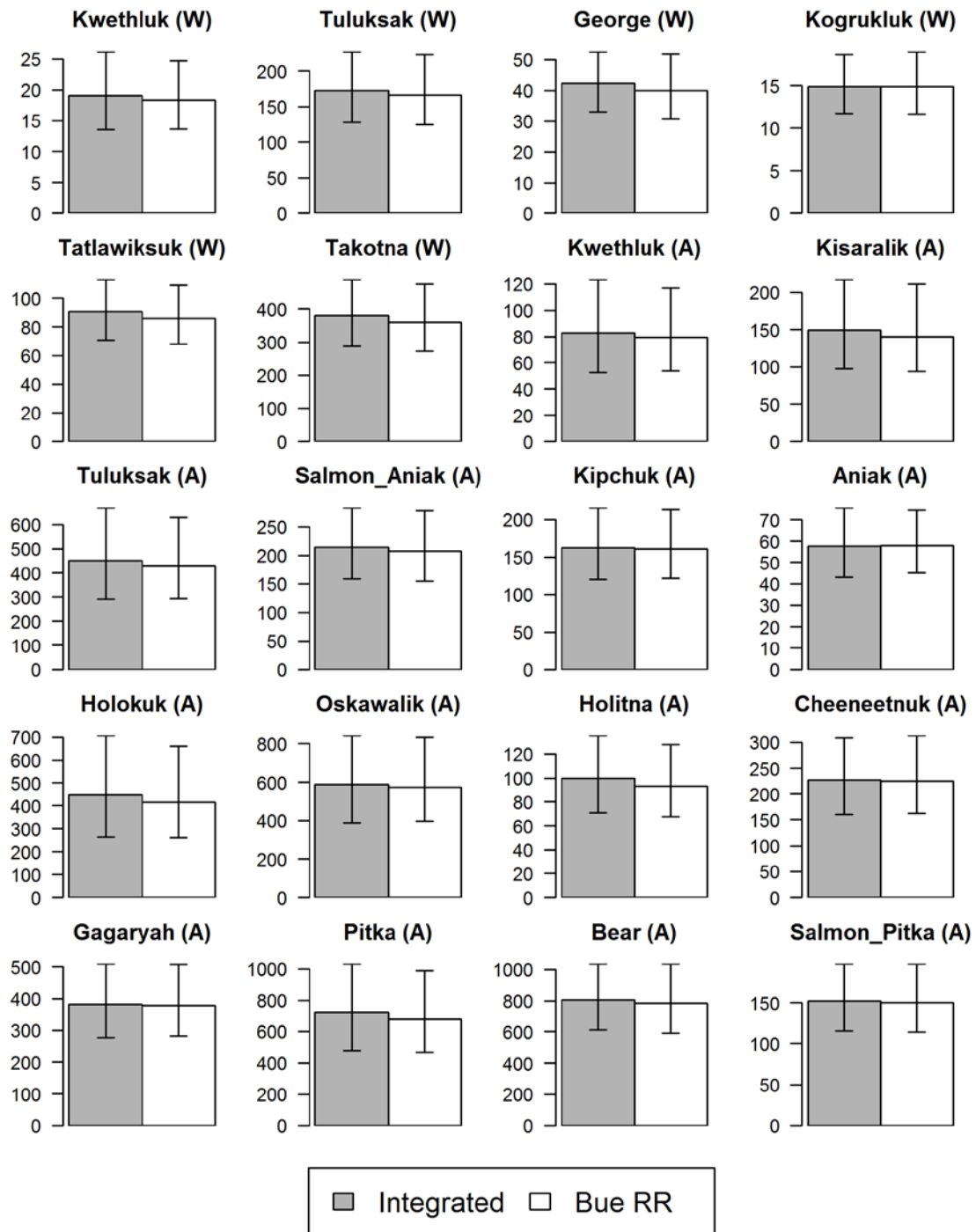


Figure 3. Proportionality coefficient estimates from the integrated model (gray bars) and the Bue et al. (2012) run reconstruction. (W) denotes a weir project and (A) denotes an aerial survey. Error bars are 95% Bayesian credibility intervals and 95% confidence intervals for the integrated and Bue et al. (2012) models, respectively.

# Over-dispersion Parameters

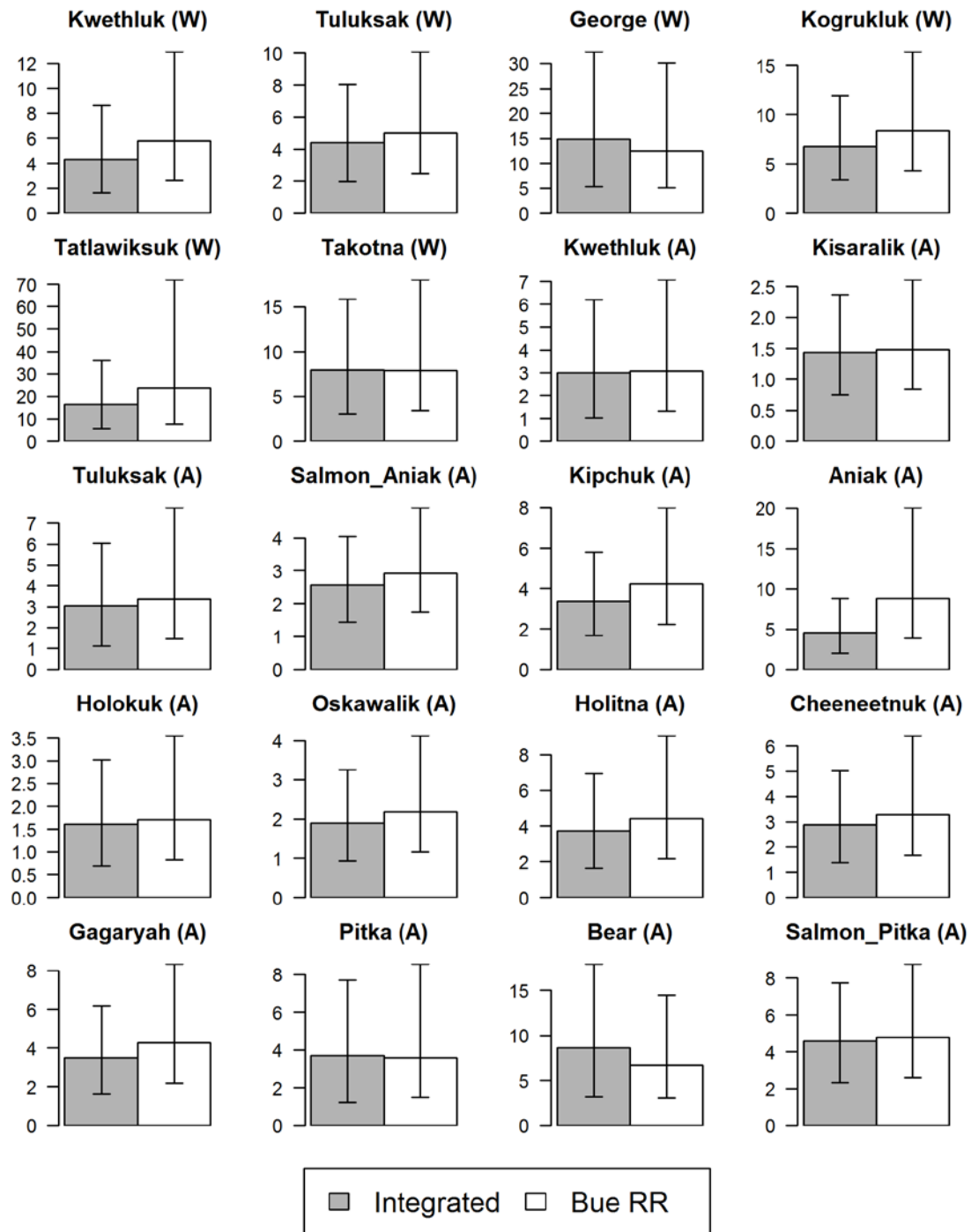


Figure 4. Over-dispersion parameter estimates from the integrated model (gray bars) and the Bue et al. (2012) run reconstruction (white bars). Error bars are 95% Bayesian credibility intervals and 95% confidence intervals for the integrated and Bue et al. (2012) models, respectively.



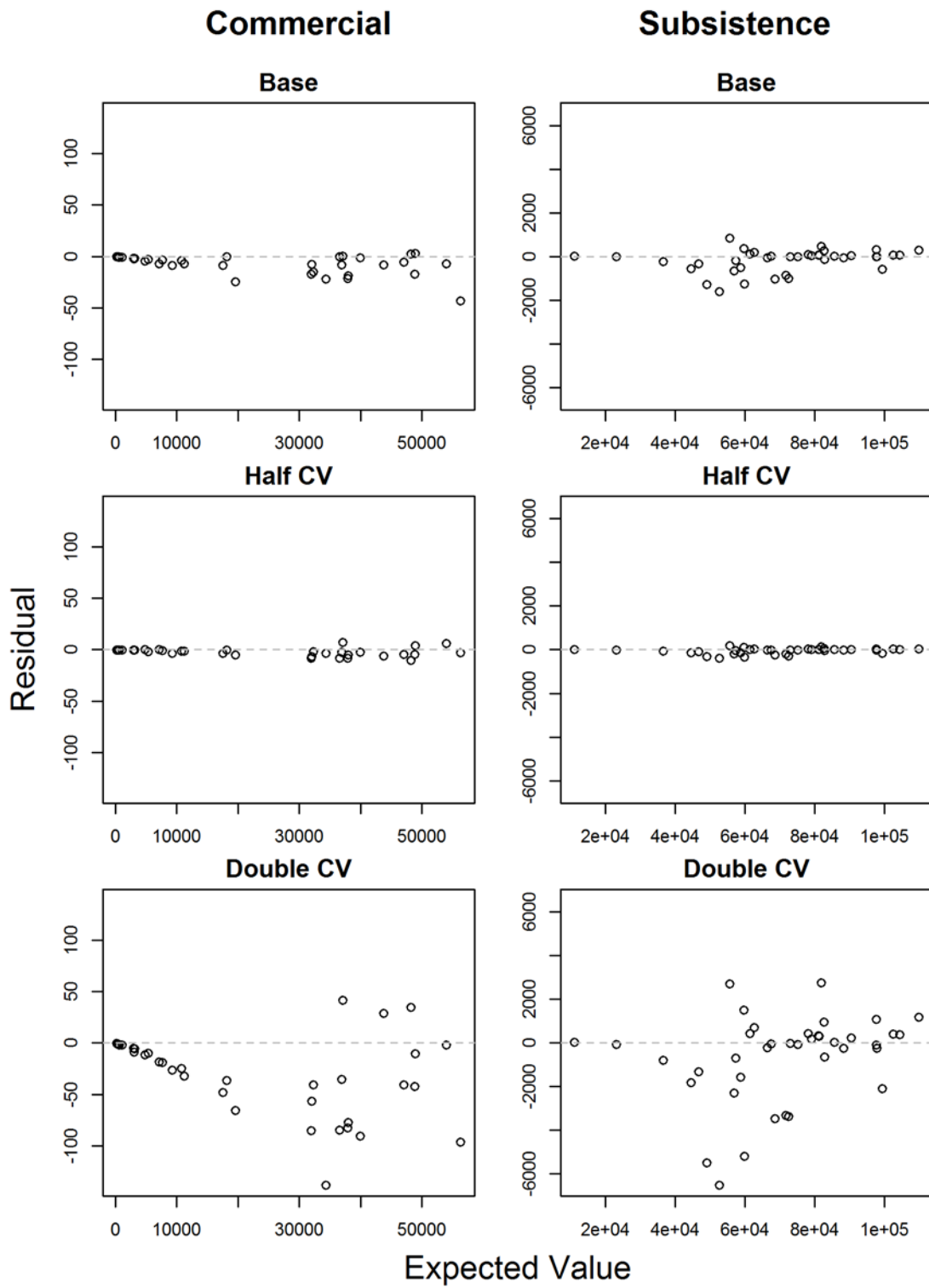


Figure 5. Harvest residuals from the integrated model under three observation CV sensitivity scenarios. Each scenario was conducted in isolation of the others.

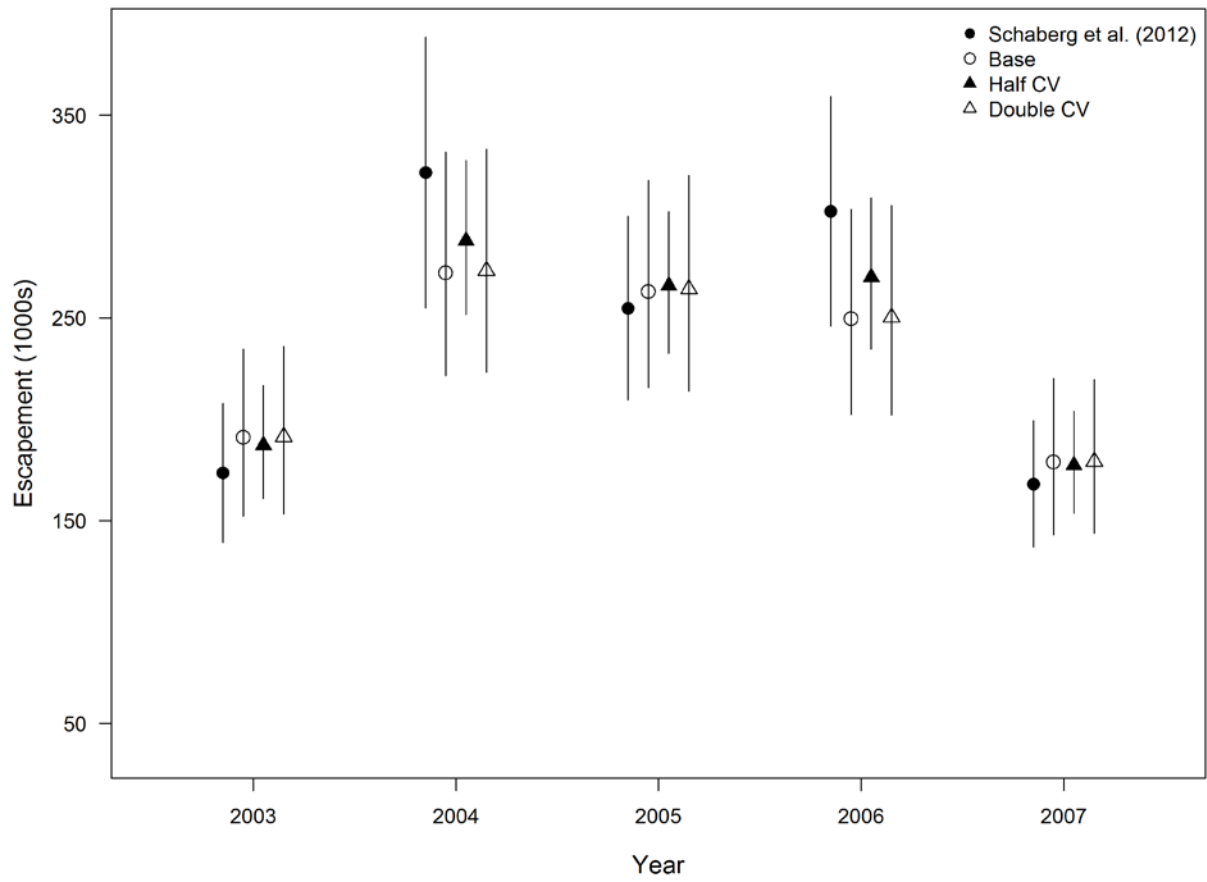


Figure 6. The effect of changing the observation variance on the mark-recapture estimates reported by Schaberg et al. (2012). Uncertainty on Schaberg et al. (2012) estimates are shown as  $\pm 1SD$ ; the Bayesian models are 95% credibility intervals. Only years with mark-recapture estimates are shown.

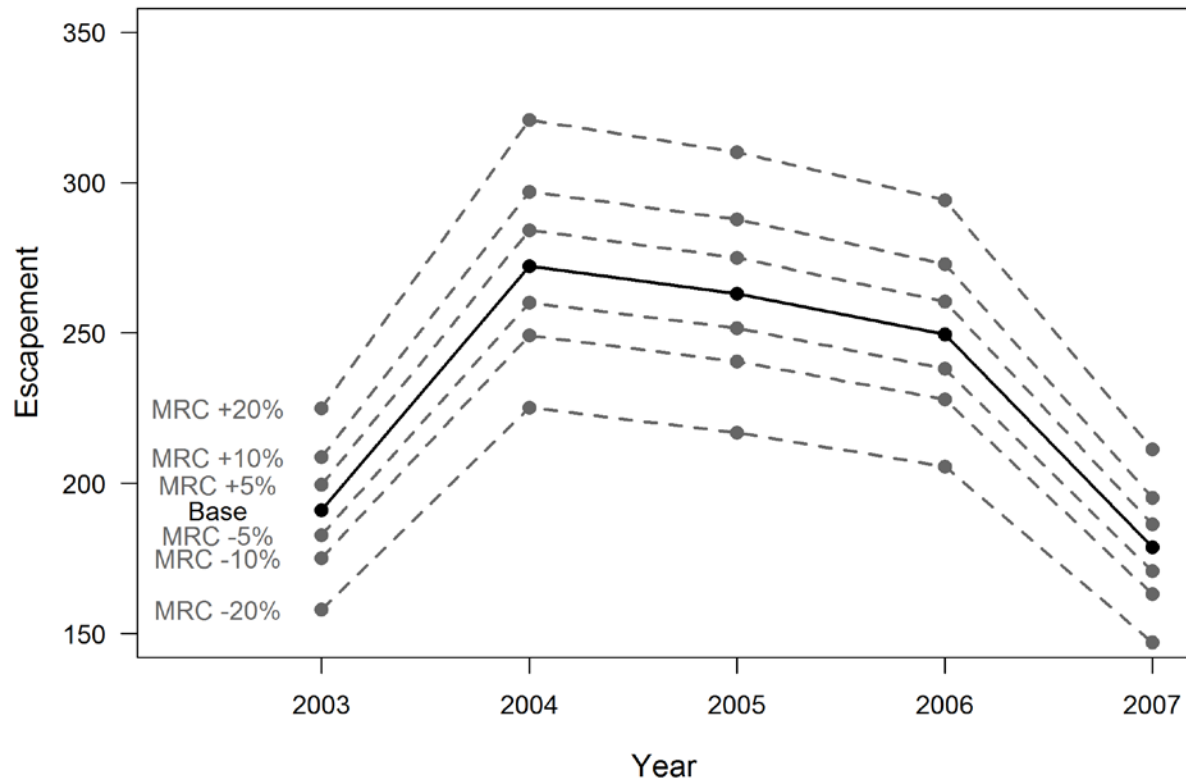


Figure 7. The effect of changing the mark-recapture estimate of drainage-wide escapement by  $\pm 5\%$ , 10%, and 20% on the expected escapement in those years.

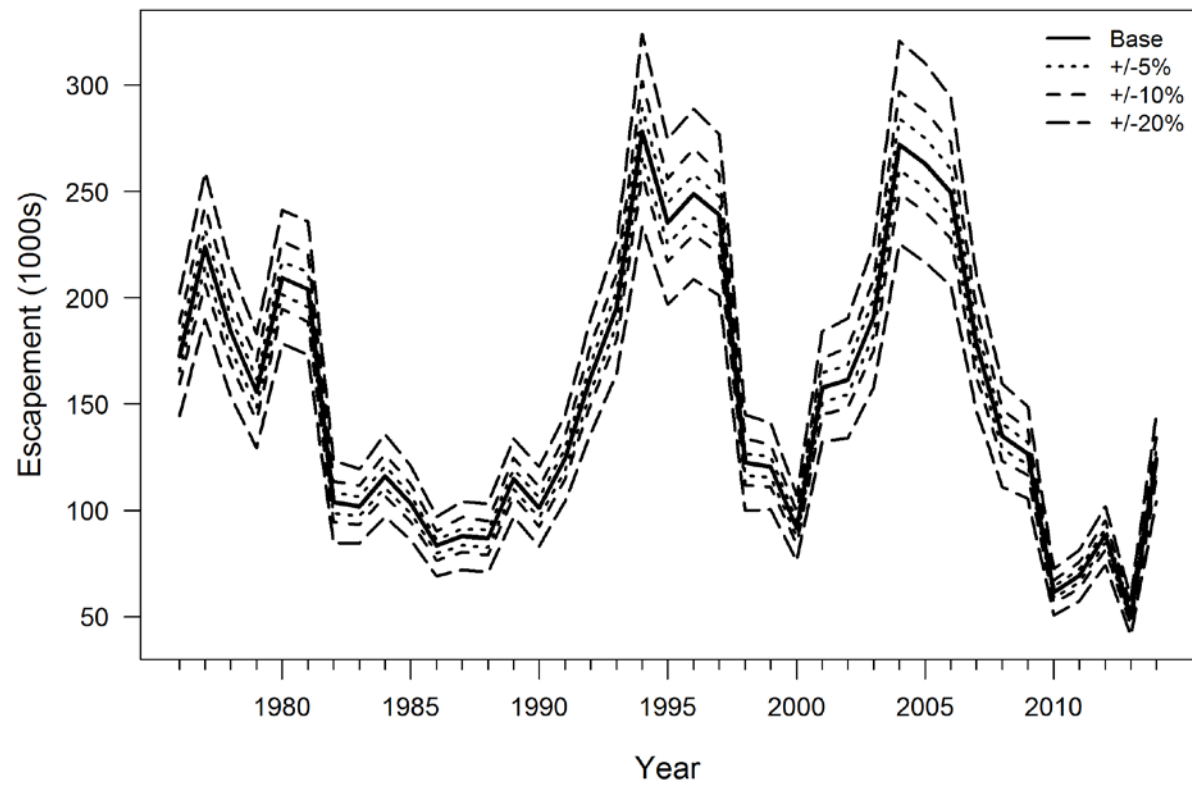


Figure 8. The effect of changing the mark-recapture estimate of drainage-wide escapement by  $\pm 5\%$ ,  $10\%$ , and  $20\%$  on the expected escapement in all years.

## APPENDIX A. INTEGRATED MODEL CODE

```
model{
  ### Spawner-Recruit with Autocorrelated lag-1 residuals: for years with spawner/recruit link
  lnalpha ~ dnorm(0,1.0E-2) %_%( 0,6)
  beta ~ dunif(0,10)
  for (y in (A+a.min):(Y+A-1)) {
    log.R[y] ~ dnorm(log.R.mean2.a[y], tau.white)
    R[y] <- exp(log.R[y])
    log.R.mean1.a[y] <- lnalpha + log(S[y-a.max]) - beta * S[y-a.max]
    log.resid.a[y] <- log(R[y]) - log.R.mean1.a[y]
    ### RPS: Return per spawner
    RPS.a[y] <- R[y]/S[y-a.max]
  }
  log.R.mean2.a[A+a.min] <- log.R.mean1.a[A+a.min] + phi * log.resid.0
  for (y in (A+a.min+1):(Y+A-1)) {
    log.R.mean2.a[y] <- log.R.mean1.a[y] + phi * log.resid.a[y-1]
  }

  ### monitoring quantities that don't start at y = 1; JAGS quirk
  log.resid <- log.resid.a[(A+a.min):(Y+A-1)]
  RPS <- RPS.a[(A+a.min):(Y+A-1)]
  log.R.mean1 <- log.R.mean1.a[(A+a.min):(Y+A-1)]
  log.R.mean2 <- log.R.mean2.a[(A+a.min):(Y+A-1)]

  ### Priors for SR portion
  phi ~ dunif(-1,0.99)
  tau.white ~ dgamma(0.01,0.01)
  log.resid.0 ~ dnorm(0,tau.red)
  tau.red <- tau.white * (1-phi*phi)
  sigma.white <- 1 / sqrt(tau.white)
  sigma.red <- 1 / sqrt(tau.red)
  alpha <- exp(lnalpha)

  ### Biological Reference Points
  lnalpha.c <- lnalpha + (sigma.white * sigma.white / 2 / (1 - phi * phi))
  S.max <- 1 / beta
  S.eq <- lnalpha.c * S.max
  S.msy <- S.eq * (0.5 - 0.07 * lnalpha.c)

  ### Brood year returns without SR link; drawn from a common lognormal dist
  mean.log.R0 ~ dnorm(0,1.0E-4)
  tau.R0 ~ dgamma(0.1,0.1)
  R.0 <- exp(mean.log.R0)
  sigma.R0 <- 1 / sqrt(tau.R0)
  for (y in 1:a.max) {
    log.R[y] ~ dnorm(mean.log.R0,tau.R0)
    R[y] <- exp(log.R[y])
  }
}
```

```

#### Generate Y+A-1=41 maturity schedules, one per brood year
#### Proportion mature (returning to spawn) at age modeled as drawn from a dirichlet distribution
across brood years
D.scale ~ dunif(0,1)
D.sum <- 1 / (D.scale * D.scale)
prob[1] ~ dbeta(1,1)
prob[2] ~ dbeta(1,1)
prob[3] ~ dbeta(1,1)
pi[1]<- prob[1]
pi[2] <- prob[2] * (1 - pi[1])
pi[3] <- prob[3] * (1 - pi[1] - pi[2])
pi[4] <- 1 - pi[1] - pi[2] - pi[3]

for (a in 1:A) {
  gamma[a] <- D.sum * pi[a]
  for (y in 1:(Y+A-1)) {
    g[y,a] ~ dgamma(gamma[a],1.0)
    p[y,a] <- g[y,a]/sum(g[y,])
  }
}

#### Calculate the numbers at age matrix as brood year recruits at age*proportion that matured that
year
for (t in 1:Y) {
  for (a in 1:A) {
    N.ta[t,a] <- R[t+A-a] * p[t+A-a,a]
  }
}

#### Calculate escapement as N-Harvest, after modeling harvest in both commercial and subsistence
fisheries
for (t in 1:Y) {
  N[t] <- sum(N.ta[t,1:A])
  S[t] <- N[t] * (1-U[t])
  log.S[t] <- log(S[t])

  #taus for harvest as known data
  sigma.H.com[t] <- sqrt(log(pow(cv.H.com[t],2)+1))
  tau.log.H.com[t] <- 1/pow(sigma.H.com[t],2)
  sigma.H.sub[t] <- sqrt(log(pow(cv.H.sub[t],2)+1))
  tau.log.H.sub[t] <- 1/pow(sigma.H.sub[t],2)

  U[t] ~ dunif(0.0001, 0.9999)
  p.com[t] ~ dunif(0.0001, 0.9999)

  u.com[t] <- U[t]*p.com[t]
  u.sub[t] <- U[t]*(1-p.com[t])

```

```

pred.com.catch[t] <- u.com[t]*N[t]
pred.sub.catch[t] <- u.sub[t]*N[t]

log.pred.com.catch[t] <- log(pred.com.catch[t])
log.pred.sub.catch[t] <- log(pred.sub.catch[t])

com.cat[t] ~ dlnorm(log.pred.com.catch[t], tau.log.H.com[t])
sub.cat[t] ~ dlnorm(log.pred.sub.catch[t], tau.log.H.sub[t])

### Multinomial scale sampling on total annual return N
for (a in 1:A) {
  q[t,a] <- N.ta[t,a] / N[t]
}
x[t,1:A] ~ dmulti(q[t,],n[t])
}

### ASSESSMENT SUBMODEL ###

### Escapement Indices
for (j in 1:20) {
  r[j] ~ dgamma(0.001, 0.001)
  k[j] ~ dnorm(10, 1E-8)
}

#Negative binomial likelihood on weirs and aerial surveys
for (i in 1:num.index) {
  est.esc[i] <- S[esc.year[i]] / k[trib[i]]
  p.esc[i] <- r[trib[i]] / (est.esc[i] + r[trib[i]])
  index[i] ~ dnegbin(p.esc[i], r[trib[i]])
}

### Weekly commercial CPUE
tau.cat~dgamma(0.001, 0.001)
for (q in 1:3) {
  ln.Q[q] ~ dnorm(0, 1E-10)
  Q[q] <- exp(ln.Q[q])
}
q.unr <- Q[1]
q.res <- Q[2]
q.mono <- Q[3]

### estimated commercial catch and fitting it to obs
for (i in 1:num.com) {
  est.c.catch[i] <- log(pp[i] * N[c.year[i]] * (1 - exp(-Q[gear[i]] * effort[i])))
  catch[i] ~ dlnorm(est.c.catch[i], tau.cat)
}

```

### in river likelihood. using total estimated escapement (mark recap+lower river expansion). Keeps model internally consistent.

```
for (i in 28:32) {  
  inr.s[i] ~ dlnorm(log.S[i], tau.inr.s[i])  
}  
}
```

### END MODEL CODE ###



## CHAPTER 3

### DESCRIPTION OF THE KUSKOKWIM RIVER CHINOOK SALMON RUN RECONSTRUCTION AND AN INVESTIGATION OF DATA WEIGHTING: A REPORT TO THE KUSKOKWIM RIVER SALMON MANAGEMENT WORKING GROUP

Ben Staton and Matt Catalano, Auburn University; Lew Coggins, US Fish and Wildlife Service; Bill Bechtol, Association of Village Council Presidents, Bechtol Research; Dan Gwinn, Biometric Research, LLC

#### ABSTRACT

- **This document is the result of a suite of independent analyses investigating the consequences of different ways of weighting weir and aerial survey data in the run reconstruction model.** All of the authors have statistical training and experience with models like the run reconstruction model and voluntarily took on this task.
- **The run reconstruction model is a tool to estimate annual Chinook salmon run abundance in the Kuskokwim River since 1976.** The model uses observed data from weirs, aerial surveys, harvest, and total abundance estimates based on a mark-recapture study to estimate a historical run abundance time series. We conclude that the model structure is generally consistent with broadly applied and accepted stock assessment modeling approaches.
- **We investigated the results of 5 different models.** Four of the models were different approaches to weighting weirs and aerial surveys, and the fifth incorporated the 2014 mark-recapture abundance estimate:
  - ADF&G base model: individual weights are estimated for each escapement monitoring project.
  - Pooled weights model: a common weight is estimated for all weir projects and one common weight is estimated for all aerial survey projects.
  - No 2014 weirs model: 2014 weir counts are excluded from the ADF&G base model.
  - No 2014 aerial surveys model: 2014 aerial surveys are excluded from the ADF&G base model
  - With MRC model: the ADF&G model including the 2014 mark-recapture estimate.
- **All of the models resulted in the same major conclusion that the 2014 run was lower than average and the drainage-wide escapement goal was met.**
- **The estimated weights for weirs were higher than for the aerial surveys,** which means that, all else being equal, the model's predictions will more closely match the weir data overall.
- **Retrospective analyses showed that the ADF&G base model run estimates varied more than the pooled weights model as more data were added,** but neither model exhibited worrisome retrospective patterns.
- **The ADF&G base model exhibited an undesirable tendency to perfectly fit to one of the weir time series in some years.** This behavior is extreme and warrants further investigation and consideration of alternate models because it suggests that in those cases the model treated one of the weirs as a perfect indicator of escapement. This extreme behavior occurred in two out of eight years in the retrospective analysis: 2007 and 2010. Both of these years were not years in which the model was used to make management decisions.
- **Future modeling efforts should continue to investigate alternate weighting schemes and model structures.** The pooled weights model we constructed represents a sensible first step in this direction, but there are likely many other potential model structures that could be considered.

## INTRODUCTION

During the Kuskokwim River Salmon Management Working Group session of the Kuskokwim Area Interagency Meeting held in Bethel, AK in March 2015 there were several important inquiries regarding the run reconstruction model. These questions included:

- How does it work?
- Which data have the most influence (i.e., weight)?
- Are there other ways of weighting the data components, and have these been considered?

The latter two questions have come up before with biologists and biometricians and were discussed prior to the Interagency Meeting. For example, some of the sensitivity analyses presented in this report have been previously conducted by the model developers and others using older versions of the model.

There are two primary objectives of this report. First is to answer the above questions and make clear why this issue of weighting has recently been raised. The second objective is to elaborate on issues related to data weighting by conducting several analyses to demonstrate how the data weighting works. This report will summarize and present the major findings from these analyses for working group and agency consideration.

The run reconstruction model is a tool that biologists and managers use to estimate how many Chinook salmon have been coming back to the river annually since the 1970s. The model was developed by Brian Bue and other assessment scientists and was completed in 2012 (Bue et al. 2012). The original report can be accessed at: <http://www.adfg.alaska.gov/FedAidpdfs/FDS12-49>. There have been several small changes to that model in 2014, as noted in this memorandum from Hamazaki and Liller (2015): <http://www.adfg.alaska.gov/static/fishing/PDFs/commercial/kuskokwim/2014RunReconMemo.pdf>.

### ***How does the Kuskokwim River Chinook Salmon Run Reconstruction work?***

To estimate run abundance, the model needs two major components. First, it needs data on how many fish are in the river each year. These data come primarily from weirs, aerial surveys, harvest estimates, and mark-recapture studies. Second, it needs some way to relate those data to actual numbers of fish, since we can't possibly count all the fish each year. The model represents a rational approach to combining information from all of the data types and sources to estimate how many Chinook salmon have been returning annually to the Kuskokwim River. The general model formulation and its assumptions are broadly consistent with commonly applied stock assessment modeling practices.

#### ***Weirs and Aerial Surveys***

The weirs and aerial surveys are treated as indices of escapement, which means that they are not complete counts, but instead represent relative changes in escapement over time in each tributary. The model predicts how many fish should be observed at each weir or in each aerial survey by assuming escapement to each tributary is proportional to the drainage-wide escapement. These proportions are assumed to be constant over time and are estimated by the model (i.e., the data is used to inform the model about the tributary proportions). The model then compares how many fish were actually counted in each tributary to these predictions. If the predictions differ greatly from the data, the model adjusts the run abundance and proportions so that the predicted counts are closer to the actual counts that were observed.

#### ***Mark-Recapture Studies***

How do we deal with tributaries that do not have weirs or aerial surveys? The majority of Chinook salmon producing tributaries are not monitored and if we want to know the total escapement each year, we need a way to expand the weir and aerial survey counts to the whole drainage. A mark-recapture study can provide information for such an approach. The mark-recapture study was an

experiment in which biologists captured and tagged fish at Kalskag, then “recaptured” the tagged fish upstream at weirs later in the run. By comparing the number of marked versus unmarked fish observed at weirs, the biologists were able to estimate how many fish passed upstream of the tagging site. The methods for the 2003-2007 mark-recapture study and the resulting estimates are provided in Schaberg et al. (2012). The mark recapture estimates, along with observed harvests and escapement estimates based on drainage-area for streams downstream of the tagging site, are used in the model to scale up the weir and aerial survey counts to a total drainage-wide abundance of fish for a given year. One way to think about this is that the weirs and aerial surveys inform the ups and downs of how many fish there are from year to year, and the mark-recapture estimates specify the actual total number of fish, which allows for the model to estimate how many fish there are in the whole drainage without actually counting them all every year.

### *Commercial Catch and Effort*

There is also a component that uses the historical catch and effort from the commercial fishery. We all know that when there are lots of fish in the river, commercial fishers don’t need to spend as much time catching fish as they would if there were fewer fish in the river. In years in which many fish were caught with little effort, the model predicts that there were many fish in the river that year. In recent years, this component has had minimal impact on the model since there is no directed commercial fishery for Chinook salmon. The component is still included because there are data from the 1970s and 1980s when the commercial component was more extensive that can provide information on total run abundance in those years.

### *End Result*

Of course there are statistical complexities surrounding how the model adjusts its predictions to match the data we observed. Essentially, the model adjusts the run abundance estimates by trial-and-error until it minimizes the differences between the model predictions and the observed data across all data sources. The resulting run abundance estimates are those that make the data we observed for the past approximately 40 years most likely to have occurred. Another way to think about this is that the model’s run abundance estimates are those that best explain all of our observations.

### ***Which data are most influential in the model?***

This question has to deal with an issue known as “data weighting”. Data weighting means that the model is influenced by some data more than others. Another way to say this is that when two datasets disagree (i.e., one says there are lots of fish, another suggests there are few fish), which dataset does the model follow more closely?

If the model weighted all datasets the same, for example weirs and aerial surveys, then the model would tend towards the average abundance suggested by all of them. However, if the weirs are weighted more heavily, then the model will follow the weir patterns in escapement trends more closely than patterns suggested by the aerial surveys.

Ideally, weirs and aerial surveys would agree perfectly all the time. However, this is not the case since the two methods of counting fish are very different. An aerial survey is a pure index of escapement, that is, its counted value is related to the true escapement in that tributary, but it is not an attempt to be a complete count. Weirs are nearly complete counts, however there are factors that prevent them from being entirely complete counts (e.g., it was not possible for the weir to be functional during the entirety of the run). Because of these missed counts, weirs are also indices, but they are more representative of the true number of fish escaping to that tributary than are aerial surveys. This might suggest that we should have more trust in what the weirs tell us than what the aerial surveys tell us, or in other words, place more weight on weirs.

As the model is currently formulated, it is “self-weighting.” That is, the model uses the data to determine which data sources provide the most consistent information about total run size and places the most weight on them. When working properly, this is very desirable since it removes the subjectivity of biologists, managers, and stakeholders deciding on which data are best and by how much. However, it can also lead the model to behave counterintuitively. For example, it is possible for the model to heavily weight a very small aerial survey that does not really reflect what is happening in the whole drainage, just because the data collected in that tributary are less variable through time than other tributaries. This is problematic when data sources are contradictory. In this case, one must decide which data to most believe: the consistent but small aerial survey or a larger one that we think is more representative of the total escapement? These types of scenarios inevitably must be confronted in nearly all stock assessment models.

### ***Why is data weighting an issue in 2014?***

In 2014, there was substantial disagreement among the various assessment projects relative to the implications of drainage-wide escapement size. In particular, the Kogrukluk, Kwethluk, and Tuluksak River weirs, along with the Cheeneetuk and Holokuk River aerial surveys, indicated that the overall run was much smaller than suggested by the Pitka Fork of the Salmon, Kipchuk, and Aniak River aerial surveys. The disagreement among the estimates based on these individual assessment projects can be easily seen in Figure 3 of the Hamazaki and Liller (2015) memo. While disagreement among individual assessment projects is not unexpected, the level of disagreement among the 2014 assessment projects was relatively severe and is partially responsible for the high uncertainty in the 2014 estimate. The coefficient of variation (a relative measure of estimator uncertainty) of the 2014 run abundance estimate was 15% and was the largest coefficient of variation since 1996. Additionally, the model potentially down-weighted the information from the Kogrukluk and Kwethluk weirs more than was appropriate based on the relative size of these systems and the greater accuracy of weir data relative to aerial survey data (Table 1). In contrast, the model possibly placed too much weight on the relatively small Tatlawiksuk River information and some of the higher, but potentially less accurate, aerial survey counts. While it is impossible to know what the perfect data weighting scheme should be, this model behavior is troubling if we assume that the weirs on the larger systems are a more reliable indicator of drainage-wide escapement trends. This issue prompted us to investigate the behavior of how the different data sources are weighted in the model, particularly for weirs and aerial surveys.

### ***How did we investigate the weighting behavior?***

We have used two approaches to investigate this question, as outlined below. These approaches are standard in stock assessments to diagnose model behavior and are often required before considering an assessment complete.

#### ***Sensitivity Analyses***

A sensitivity analysis is a way of testing what impact a component of the model has on the output (e.g., the abundance estimates). Essentially, we change something about the model formulation and see how much the output changes. If the output changes a lot compared with the original model, then we infer that the component we changed has a large influence in the model. This exercise is important to allow us to see what would happen if we make different assumptions about how the model works. The idea is that when we cannot be perfectly certain of the best way to formulate the model, we can try several different variations of the model to see if they paint a different picture regarding run abundance. With regard to the question of how the model should weight the aerial survey data versus the weirs, for example, we can force the model to weight weirs and aerial surveys in a variety of different ways, and see how much it matters in terms of the run abundance estimates. This

approach allows us to explore a range of possible models rather than having to decide on a single model when we lack an objective way to decide.

### *Retrospective Analyses*

A retrospective analysis is an exercise in which we go back in time, so to speak. We pretend that we were back in 2007 and ignore all of the data collected between then and now and run the model. Then we include data from 2008 and run the model, then 2009, 2010, and so on up until the present. By doing this, we can look at how the model results change as we accumulate more and more data. If there are conflicts in the data, or the data weightings change over time then we might expect the model estimates to change more as we accumulate data. We are particularly interested in how the model chooses which data sources to weight over time and how much the weightings change when we add new data. This can help shed light on how the model selects which data are most important (i.e., aerial surveys versus weirs), which could indicate whether there are potential errors in the model estimates.

## **METHODS**

### *Sensitivity Analyses*

We formulated five models that differed mainly in how weirs and aerial surveys were weighted and also included a model that incorporated the preliminary 2014 mark-recapture drainage-wide abundance estimate. Table 2 describes the different model structures and combinations included in the sensitivity analysis.

Each of these scenarios represents a separate run of the model and each one was chosen strategically to illustrate how the weirs and aerial surveys impacted the model estimates. The first two scenarios were different model structures: the first allowed each individual project (i.e., individual weirs and aerial surveys) to have a different weight whereas the second forced all weirs to have the same weight and all aerial surveys to have the same weight. This approach was useful because if the model tended to track a particular aerial survey or weir count very closely at the expense of the others for no rational reason, then it may make sense to consider another weighting scheme such as pooling the weights by project type (i.e., weirs versus aerial surveys). This approach would prevent the model from picking one particular survey and following it too closely. We call this the “pooled weights model” for comparison with the ADF&G base model, which we call the “separate weights model.”

A more extreme data weighting scheme that is often employed in sensitivity analyses is to assign zero weight to a particular dataset to assess the effect of its complete removal. With this approach, we can bound the extremes suggested by each type of data. We used this approach by removing (i.e., assigning a zero weight) the 2014 weir data and look at how much larger the aerial surveys suggest the 2014 run was. We then did the same for the aerial surveys. Finally, the last model scenario included the 2014 mark-recapture estimate, which is an independent estimate of drainage-wide abundance. Its inclusion allowed us to investigate which data were consistent with this estimate. Please note that the 2014 mark-recapture drainage-wide abundance estimate provided by ADF&G is preliminary and could change depending on the outcome of continued internal review by staff biometricians.

### *Retrospective Analysis*

We conducted two separate retrospective analyses: one was conducted using the separated weights model and one used the pooled weight model. We conducted a separate retrospective analysis for each of these models to investigate the influence of the model formulation (i.e., how it dealt with weights: separate or pooled) on (1) which projects had the largest estimated weight overall and (2) whether or how much the weights changed as more data accumulated. We started the retrospective analysis in 2007 as this was the last year of the original mark-recapture studies (Schaberg et al. 2012).

Starting after all of the original mark-recapture studies were completed allowed for all of the scaling information provided by those projects to be included in all of the retrospective model runs. Thus, any differences in the model estimates across subsequent model runs (i.e., as additional years of data are included) could be attributed mostly to how the model dealt with weir and aerial survey counts in those years.

All models were run using the model code provided in Hamazaki and Liller (2015). We made minor changes when warranted by the particular model scenario (e.g., removing particular data, pooling weights, etc.).

## RESULTS

When conducting these analyses, we discovered that we could not exactly reproduce the estimates reported by Hamazaki and Liller (2015). A detailed investigation of this phenomenon indicated that the parameter bounds in the ADF&G model (separate weights model) were too narrow to allow the model to freely estimate all of the parameters. Parameter bounds are simply upper and lower caps on parameter values (e.g., run abundance). They are commonly applied in stock assessments to prevent the models from inadvertently choosing implausible parameter values during the statistical trial-and-error estimation process. In the Kuskokwim Chinook model, the parameter bounds for the weights were not high enough such that one of the model weight estimates was pushed up against the bound, which means that the true estimate was likely outside the bound and suggests that the model was overly constrained by the bound. That is, the model was not allowed to estimate parameter values that fully satisfied the data. For our analyses, we loosened this constraint, which resulted in changes of 1-2% for the 2014 run estimate. For this reason, the estimates in this document do not exactly match those presented in the memo by Hamazaki and Liller (2015), but the difference is small.

Additionally, when comparing our estimates to those presented in Hamazaki and Liller (2015), we discovered an error in how the estimates were reported. The order of the over-dispersion parameters (or weights, as they are referred to in this document) for the aerial surveys was inadvertently shuffled in the memo (see Table 2 in Hamazaki and Liller [2015]). For this reason, the weights that we report will differ from the memo. We have informed ADF&G of these issues.

### ***Sensitivity Analyses***

The primary measure we used to compare the different scenarios was the 2014 run abundance estimated by each model, as shown in Table 3 and Figure 1, below.

The model structure clearly influenced the 2014 run estimate, but relative to the larger picture, these differences were small. The error bars are 95% confidence intervals, which is an estimate of uncertainty and represents how confident we are in the abundance estimate. The 95% confidence intervals on escapement were calculated by subtracting the total harvest estimate from the upper and lower bounds of the total run confidence interval, which ignores the uncertainty in the 2014 harvest estimate (which is considered to be minimal in this model). When we pooled the weights across project type (one weight for all the weirs and one weight for all the aerial surveys), the estimate was 15% smaller than when we estimated a different weight for each project. This finding resulted from the pooled model placing more weight on weirs that suggested a smaller run than did the aerial surveys in 2014. This model behavior was further demonstrated when we examined the next two points in Figure 1. When we excluded the 2014 weir data, the model predicted a higher 2014 run abundance and when we totally excluded the 2014 aerial survey data, the model predicted a lower 2014 run abundance. Finally, the model that considered the preliminary abundance estimate from the 2014 mark-recapture project produced the lowest estimated 2014 run abundance. Considering estimates across this set of sensitivity analyses provided a relatively complete characterization of the uncertainty in the run reconstruction among a wide set of options related to model structure and data considered. The

important thing to keep in mind is that no matter which scenario you look at, the main conclusion is the same: the 2014 run was smaller than average and the drainage-wide escapement goal was met.

However, these findings pertain just to 2014. Since the model reconstructed the run abundance time series going back to 1976, it is important to look at how changing the model structure affected the whole time series. Figure 2 shows what happened when we compared the separated weights model (which is the ADF&G base model) to the pooled weights model:

The run estimates from the pooled weights model fell within the 95% confidence intervals of the separated weights model nearly every year, with the exception of 1988. Based on this finding, we can conclude that the run estimates from these two models are similar. Several of the estimates from the other scenarios fall within this confidence interval as well. The only scenario that did not result in a 2014 run estimate within the confidence intervals of the ADF&G base scenario model (separated weights model) was the model with the 2014 mark recapture estimate included (solid triangle), but it was very close to being within the interval.

### ***Retrospective Analyses***

Figure 3 depicts how the weights for the two project types changed overtime as more data were added using the pooled weights model.

Weir data were weighted more heavily than aerial survey data in the pooled weights model. The estimated weights from this model were consistent as more data were added. In contrast, estimated data weightings from the separate weights model varied drastically for some datasets (larger numbers mean it had more influence on the model; note the different magnitude of the axes for certain projects, Figure 4).

In particular, the model estimated extremely large weights for the Kwethluk weir in 2007 and Tatlawiksuk weir in 2010, for reasons that are not obvious. We investigated this behavior further by plotting the model predictions versus the observed data points for those years. Figure 5 shows observed and predicted escapement counts for the retrospective model run using data through 2007 when the Kwethluk weir received an extremely large weight. Figure 6 shows the same information but for the retrospective model run using data through 2010 when the Tatlawiksuk weir received an extremely large weight. The line on each graph depicts a model prediction that is exactly the same as the observed data, so points that fall closer to the line indicate that the model followed that dataset very closely for that year. Figures 5 and 6 indicate that the model fit the Kwethluk weir and Tatlawiksuk weir data perfectly in 2007 and 2010, respectively, because all of the data points fall exactly on the line. We were unable to identify the reason that the model adjusted its estimates so that they perfectly fit the observed data for those weirs in those model runs.

The separated weights model behavior of choosing one dataset and following it precisely at the expense of all other information is extreme and abnormal, especially when the chosen dataset changed as more years of data were considered. One way to think about this model behavior is that in 2007 for example, the model treated the Kwethluk weir as a perfect representation of escapement for the entire drainage, which strains the credibility of the separated model. The extreme weighting behavior in the 2007 and 2010 retrospective model runs indicates that there is currently insufficient information in the data to allow consistent estimation of a stable and sensible weighting scheme under the separated weights model. The retrospective analysis revealed that if this model had been used to produce estimates following either the 2007 or 2010 sampling seasons, it is likely that the analyst would have chosen some other method that resulted in a more justifiable set of data source weights.

In addition to the weights under the pooled and separated model structures, we also investigated how the estimated run abundance time series changed under these scenarios as we added more years of data, as shown in Figure 7.

The run abundance time series from the separated weights model fluctuated more as we added data than did the pooled weights model. Although the overall trend was similar between models, the peak years in the early 1990s fluctuated by 10-20% (nearly 100,000 fish) between retrospective runs, whereas in the pooled model they fluctuated by 5-10% (~50,000 fish). Neither model demonstrated systematic retrospective bias that is commonly observed in many stock assessments. An example of a systematic bias would be if the abundance estimates consistently trended downward across the board with the inclusion of each additional year of data. Although both of these models behave reasonably well in this respect, the pool weights model clearly has less retrospective variation in the abundance estimates and does not suffer from the extreme tendency to perfectly fit one of the escapement surveys in some years.

## **DISCUSSION**

These analyses shed light on how the run reconstruction model assigns weights to individual datasets. In general, weirs were weighted more heavily than aerial surveys, which makes sense because we should expect the aerial surveys to have larger sampling error. When we look retrospectively, abundance estimates and data weights from the separated weights model fluctuated more as more data were added when compared to the pooled weights model. We have not yet figured out a way to quantify what a weight of about 9 for weirs versus a weight of about 3 for aerial surveys exactly means (how much more do weirs influence the model than aerial surveys?), but it is sufficient to say that an individual weir carries more weight in the model than does an aerial survey with an equivalent number of observations.

While there is still some question of whether it is better to separate the weights by tributary or to pool them by project type, our analyses suggest that a pooled weight model should be seriously considered. It is important to note that the pooled weight model is only one approach for dealing with the extreme and unsuitable “perfect-fitting” behavior of the separated weights model. Other weighting schemes that we did not investigate could potentially deal with this issue better than the pooled weights model. For example, it would be beneficial to explore whether tributary projects that count more of the total escapement should be weighted more heavily (tributaries with more fish get weighted more heavily). For the analyses presented in this document, our intent was to highlight the weighting behavior of the current model, not to investigate which (if any) alternate model is better. These are statistical issues and should be discussed and addressed eventually, but they are beyond the scope of this document.

One option for dealing with differences among model formulations would be to run multiple models each year and then average their results. Such a strategy would help to address uncertainty in which model is best. For example, there may be three models: one that separates the weighting by each tributary project, one that pools the weights by project type, and one that weights data sources by the contribution of that tributary to the total escapement. The results could then be averaged across these models, and this average could be used for the run abundance estimate. Under this approach, it would need to be decided whether each of the models would receive equal treatment, or if there are some formulations that are considered more plausible than others. There are well-established methods that have been developed to do these kinds of computations, if this approach were to be pursued in the future.

## **Acknowledgements**

This work would not have been possible without the cooperation of ADF&G. In particular, we would like to thank Zach Liller and Toshihide “Hamachan” Hamazaki for their interest and support in these analyses. ADF&G was involved in the preliminary discussion that led to these analyses and provided model code, data, and the preliminary 2014 mark-recapture estimate. We would like to thank



them for their support and assistance in these analyses. We, as independent analysts, and ADF&G collectively agree that analyses of this sort are critical in advancing our understanding and that they lead to overall better assessment techniques. While this document is the result of a set of independent analyses, it would not exist without the support of ADF&G and we look forward to continued collaboration in the future.

#### REFERENCES

- Bue, B.G., K.L. Schaberg, Z.W. Liller, and D.B. Molyneaux. 2012. Estimates of the historic run and escapement for the Chinook salmon stock returning to the Kuskokwim River, 1976-2011. Alaska Department of Fish and Game, Report to the Board of Fisheries, Anchorage. Available online at: <http://www.adfg.alaska.gov/FedAidpdfs/FDS12-49>
- Hamazaki, T., and Z.W. Liller. 2015. 2014 Kuskokwim River Chinook Salmon Run Reconstruction and Model Revisions. Memorandum from the Alaska Department of Fish and Game, Division of Commercial Fisheries. Available online at: <http://www.adfg.alaska.gov/static/fishing/PDFs/commercial/kuskokwim/2014RunReconMemo.pdf>
- Schaberg, K.L., Z.W. Liller, D.B. Molyneaux, B.G. Bue, and L. Stuby. 2012. Estimates of total annual return of Chinook salmon to the Kuskokwim River, 2002-2007. Alaska Department of Fish and Game, Fishery Data Series No. 12-36, Anchorage. Available online at: <http://www.adfg.alaska.gov/FedAidPDFs/FDS12-36.pdf>

Table 1. Overdispersion parameter estimates for each escapement project

Project Type	Tributary	Mean	Lower 95% CI	Upper 95% CI
Weir	Kwethluk	5.81	2.61	12.93
	Tuluksak	5.00	2.48	10.09
	George	12.44	5.14	30.10
	Kogrukluk	8.36	4.26	16.41
	Tatlawiksuk	23.58	7.72	72.00
	Takotna	7.85	3.42	18.02
Aerial Survey	Kwethluk	3.07	1.33	7.07
	Kisaralik	1.48	0.84	2.61
	Tuluksak	3.37	1.47	7.73
	Salmon (Aniak)	2.92	1.74	4.92
	Kipchuk	4.22	2.23	8.00
	Aniak	8.80	3.86	20.06
	Holokuk	1.71	0.82	3.55
	Oskawalik	2.18	1.16	4.12
	Holitna	4.42	2.16	9.04
	Cheeneetnuk	3.27	1.67	6.40
	Gagaryah	4.26	2.18	8.34
	Pitka	3.58	1.50	8.56
	Bear	6.70	3.11	14.43
	Salmon (Pitka)	4.78	2.60	8.77

Table 2. Descriptions of alternative model formulations that were run.

General Name	Description
Separate weights	The model estimates separate weights for each weir and aerial survey. This is the ADF&G base model.
Pooled weights	The model estimates a single weight for all weirs and a single weight for all aerial surveys
Without 2014 Weir Data	Run the model without 2014 weir project data, separate weights for all projects
Without 2014 Aerial Data	Run the model without 2014 aerial survey data, separate weights for all projects
With 2014 MRC	Run the model with 2014 weir and aerial survey data and include the 2014 mark-recapture estimate (using the separated weights model)

Table 3. Run abundance and escapement estimates for alternative model formulations

Scenario	Run Abundance			Escapement		
	Mean	Lower 95% CI	Upper 95% CI	Mean	Lower 95% CI	Upper 95% CI
ADF&G Base Model (Separate Weights)	137,932	102,364	185,857	126,170	90,602	174,095
Pooled Weights	117,411	90,673	152,035	105,649	78,911	140,273
Without 2014 Weir Projects	154,962	108,046	222,250	143,200	96,284	210,488
Without 2014 Aerial Surveys	105,842	70,023	159,985	94,080	58,261	148,223
With 2014 MRC Estimate	97,087	83,423	112,989	85,325	71,661	101,227

## Comparison of 2014 Run Abundance and Escapement Between Sensitivity Analyses

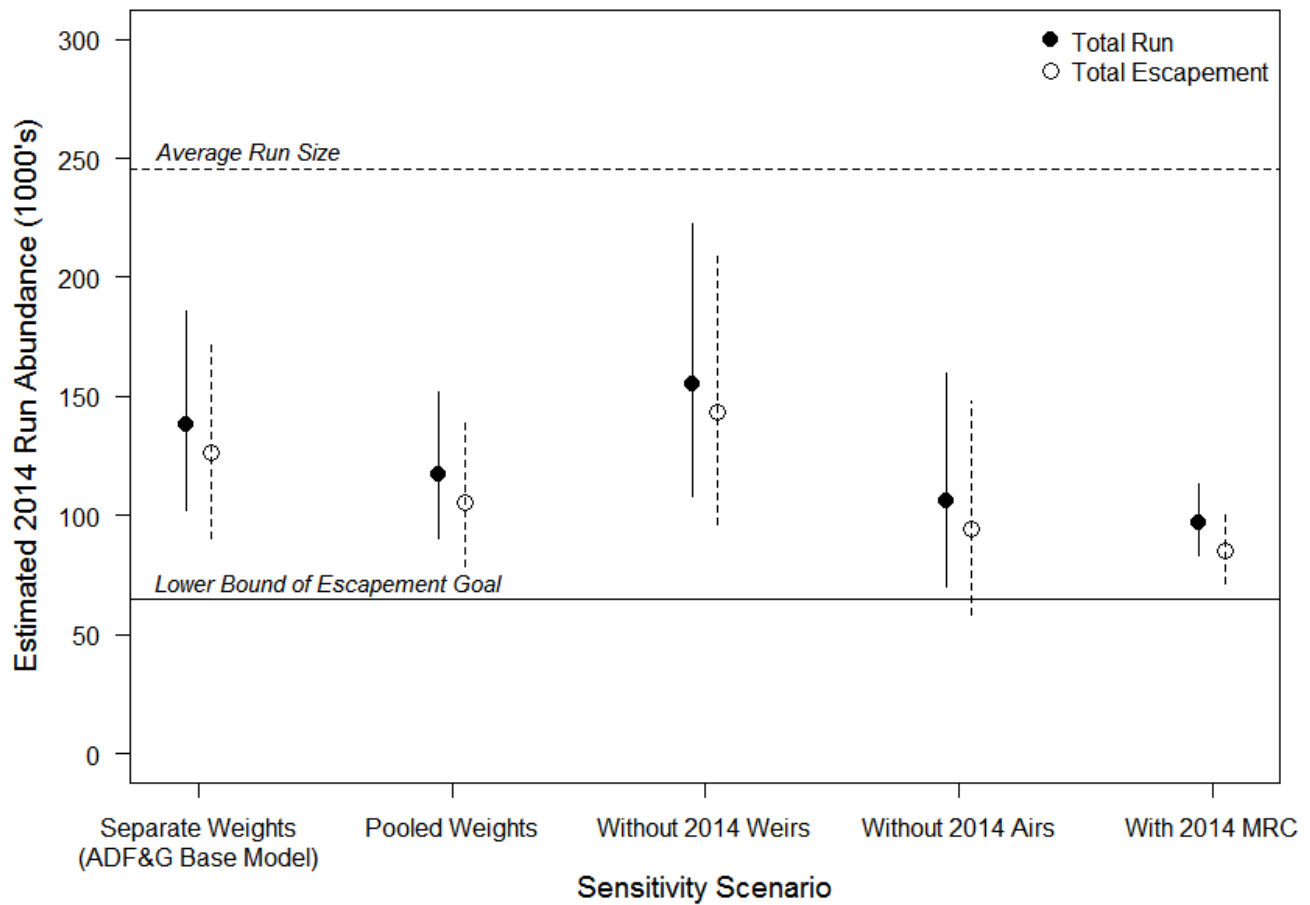


Figure 1. Run abundance and total escapement estimates (95% confidence intervals) for the year 2014 from alternative model formulations.

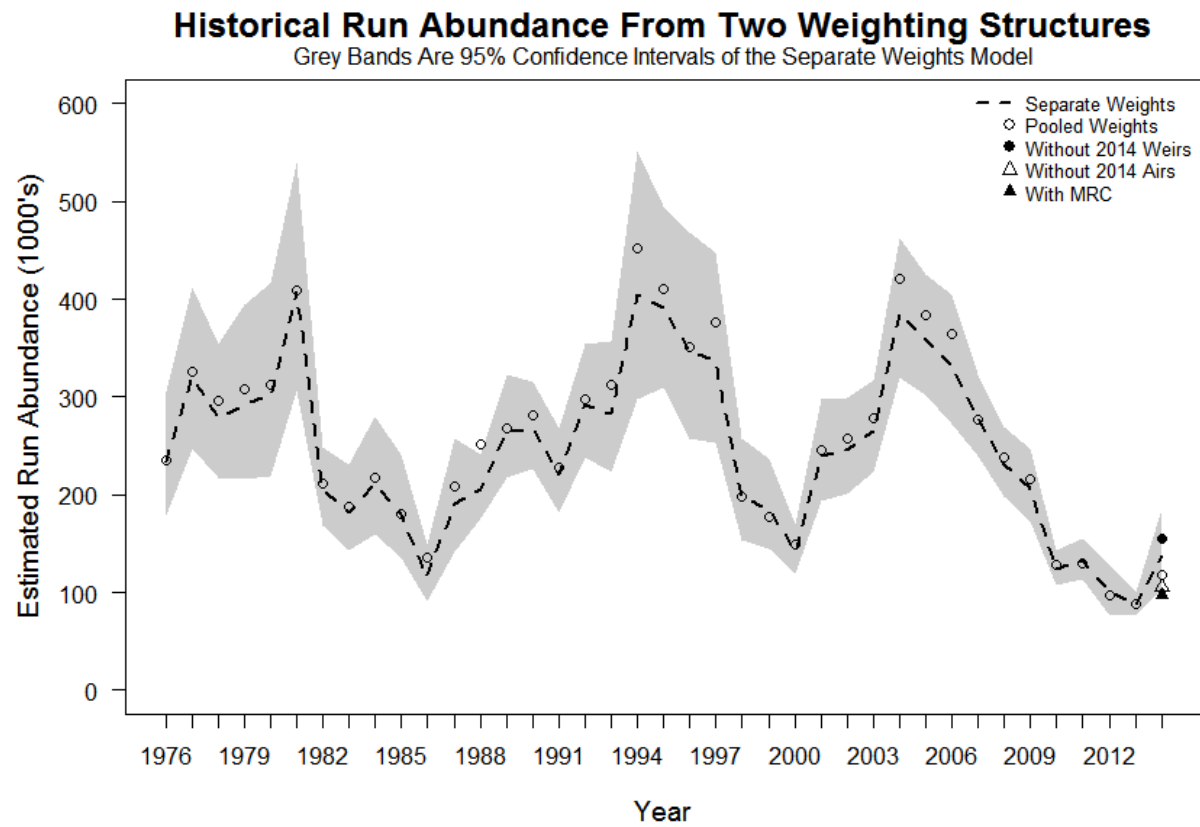


Figure 2. Time series of run abundance estimates from alternative model structures.

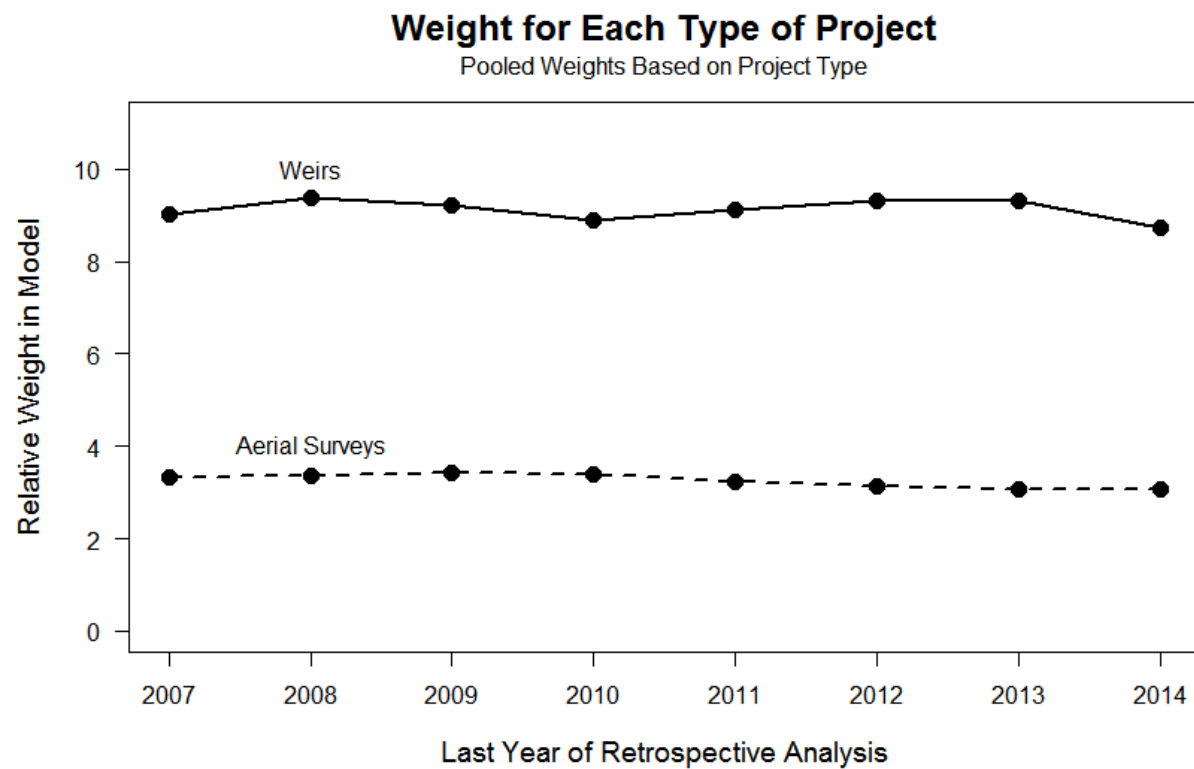
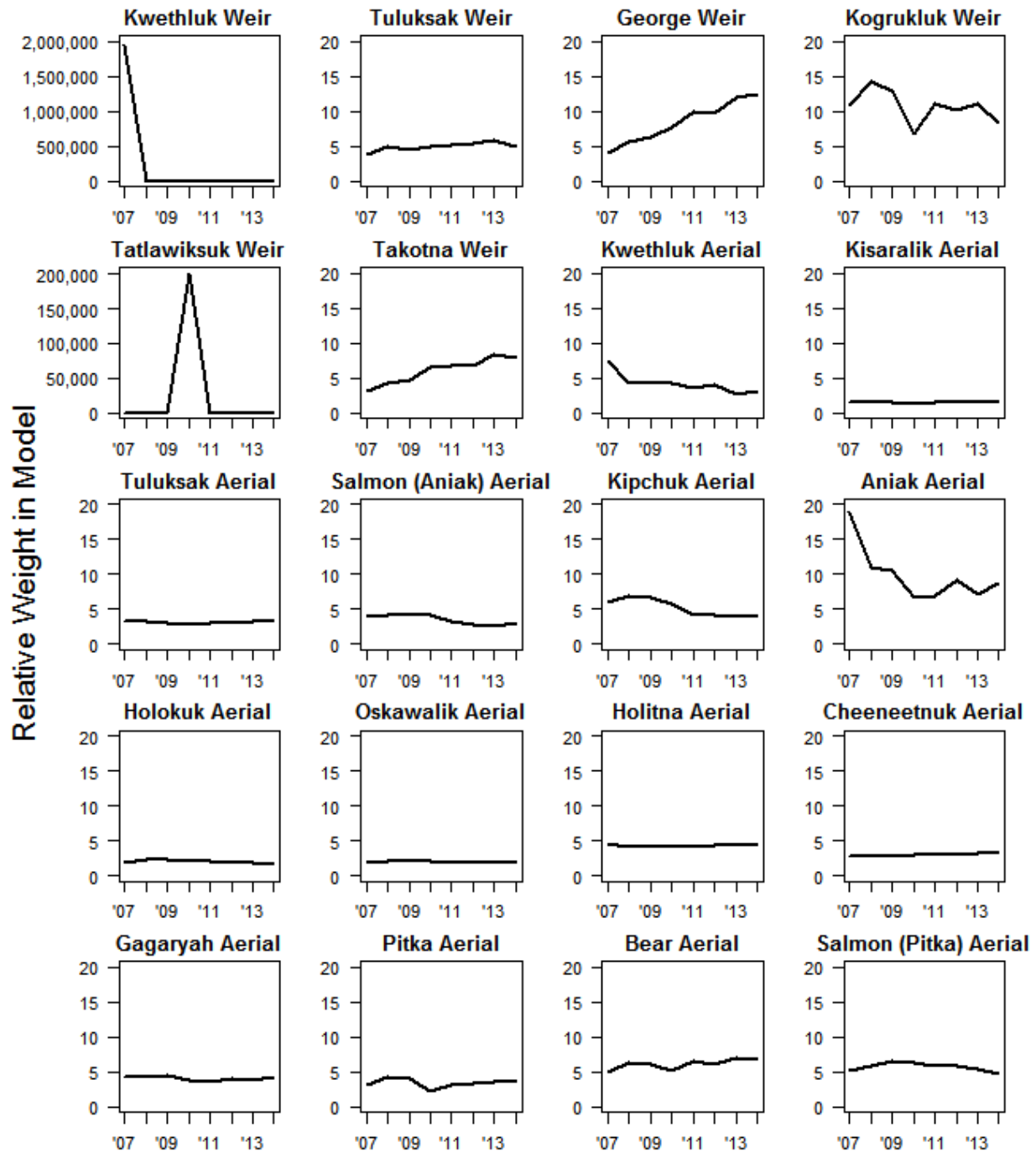


Figure 3. Variation in escapement index weight estimates (overdispersion parameters) for aerial surveys and weirs from the pooled weights model.

# Weight of Each Tributary Project

Separate Weights Model



Last Year of Retrospective Analysis

Figure 4. Retective estimates of escapement index weights from the individual weights model.

# Observed and Predicted Index Counts

Separated Weight Retrospective Analysis Ending with 2007

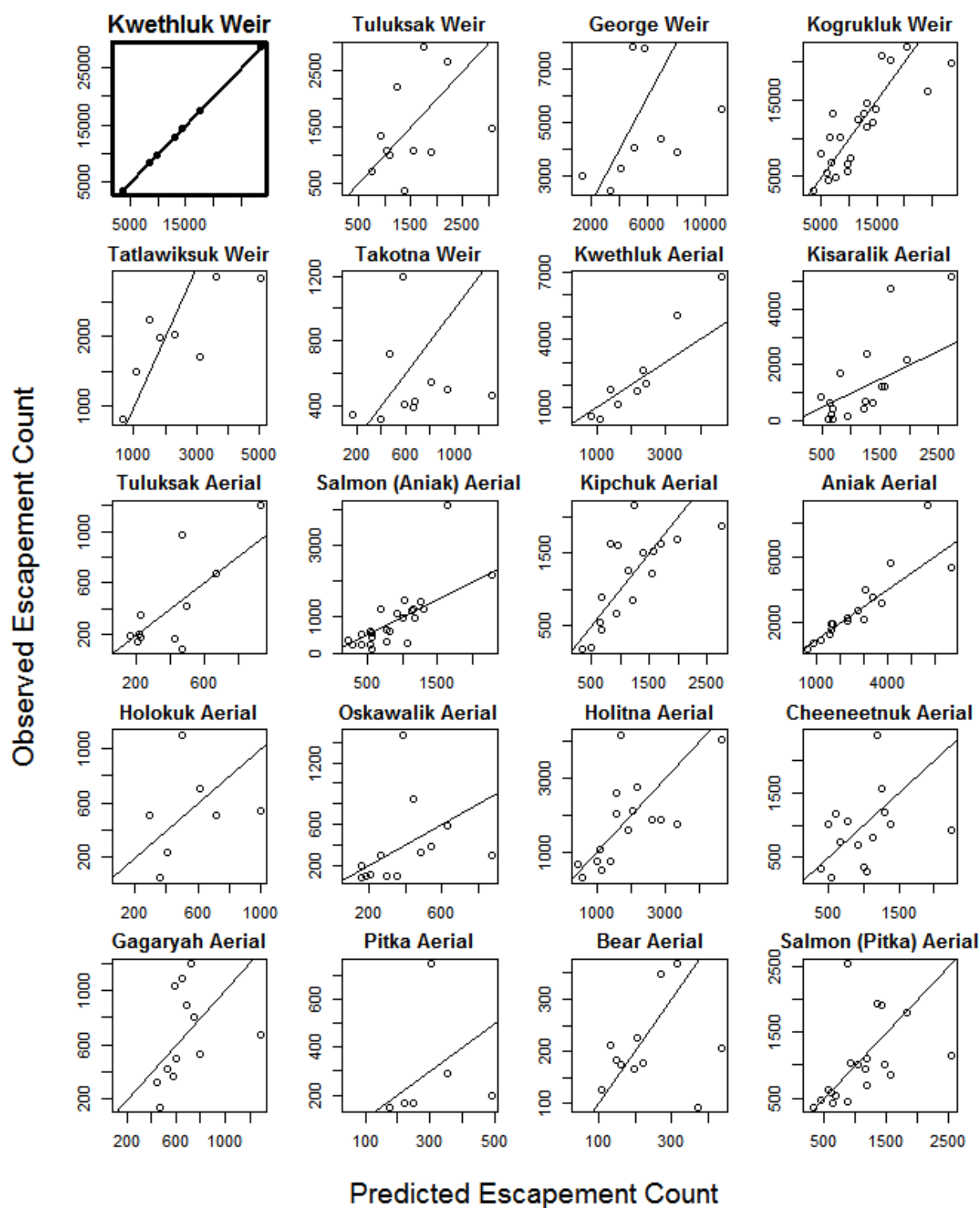


Figure 5. Observed and predicted escapement index counts from the 2007 run of the individual weights model.



# Observed and Predicted Index Counts

Separated Weight Retrospective Analysis Ending with 2010

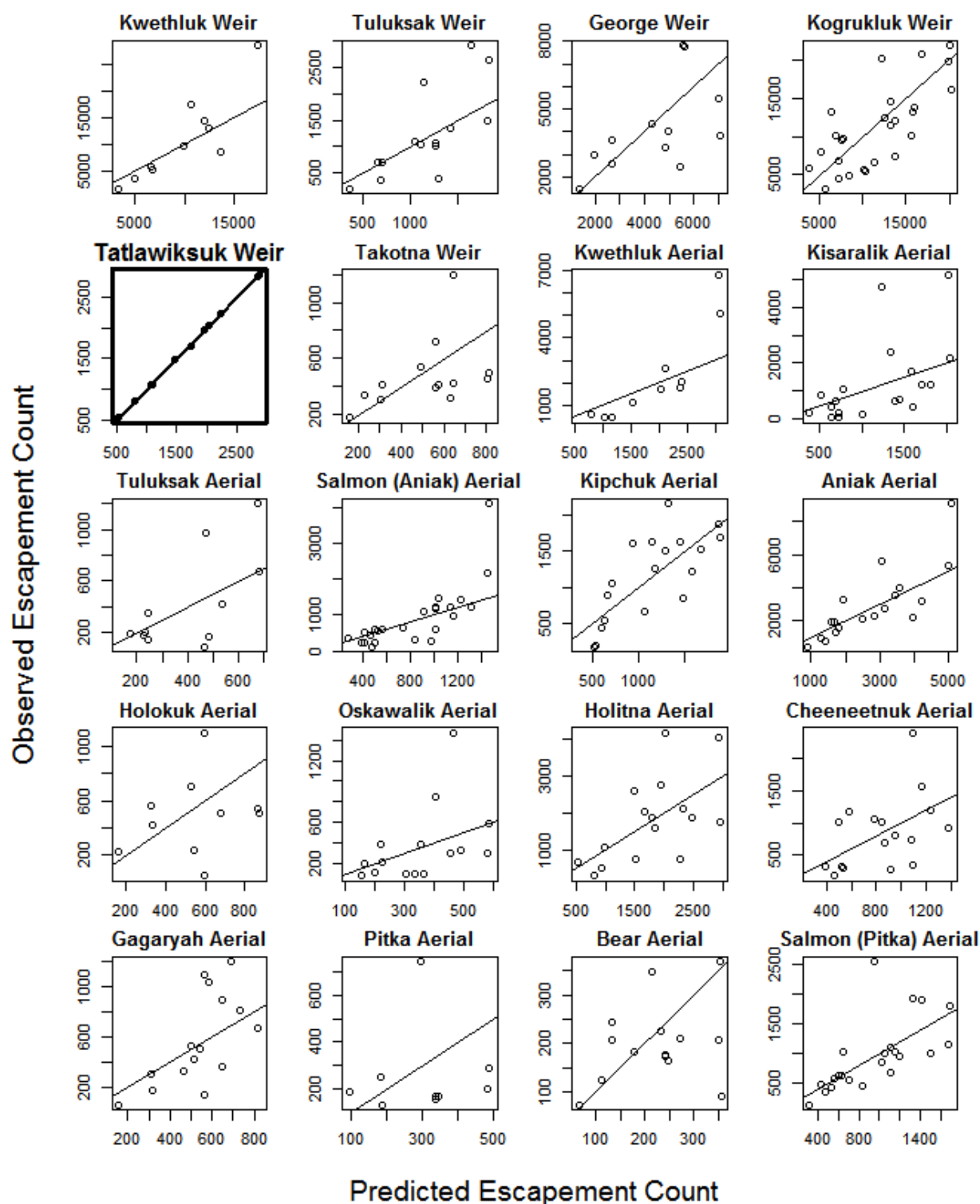


Figure 6. Observed and predicted escapement index counts from the 2010 run of the individual weights model .

## Run Abundance Changes with Retrospective Analysis

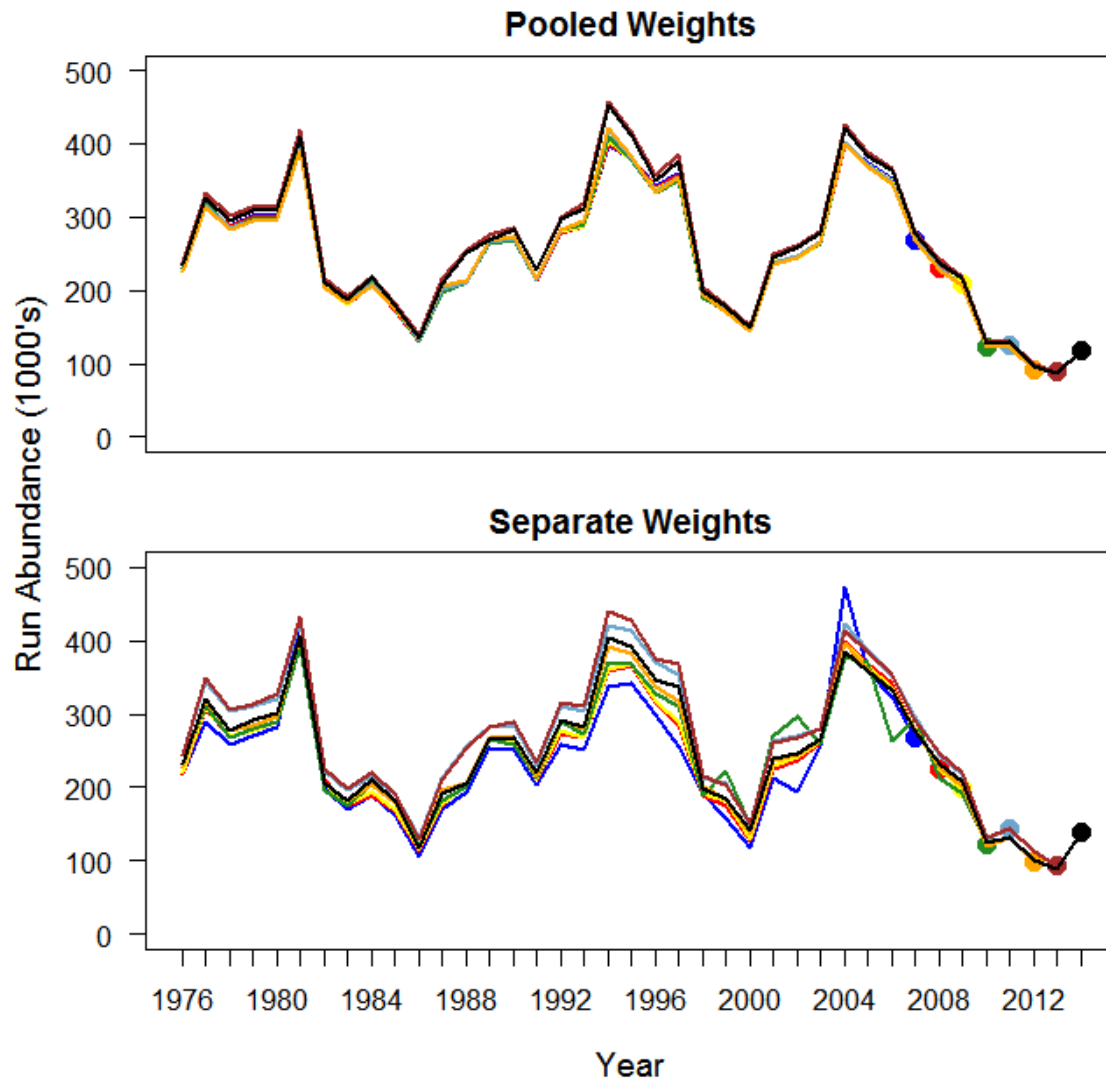


Figure 7. Retrospective run abundance estimates from the pooled weights (upper panel) and separate weights (bottom panel) models.

## CHAPTER 4

### A BAYESIAN MODEL TO ESTIMATE KUSKOKWIM RIVER CHINOOK TRIBUTARY RUN TIMING AND ABUNDANCE FROM TELEMETRIC MARK-RECAPTURE DATA

#### ABSTRACT

We conducted a new analysis of the telemetric mark-recapture data from Kuskokwim River Chinook salmon from 2003-2007 to demonstrate a new analytical approach that has the potential to take advantage of the full range of spatial and temporal information contained in these data. Specifically we sought to quantify estimates of abundance and run timing for tributary-specific Chinook salmon stocks in the Kuskokwim drainage while simultaneously estimating total escapement upstream of the tagging site. To accomplish this goal, we used a Bayesian modification of the maximum likelihood methods described in Bromaghin et al. (2010). Across all years, five of the twelve tributary groups (Mainstem [mean proportion = 0.15], Aniak & Middle Tributaries [0.19], Holitna River [0.09], Hoholitna River [0.18], and Kogruklu River [0.11]) were consistently the largest contributing groups to the run across all time strata. On average these groups composed 0.72 of the total run. Tributary-specific temporal patterns also existed. Chinook salmon from upstream tributary groups tended to complete their migrations earlier than those from middle and Holitna tributary groups. Two upstream tributary groups, Upper Kuskokwim Tributaries and Swift River, contributed large proportions to the run during the first two time strata (22 June & earlier), but their contributions declined during the final three time strata (23 June & later). While this trend was clearly evident from the run timing proportions from each tributary group, the proportional composition of the run indicated that even though these upper tributary groups had higher run timing proportions early in the run they were less abundant than middle or Holitna tributary groups at the tagging site at this time due to the high abundance of the middle tributary stocks. Annual trends in total abundance were similar to those found in Schaberg et al. (2012) using the Lincoln-Petersen estimator. While annual trends were similar, the Bayesian model tended to predict lower total abundance than Schaberg et al. (2012) in most years (2003-2006; range: 15 – 23% lower), with the difference between the Bayesian model and Schaberg et al. (2012) model estimates increasing as population abundance increased (Figure 1). In conclusion, the finer-scale information on run timing and abundance provided by the Bayesian model should be useful to ongoing management efforts aimed at understanding the diversity of Chinook salmon spawning stocks within the Kuskokwim River Drainage.

## INTRODUCTION

Pacific salmon populations in large river systems are typically comprised of many distinct spawning stocks that mix in the mainstem of rivers, where they are exposed to harvest, before segregating into their distinct spawning tributaries. Distinct spawning stocks may differ in their abundance, run timing, and other life-history traits due to site-specific factors (Hilborn et al. 2003; Waples et al. 2005). The existence of such locally-adapted spawning stocks has been suggested to buffer against ecological and environmental stressors because of complementary dynamics of locally-adapted stocks (Schindler et al. 2010; Griffiths et al. 2014). Therefore, a key goal of managing Pacific salmon species in recent years has been to promote watershed and harvest regulations that protect and restore the diversity life-history traits contained within discrete salmon populations (Griffiths et al. 2014).

A key challenge in managing mixed-stock fisheries, such as Chinook salmon, can be allocating harvest so that a few stocks are not disproportionately exploited. Historically, in the Kuskokwim Basin, subsistence Chinook harvest has occurred near the mouth of the river and has been loaded heavily towards the front end of the run (Linderman and Bergstrom 2009). This creates the potential to adversely affect discrete stocks that consistently run earlier than other stocks. In other Alaskan drainages, previous research has shown that Chinook salmon stocks with the longest distance to travel begin their upriver migration earlier than stocks that have shorter distances to travel (Wuttig and Evenson 2001; Clark et al. 2015). The same pattern has also been observed in Columbia River Chinook salmon stocks (Keefer et al. 2004). If this relationship between migration distance and run timing exists for Chinook salmon in the Kuskokwim drainage, the potential may exist for upriver stocks to be subjected to a disproportionate amount of exploitation, if the majority of harvest is allocated during the early portion of the run. Upriver stocks may also experience higher mortality rate because they are vulnerable to a large proportion of the fishing effort that occurs in the drainage.

Chinook salmon fisheries in the Kuskokwim River are actively managed through the use of fixed escapement goal policies, as are numerous salmon stocks across Alaska. Escapement goals set the number of spawners that should be allowed to escape harvest and spawn each year to produce future, sustainable yield (Clark et al. 2009). Typically, escapement goals are set using well-known production models based on stock-recruit theory (e.g., Ricker et al. 1950). However, low quality and incomplete enumeration of salmon runs across large, remote drainages has precluded the use of traditional spawner-recruit models to set escapement goals for several Chinook salmon stocks in western Alaska, including those in the Kuskokwim River. As an alternative to spawner-recruit models, Bue et al. (2012) adapted an existing model, used to estimate drainage-wide run abundance and escapement in data-limited circumstances (Shotwell and Adkison 2004), to Chinook salmon runs in the Kuskokwim River. The Bue et al.(2012) model used escapement data from weirs, counting towers, air surveys, along with harvest estimates, and, importantly, was anchored by five years of abundance estimates generated from a telemetric mark-recapture study (Schaberg et al. 2012), which allowed for the Bue et al. (2012) model to be calibrated to empirical abundance estimates.

Telemetric mark-recapture studies have become an important method for estimating drainage-wide escapement for Kuskokwim River Chinook salmon. The analytical approach for these studies has been the two-sample Lincoln-Peterson estimator. One key assumption of these studies is that capture probabilities are similar across all stocks, during the entire duration of the study. Analysis of Kuskokwim Chinook mark recapture models has indicated that the Lincoln-Peterson model is appropriate for Kuskokwim Chinook, specifically that the equal capture probability assumption has been satisfied. Nevertheless, maximum likelihood and Bayesian approaches are increasingly being used to take fuller advantage of the spatial and temporal nature of these data (e.g., Reimer et al. 2016). Using temporal stratification, these methods can relax the assumption of equal catchability. Additionally, these approaches can generate stock-specific estimates of abundance and run-timing that will provide

increased knowledge of the diversity of abundance and migratory timing for these discrete stocks. In short, these new approaches have the potential to “squeeze” more out of the data than the traditional Lincoln-Peterson estimator.

Herein, we provide a new analysis of the telemetric mark-recapture data from Kuskokwim River Chinook to demonstrate a new analytical approach that has the potential to take advantage of the full range of spatial and temporal information contained in these data. Specifically we sought to quantify estimates of abundance and run timing for tributary-specific Chinook salmon stocks in the Kuskokwim Drainage while simultaneously estimating total escapement upstream of the tagging site. To accomplish this goal, we used a Bayesian modification of the maximum likelihood methods described in Bromaghin et al. (2010). Fortunately, all of the data required for such a model were collected during a previous telemetric mark-recapture survey conducted in the Kuskokwim drainage from 2003-2007. Therefore, the potential exists for this Bayesian model to provide simultaneous estimates of tributary-specific run timing and abundance while removing the requirement for equal capture probability over time. Such information would provide managers useful information that they could use for adjusting management actions and, potentially, develop new strategies to prevent the over-harvest of discrete stocks and maintain a diverse portfolio of Chinook salmon spawning stocks in the Kuskokwim River.

## **Methods**

### *BART Model Overview*

We used a state-space Bayesian Abundance and Run Timing Model (hereafter, referred to as BART), which was a modification of the maximum likelihood- based model developed by Bromaghin et al. (2010). This model simultaneously estimates both run timing parameters and abundance of specific tributary groups using previously collected data from a telemetric mark-recapture study. As in the Bromaghin et al. (2010) model, temporal differences in capture probability are incorporated directly into the BART model and do not result in biased estimates. The model stratifies Chinook salmon abundance by space and time, where twelve tributary groups (Kuskokwim Mainstem, Aniak & Middle Tributaries, Salmon River, George River, Holitna River, Hoholtna River, Kogruklu River, Stony River, Swift River, Tatlawiksuk River, Upper Kuskokwim Tributaries, and Takotna River) represent the spatial stratification and five approximately 1-week intervals (15 June and earlier, 16-22 June, 23-29 June, 30 June – 6 July, 7 July and later) represent the time strata. Additionally, we aggregated tributaries into four large composite groups (Mainstem, Middle Tributaries, Holitna Drainage, and Upper Tributaries) to investigate large-scale patterns in total abundance and run timing. A list of tributaries contained within each both tributary and composite groups is presented in Table 1.

To estimate run timing and abundance, the BART model combines the temporal distribution of catch rates at a marking location with known final fate information (tributary destination) to estimate run timings of tributary groups while simultaneously estimating abundance from weir counts and recapture data. Information on relative abundance across tributary groups is provided by the final locations of radiotagged Chinook salmon. Information on the relative abundance by temporal strata is provided by CPUE data from gillnet collection efforts to radiotag Chinook salmon in the Kuskokwim River. Total Chinook salmon counts at tributary weir projects anchor the analysis by providing known escapement for a subset of stocks in each year. The BART model was run separately for each year, using telemetric mark-recapture data from 2003-2007.

### *Radiotagging Background*

Complete details of telemetric mark-recapture studies conducted in the Kuskokwim River during 2003-2006 are provided in Stuby et al. (2007). Study details from the 2007 mark-recapture study are provided in Schaberg et al. (2012). Here, we provide a brief overview of collection, tagging, and tacking methods used during these years. During 2003-2007, Chinook salmon returning the middle and upper

tributaries of the Kuskokwim River were captured across the entire run near Kalskag, Alaska (rkm 270) using drift gillnets and fish wheels. The exact location of collection differed across years (see Stuby et al. 2007 for details), but was within 10-12 km on the upstream or downstream side of Kalskag during each year. Collection efforts were conducted 6 days a week from the start to the end of the run. Drift gillnets (30.48 m in length, 20.32 cm mesh size, and 29 panels deep) fishing effort alternated between the north and south banks of the Kuskokwim River every 45 minutes of soak time, to ensure that all spatial components of the run had a non-zero probability of capture. During each day of collections gillnets were fished for approximately 3 hours of soak time. Fish wheels ( $N = 3$ ) were fished for 24 hours a day during the run. Two fish wheels were operated on the north bank and one fish wheel was operated on the south bank of the Kuskokwim River near Kalskag.

Upon collection, a subset of Chinook salmon ( $\geq 450$  mm) were tagged with esophageal implant radio tags (Model Five pulse encoded transmitters made by Advanced Telemetry Systems; Isanti, Minnesota) that were distinguishable by a unique frequency and encoded pulse pattern. Tags were deployed proportional to run strength, based on catches from a test gillnet fishery operated near Bethel, Alaska by the Alaska Commercial Fisheries Division. Additionally, weekly tagging goals were developed for small ( $< 650$  mm) and large ( $\geq 650$  mm) Chinook salmon, based on historical length data from upriver weirs. Based on these data, approximately 20% of Chinook salmon escapement past weirs was by fish that were  $\geq 650$  mm (Stuby et al. 2007).

Radiotagged fish were tracked upriver using a network of 13 (2003) to 17 (2007) ground-based tracking stations similar to those in Eiler et al. (2005). Tracking stations extended from just downstream of the tagging area, near the abandoned village of Uknavik (rkm 264), to the far upstream reaches of the drainage near the villages of McGrath and Takotna (rkm 835). Additionally, two aerial surveys were conducted in each year to obtain final locations of fish on the spawning grounds. During 2003-2006 these two flights were conducted during mid-July and mid-August (Stuby et al. 2007), while in 2007 they were conducted during late-August and late-September (Schaberg et al. 2012). These flights covered the vast majority of the middle and upper Kuskokwim drainage, upstream of rkm 233.

#### *Required Data Sources*

CPUE Index. Mean daily catch-per unit effort (CPUE) was calculated for each day that gillnetting was conducted to collect Chinook salmon for radiotag implantation. CPUE was averaged across days within each time strata to generate an index of abundance in each strata. Each day, drift gillnets were fished for approximately 3 hours of soak time. Individual drifts lasted for a maximum duration of  $\sim 15$  minutes, but could be as brief as 1 minute if a Chinook salmon was caught immediately upon deployment. Once a fish was caught in the net, the tagging crew would immediately pull the net, to minimize stress on the captured fish. We used simulated data to investigate the potential of this sampling behavior (i.e., pulling the net as soon as fish was captured) to bias estimate of relative abundance. The results indicate that while pulling the gillnet as soon as a fish is captured may bias estimated abundance when compared to known or true abundance, this bias scales linearly with true abundance. Therefore, CPUE from gillnet tagging efforts should be a suitable index of relative abundance across time strata.

Final Locations for Radiotagged Fish. Final locations of tagged fish were determined by a combination of aerial surveys and ground-based telemetry tracking stations. To identify fish that successfully continued upstream after tagging and required a final location assignment, we used the same criteria as Schaberg et al. (2012). Specifically, a tagged fish was deemed part of the sample population if any of the following conditions were met: 1) it was recorded at any ground-based tracking station located upstream of Birch Tree Crossing station (rkm 294) and remained upstream of Birch Tree Crossing for more than 7 days, 2) it was located upstream of Birch Tree Crossing during one of the aerial tracking surveys, or 3) it was harvested upstream of Birch Tree Crossing. Annual counts for the number

of fish tagged as well as the number of fish that entered the sample population and were assigned a final location are presented in Table 2.

For fish that met these criteria, we used either the final locations from aerial tracking or ground-based tower detections to assign each fish to one of the 12 tributary groups. Using GIS software, final GPS coordinates for each fish from aerial surveys was overlaid on drainage boundaries for specific tributaries (Table 1). Subsequently, each Chinook with a final location was assigned to one of the 12 tributary groups based on the drainage boundary it was located within. For fish not located during the final aerial surveys, we used telemetry tower data to determine the most upstream recorded location of each fish. For some fish, the final location from aerial tracking surveys and/or the most upstream tower detection was along the mainstem. For these fish we were unable to determine a final location, and grouped these fish into the Kuskokwim Mainstem tributary group.

**Weir Counts.** Weir counts provided known escapement for some stocks in each year of the study. This allowed for annual abundance estimates for the BART model to be anchored to annual escapement for these stocks. Annual Chinook salmon counts were used from the Salmon River (Aniak River drainage, rkm 404), George River (rkm 453; Thalhauser et al. 2008), Tatlawiksuk River (rkm 568; Stewart et al. 2008), Kogruklu River (Holitna River drainage, rkm 710; Williams and Shelden 2010), and Takotna River (rkm 835; Costello et al. 2008), although all weirs were not operated in each year (Hansen and Blain 2014; Table 3). Additionally, because count data from several weir projects used in Schaberg et al. (2012) are now considered to be of low quality, due to incomplete operation during the Chinook salmon run (Hansen and Blain 2014), we ran the BART model using either: 1) updated weir counts, as reported in Hansen and Blain (2014), or 2) weir counts previously reported in Schaberg et al. (2012). This allowed us to investigate if the choice of weir data affected our estimated population abundance. Hansen and Blain (2014) used a Bayesian framework to estimate Chinook passage during inoperable periods (i.e., caused by flooding or disassembly for maintenance) and also omitted counts from three weir projects (i.e., George River in 2003, Salmon River in 2006, and Kogruklu River in 2007) that were used in Schaberg et al. (2012). Weir counts used in Schaberg et al. (2012) and Hansen Blain (2014) are presented in Table 3.

### *Model Details*

The Bayesian state-space BART model is comprised of two submodels, which are fully described below. The process model describes the true unobserved state of the system, including probability structures of key population parameters. The observation model then provides sampling distributions of the observed data that link the true states with the observed data from Kuskokwim Chinook salmon projects conducted during 2003-2007. The full probability model describes the joint probability distribution of the observed data and the population parameters of interest. The complete BART model is provided in Appendix 2.

Additionally, we ran several iterations of the BART model to test the effect of data sources and sharing of information across years on estimated run timing and abundance. In the base BART model, we used weir counts as presented in Hansen and Blain (2014) and ran the model for each year (i.e., 2003-2007) separately, without sharing any information across years (Appendix 1A). We also ran the base BART model for each year using weir data from Schaberg et al. (2012), to determine the effect of differing weir data on estimated abundance. Finally, we also ran an alternative form of the BART model, which shared run timing information for specific tributary groups across years (Appendix 1B). We ran this version of the shared run timing BART model separately using either our base weir data (Hansen and Blain 2014) or the Schaberg et al (2012) weir data.

### *Process Model*

The number of Chinook salmon from tributary group  $i$  that pass the tagging site during time strata  $j$  follows:

$$N_{ij} = N \pi_{ij}$$

where  $N$  is the total estimated Chinook abundance north of Birch Tree Crossing in a given year, drawn from a normal distribution with mean  $\mu_N$  and standard deviation  $\sigma_N$ . The value of  $\mu_N$  was set to 0 and  $\sigma_N$  was extremely large ( $10^6$ ), so that the normal distribution would be equivalent to a uniform distribution over the range of the likelihood (Fleischman et al. 2013). The parameter  $\pi_{ij}$  represents the proportion of  $N$  in tributary group  $i$  passing the tagging site during strata  $j$ . The parameter  $\pi_{ij}$  sums to 1.0 across all tributary groups and time strata within each year and is calculated from an array of expected frequencies of occurrence of each tributary group during each strata ( $\gamma_{ij}$ ) that follow a Dirichlet distribution. The Dirichlet is the conjugate distribution for the multinomial distribution, which is used to model the spatial and temporal distributions of radiotagged Chinook salmon in the observation model (described below). For tributary groups  $i = \{1-10\}$ , the Dirichlet distribution was constructed by using a series of independent Gamma distributions where:

$$\pi_{ij} = \frac{\gamma_{ij}}{\sum_{j=1}^5 \sum_{i=1}^{11} \gamma_{ij}}$$

$$\gamma_{ij} \sim \text{gamma}(\text{shape} = 1, \text{rate} = 0.1)$$

For the most upstream tributary group, Takotna River  $i = \{12\}$ , there were insufficient final location observations to allow for  $\gamma_{ij}$  parameters to be fit to the data. Therefore, we estimated  $\gamma_{ij}$  independently for tributary groups  $i = \{1-11\}$  and then shared information from an adjacent tributary group, Upper Kuskokwim Tributaries  $i = \{11\}$ , with Takotna River. To accomplish this we fit a proportional parameter,  $\varphi$  to describe the proportion of occurrences of  $\gamma_{11}$  in each time strata  $j$  that should have occurred in the Takotna River. This sharing of information allowed for Takotna River to share run timing exactly with the Upper Kuskokwim Tributaries, but for it to have an independent abundance estimate. As such  $\pi_{ij}$  was for  $i = \{11\}$  was constructed as:

$$\pi_{ij} = \frac{\gamma_{ij} \varphi}{\sum_{j=1}^5 \sum_{i=1}^{11} \gamma_{ij}}$$

while  $\pi_{ij}$  was for  $i = \{12\}$  was constructed as:

$$\pi_{ij} = \frac{\gamma_{ij} (1 - \varphi)}{\sum_{j=1}^5 \sum_{i=1}^{11} \gamma_{ij}}$$

Both tributary-specific total abundance ( $N_i$ ) and strata-specific total abundance ( $N_j$ ) are derived from  $N_{ij}$  as follows:

$$N_i = \sum_j N_{ij}$$

$$N_j = \sum_i N_{ij}$$

Additionally, several other metrics were calculated, for both use in the Observation Model, and for ease of comparing run timing among stocks. The parameter  $\theta_{ji}$ , the proportion of Chinook salmon in strata  $j$  from each tributary group ( $i$ ) was calculated as:

$$\theta_{ji} = \frac{N_{ij}}{N_j}$$



so that  $\theta_{ji}$  sum to 1.0 for each time strata. Similarly,  $\rho_{ij}$ , the proportion of Chinook salmon from each stock ( $i$ ) present in strata  $j$  was calculated as:

$$\rho_{ij} = \frac{N_{ij}}{N_i}$$

so that  $\rho_{ij}$  sums to 1.0 for each stock.

#### *Observation Model*

Observed data used by the BART model included annual weir counts from 5 weir projects conducted across the Kuskokwim River drainage north of Birch Tree Crossing (Table 3), gillnet CPUE by time strata (Table 4), and final spawning destinations of radiotagged fish for each tagging strata (Table 5).

Observed weir counts ( $\hat{N}_i$ ), which served as an estimate of total annual abundance for stocks with weirs, were assumed to be lognormally distributed with parameters  $\ln(N_i)$  and  $\sigma_i$ , from the process submodel. The value of  $\sigma_i$  was calculated from CVs (following Evans et al. [1993]), which were assumed to be 0.1 for all weirs. This assumption seems reasonable based on the credible intervals on total weir count from the Kuskokwim drainage analyzed using Bayesian methods in Hansen and Blain (2014). Also, sensitivity analyses revealed that increasing or decreasing CVs on weir counts from 0.01 – 0.2 had little effect on total or stock-specific abundance estimates.

Catch per unit effort (CPUE) from gillnet catches in each time strata during the original tagging study served as the index of abundance that scaled total estimated abundance ( $N$ ) across  $j$  time strata. Catch per unit effort in time strata  $j$  was modeled as linearly related to abundance:

$$\ln(CPUE_j) \sim \text{norm}(\ln[Nq_j], \sigma_{CPUE})$$

$$Nq_j = N_j q$$

with  $q$  serving as the constant of proportionality between total estimated abundance and gillnet CPUE in each time strata.

The number of Chinook salmon with final locations in each  $i^{\text{th}}$  tributary group within each  $j^{\text{th}}$  time strata ( $x_{ji}$ ) was modeled as multinomially distributed, with order parameter  $n_j$  (sum of final locations across all stocks in each time strata) and proportion parameter  $\theta_{ji}$  from the process submodel.

#### *Prior Distributions and Alternative Model Forms*

A key requirement for Bayesian analyses is that prior probability distributions be specified for all unknown parameters in the model (Table 6). Most of our prior distributions in this model were designed to be uninformative and have minimal effect on the posterior. However, a few exceptions were made. Specifically, the prior distribution for  $N$  excluded negative values, as it would be biologically impossible to have a negative value for population abundance. Also, an inverse gamma (100,1) was given to  $\sigma_{CPUE}$ , which is equivalent to a CV = 0.1. To test the effect of our assumption, we conducted sensitivity analyses with differing inverse gamma distributions, with CV equivalents ranging from 0.05 – 0.2 (Table 6). Also, while a relatively uninformative prior was assigned to the gamma distribution (shape=1, rate=0.1) used to construct the Dirichlet distribution, we conducted sensitivity analysis to quantify the effect of a less informative prior gamma distribution (shape=0.1, rate=0.01) on posterior distributions.

Finally, we explored the use of a hierarchical modification of the BART model to share run timing information across years. To accomplish we used the hyperparameter  $\alpha_{ij}$  to share the expected

frequencies of occurrence of each tributary group during each strata across years (see Appendix 1B for JAGS code). To accomplish this, which we assigned a hyperprior:

$$\alpha_{ij} \sim \text{gamma}(\text{shape} = 0.001, \text{rate} = 0.001)$$

which was subsequently used as the shape parameter in our independent gamma distribution ( $\gamma_{ij}$ ) to construct the Dirichlet distribution, as follows:

$$\gamma_{ij} \sim \text{gamma}(\text{shape} = \alpha_{ij}, \text{rate} = 0.1)$$

As in construction of the Dirichlet distribution in the base, single year BART models,  $\alpha_{ij}$  was only estimated for  $i = \{1:11\}$ , due to insufficient data from Takotna  $i = \{12\}$ , to independently estimate  $\alpha$ . As in the base, single year BART models, a proportional parameter  $\varphi$  was fit in each year to model the proportion of occurrences of  $\gamma_{11}$  in each time strata  $j$  that should have occurred in Takotna.

### *Model Fitting*

All model versions (e.g., single year base models, multi-year models with shared run-timing parameters) were fit using Bayesian integration with MCMC methods to sample from the joint posterior probability distribution. MCMC sampling was conducted using the JAGS software (“Just Another Gibbs Sampler”, Plummer 2013) implemented through R (R Core Development Team 2014) using the R package “R2jags” (Su and Yajima 2015). MCMC sampling was conducted with two chains with differing initial values to verify convergence and confirm that multiple solutions were not found. Convergence of the chains was verified by visual inspection of trace plots of the posterior distributions and by confirming that the 95% credible interval for the Gelman-Rubin statistic for each estimated parameter was  $\leq 1.1$  (Gelman et al. 2004). MCMC sampling of all base, single year BART models involved a burn-in period of 50,000 iterations, 500,000 post burn-in iterations, and a thinning interval of 100 iterations. This resulted in a total of 5,000 posterior samples retained for analysis. MCMC sampling for the shared run-timing model also involved a burn-in period of 50,000 iterations, but used 1,000,000 post burn-in iterations with a thinning interval of 200 iterations, resulting in 5,000 posterior samples being kept. Means, standard deviations, and 95% credible intervals (2.5 and 97.5 percentiles of the posterior distributions) were calculated for quantities of interest from their marginal posterior distributions.

## **RESULTS**

### *Total Abundance*

Estimated Chinook salmon abundance for the Kuskokwim River drainage north of Birch Tree Crossing (rkm 294) from the base BART model ranged from a high of 195,472 (95% CI: 148,433 – 254,391) in 2006 to a low of 108,290 (78,238 – 146,542) in 2003 (Figure 1). Annual trends in total abundance were similar to those found in Schaberg et al. (2012) using the Lincoln-Petersen estimator. While annual trends were similar, the BART model tended to predict lower total abundance than Schaberg et al. (2012) in most years (2003-2006; range: 15 – 23% lower), with the difference between BART and Schaberg et al. (2012) model estimates increasing as population abundance increased (Figure 1). The one year in which abundance estimates from the base BART model and Schaberg et al. (2012) were similar (~1% difference) was 2007 (Figure 1). Running the base BART model using weir counts from Schaberg et al. (2012) did not resolve the differences in total abundance between the base BART model and Schaberg et al. (2012) (Figure 1). Running the BART model with shared run timing information across years resulted in slightly elevated abundance estimates (range: 1% - 13% increases) compared to the base BART model in most years (2003-2006). However, estimated abundance from BART models

with shared run timing (using weir data from either Hansen and Blain 2014 or Schaberg et al. 2012) was still lower than estimates from Schaberg et al. (2012)(Figure 1). The exception to this trend was 2007, for which estimated abundance from the BART model with shared run timing using Hansen and Blain (2014) weir counts was much higher than estimates from both the base BART model (32% increase) and those from Schaberg et al. (2012) abundance estimate (33% increase; Figure 1).

While total abundance fluctuated during 2003-2007, there were similar temporal trends within each year. Across all years, total abundance was generally low during the 1<sup>st</sup> time strata (15 June & earlier), much higher during the 2<sup>nd</sup> to 4<sup>th</sup> time strata (16 June – 6 July) and low again during the 5<sup>th</sup> time strata (7 July & later; Table 7 and Figure 2). During 2003, 2004, and 2007, total abundance peaked during the 2<sup>nd</sup> and 3<sup>rd</sup> time strata (16-29 June) and during 2005 and 2006 total abundance peaked during the 3<sup>rd</sup> and 4<sup>th</sup> time strata (23 June – 6 July; Table 7 and Figure 2).

#### *Run Timing by Tributary Group and Year*

The base BART model also estimated parameters for the proportion of each tributary group passing the Kalskag tagging site during each time strata. Parameter estimates and standard deviations for the proportion of each tributary group passing Kalskag during each time strata are presented in Tables 8a-8b. When these run timing proportions were averaged across years (2003-2007) for each tributary group, three main patterns emerged. The first run timing pattern was represented by four tributaries (Upper Kuskokwim Tributaries, Takotna River, Tatlawiksuk River, and the Kuskokwim Mainstem). These tributaries had the highest proportions in the first strata (15 June & earlier), reached a plateau during the second and third time strata (16-29 June) and declined during the last two time strata (30 June & later; Figure 3). The second run timing pattern was represented by two tributaries (Swift River and Salmon River), that had a distinct peak in the proportion of the run passing Kalskag during the second time strata (16-22 June) and declined during the remaining time strata (Figure 3). Finally, the third run timing pattern was represented by six tributary groups (Aniak River, George River, Holitna River, Hoholitna River, Kogruklu River, Stony River). For these tributaries, the run proportions passing Kalskag resembled a normal curve, with gradually increasing proportions in early time strata, a peak during the middle time strata (23-29 June) and gradually decreasing proportions during the later time strata (Figure 3). Generally, tributaries exhibiting the first and second run timing patterns were upstream tributaries and those exhibiting the third pattern were middle tributaries.

There also appeared to be annual differences in run timing that affected all tributary groups in a given year. For example, runs of all tributary groups in 2006 and 2007 were delayed compared to runs in 2003-2005 (Tables 8a-8b). In 2007, all tributary groups had lower proportions passing Kalskag ( $0.01$  [mean],  $\pm 0.006$  [SD]) during the first time strata (15 June & earlier) than in 2003-2005 ( $0.13 \pm 0.07$ ; Table 8a-8b). However, despite the late start to the run in 2007, the average proportion of tributary groups passing Kalskag during the 5<sup>th</sup> time strata ( $0.07 \pm 0.03$ ) were similar to those in 2003-2005 ( $0.07 \pm 0.03$ ; Table 8a-8b), indicating that most tributary groups in 2007 appeared to experience a short, compressed run. In 2006, the proportions of most tributary groups passing Kalskag ( $0.02 \pm 0.015$ ) was also low during the 1<sup>st</sup> time strata, with the exception that the proportion of Upper Kuskokwim Tributary Chinook that passed Kalskag during the first time strata ( $0.18$ ) was much greater than for all other tributary groups (Table 8a-8b). However, unlike 2007, many tributary groups in 2006 had their highest proportions passing Kalskag during the 4<sup>th</sup> time strata (30 June – 6 July), and proportions of each tributary group in the 5<sup>th</sup> strata (7 July & later) were elevated compared to the 5<sup>th</sup> time strata in other years. This seemed to indicate that the overall run during 2006 was not compressed, but that many tributary groups experience a phase shift in run timing, with their proportions increasing during the latter time strata compared to other years (Table 8a-8b).

#### *Composition of the Run by Tributary Group and Year*

During 2003-2007, Chinook salmon from Holitna Drainage were the most numerous out of the four large composite tributary groups. Across these years in almost all strata, Chinook salmon from the Holitna Drainage made up a plurality of the run (Figure 4). Holitna Drainage Chinook salmon were the most numerous during 21 of the 25 time strata across all years, ranging from 0.19 – 0.5 (mean = 0.38; Figure 4). The few strata in which Holitna Drainage Chinook salmon were not the most abundant composite tributary group, were usually the first or last strata, when overall abundance was low (Figure 4). While the Holitna Drainage Chinook appear to be the most abundant composite tributary group, other composite tributary groups also support substantial proportions of the run (Figure 4). The Middle Tributary composite group ranged from 0.11 – 0.39 and averaged 0.25 across all years (Figure 4). The Upper Tributary composite group ranged from 0.9 – 0.53 and averaged 0.22 (Figure 4) and Chinook with final locations in the Kuskokwim Mainstem averaged 0.15, ranging from 0.06 – 0.31 (Figure 4). In general the Chinook salmon run in the Kuskokwim at Birch Tree appears to be fairly diverse throughout the run, with all four composite groups contributing substantial proportions of the run across multiple years. While interannual differences exist, there appears to be consistent evidence that Chinook salmon from the Upper Tributaries compose a large portion of the run in 1<sup>st</sup> strata, but their contribution to the overall run decline as the run continues. Notably, across all years, abundance was typically low in the 1<sup>st</sup> strata (Figure 2), especially in 2006, when Upper Tributary Chinook composed the highest proportion in any time strata of the study (Figure 4).

Examination of tributary-specific proportions composing the run (Table 9), a large degree of temporal variability was evident both within and across years. However, despite this variability, some general trends exist. Across all years, five of the twelve tributary groups (Mainstem [mean proportion = 0.15], Aniak & Middle Tributaries [0.19], Holitna River [0.09], Hoholitna River [0.18], and Kogruklu River [0.11]) were consistently the largest contributing groups to the run across all time strata (Table 9). On average these groups composed 0.72 of the total run. Tributary-specific temporal patterns also existed. Two upstream tributary groups, Upper Kuskokwim Tributaries and Swift River, contributed large proportions to the run during the first two time strata (22 June & earlier), but their contributions declined during the final three time strata (23 June & later; Table 9). Conversely, several middle tributary groups including Aniak & Middle Tributaries, Holitna River, and Kogruklu River generally had increasing contributions to the run moving from the first to last time strata (Table 9). Finally, of the larger tributary groups, Chinook salmon from the Hoholitna River contributed the most consistent proportions to the run, contributing relatively high proportions to the run across all five time strata (Table 9).

### *Sensitivity Analysis*

We evaluated how changes in our prior distributions for the variance on the index of abundance ( $\sigma_{CPUE}$ ) affected estimates of abundance within each time strata (Figure 5) and for each year. Estimates showed that increasing  $\sigma_{CPUE}$  from 0.1 up to 0.2 or decreasing  $\sigma_{CPUE}$  down to 0.05 had virtually no effect on the estimated abundance within each time strata across years (Figure 5). Similarly, such changes in  $\sigma_{CPUE}$  also had virtually no effect on overall abundance estimated for each year.

Finally, we altered the assumed variance on  $\gamma[i,j]$  (used to create the Dirichlet distribution of expected frequencies of Chinook salmon observed in each tributary group during each time strata) to investigate how a 10-fold increase in the variance of the prior distribution (Table 6) would affect annual estimates of total abundance. Increasing the variance of the prior distribution resulted in increased estimates in total abundance in most years (Figure 6), with increases of 14%, 8%, 15% and 36% compared to the base BART model during 2004-2007, respectively. Despite these increases, estimated total abundance in 2004-2006 was still below that estimates from Schaberg et al. (2012). In 2007, increasing the variance of the prior distribution resulted in an estimated total abundance that exceeded Schaberg et al. (2012) by 38%. In 2003, estimated population abundance was similar using either prior distribution for  $\gamma[i,j]$ .

## DISCUSSION

Overall, across the five years of this study, Chinook salmon from upstream tributary groups tended to complete their migrations earlier than those from middle and Holitna tributary groups in the Kuskokwim River Basin. This finding supports those from other Alaskan river systems that Chinook salmon with the furthest distance to travel initiate their upriver migrations earlier than those with shorter distances to travel (Wuttig and Evenson 2001; Clark et al. 2015). While this trend was clearly evident from the run timing proportions from each tributary group (Tables 8a-8b; Figure 3), the proportional composition of the run (calculated using the estimated abundance of each tributary group in each time strata; Figure 4) indicated that even when these upper tributary groups had higher run timing proportions than middle or Holitna tributary groups (22 June and earlier; Table 8a-8b; Figure 3) they were less abundant at the tagging site than Chinook salmon from middle or Holitna tributary groups (Table 10a-10b; Figure 4). Potentially, the presence of Chinook salmon from the productive middle and Holitna drainages during the early run could serve to alleviate harvest pressure on upper tributary stocks during this time. Analysis of run timing of tributary groups on a finer scale (e.g., daily) than conducted in the BART model (~ weekly) may better indicate the extent to which upper tributary groups co-occur with spawners from middle and Holitna tributary groups during the early run.

Our results also indicate that there were annual differences in run timing that affect all tributary groups in a given year. For example, runs of virtually all tributary groups in 2006 and 2007 were delayed compared to runs in 2003-2005. This could indicate that large-scale regional environmental variables (i.e., temperature, ice out data) may affect run timing of all Chinook salmon tributary groups, as has been documented in other systems (Keefer et al. 2008; Mundy and Evenson 2011). Future research investigating the relationship between environmental variables and Chinook salmon run timing in the Kuskokwim drainage could benefit managers by providing estimates of the proportion of the run that has already passed a given location on a given date (Adkison and Cunningham 2015).

It is important to note that this study quantifies only trends in Chinook salmon escapement upstream from the tagging site at Birch Tree Crossing. During 2003-2007, an average of 88% of subsistence Chinook harvest occurred downstream of the tagging site (Brown et al. 2005; Fall et al. 2007a,b; Fall et al. 2009a, b). Therefore, our results may not be representative of the actual stock structure before harvest. Based on our results, it seems possible that upper tributary groups may experience higher exploitation during the early portion of the run than middle or Holitna tributary groups, which have a larger proportion of their tributary groups running later in the season. Previously, Linderman and Bergstrom (2009) reported that an average of 76% of the subsistence harvest for Chinook salmon occurs during the first half of the run. Comparing results of our study with similar mark-recapture telemetric studies conducted during recent years when early-season subsistence harvest restrictions were in place (e.g., 2014, 2015) may help determine if downstream harvest disproportionately targets Chinook salmon from upstream tributary groups.

Estimated total abundance from the base BART model had similar annual trends to those estimated using traditional Lincoln-Petersen mark-recapture methods in Schaberg et al. (2012). The BART model estimates, however, were lower in most years (2003-2006) than those from Schaberg et al. (2012) (Figure 1). As we used the same final location information from Schaberg et al. (2012) (see Table 2 in Schaberg et al. 2012), differences in estimated abundance are likely due to model structure (Bayesian methods versus traditional Lincoln-Petersen analysis). Furthermore, rerunning the BART model using the exact weir count data from Schaberg et al. (2012) (Table 3) produced very similar estimates to our base BART model, which used updated weir counts from Hansen and Blain (2014; Table 3). As the BART model directly incorporates temporal differences in capture probability by using temporal stratification, it seems likely that this temporal stratification may be responsible for such consistent differences in abundance estimates across years.

The one year in which abundance estimates were similar between the base BART model and Schaberg et al. (2012) was 2007 (Figure 1). In this year, the count data from the Kogruklu River weir was used in Schaberg et al. (2012) but was not used in the base BART model (Table 3). Of the 5 weir projects used in our analysis, the Kogruklu weir counts had, by far, the most Chinook salmon on an annual basis (Table 3). Omitting the Kogruklu weir counts in the base BART model for 2007 reduced the overall ratio of marked to unmarked fish, which increased the  $q$  parameter (constant of proportionality between total estimated abundance and gillnet CPUE in each time strata), relative to when the Kogruklu weir data was included. This led to an increase in our estimated abundance in the 2007 base BART model.

When we shared information on the proportions of each tributary group present in each time strata across years, estimated abundance tended to increase (Figure 1). However, these increases were relatively modest in most years (2003-2006; Figure 1), and still below abundance estimates from Schaberg et al. (2012). The exception was 2007, when shared run timing increased the abundance estimate by 33% over the based BART model. In 2007, there were rather large differences in the run timing evident across all tributary groups. Compared to other years, 2007 was a late, compressed run in which all tributary groups were late arriving and then abundance quickly declined after a peak during the 2<sup>nd</sup> and 3<sup>rd</sup> time strata (Tables 8a-8b; Figure 4). Sharing proportions of each tributary groups present in each time strata across years caused proportions for 2007 to increase during the first and later time strata, which did not match gillnet catch or CPUE data. The combination of lacking Kogruklu weir data and changes in the proportions of each tributary group present in each time strata appeared to dramatically alter the estimated abundance for 2007 using the base BART with shared run timing information (Figure 1).

In conclusion, our results offer a fine-scale view of run timing and abundance for Kuskokwim River Chinook salmon tributary groups. While our results only represent trends in escapement after the majority of harvest has occurred, valuable insights can still be drawn from our results. For example, it appears that specific tributary groups have consistent trends in run timing across years. Upriver tributary groups were present in larger proportions during the early part of the run and decline later in the run. These upriver tributary groups may be more vulnerable to exploitation than other tributary groups (middle and Holitna tributary groups) that are present in larger proportions later in the run. Total abundance estimates from the BART model, while similar to trends estimated previously using traditional mark-recapture methods (Schaberg et al. 2012), are more robust to temporal differences in capture probabilities. Additionally, the BART model estimates abundance on a finer spatial scale than previous studies by using Bayesian methods to distribute overall abundance across tributary groups using the spatial distribution of final locations from radio-telemetry surveys. This fine-scale information on run timing and abundance should be useful to ongoing management efforts aimed at preserving the diversity of Chinook salmon spawning stocks within the Kuskokwim River Drainage.

## REFERENCES

- Adkison, M.D. and C.J. Cunningham. 2015. The effects of salmon abundance and run timing on the performance of management by emergency order. *Canadian Journal of Fisheries and Aquatic Sciences*. 72: 1518-1526.
- Brown, C. L., and coauthors. 2005. Alaska subsistence salmon fisheries 2003 annual report. Alaska Department of Fish and Game, Technical Paper No. 316, Anchorage.
- Clark, S. C., T. L. Tanner, S. A. Sethi, K. T. Bentley, and D. E. Schindler. 2015. Migration timing of adult Chinook salmon into the Togiak River, Alaska, watershed: is there evidence for stock structure? *Transactions of the American Fisheries Society* 144:829-836.
- Costello, D. J., D. B. Molyneaux, and C. Goods. 2008. Takotna River salmon studies, 2007. Alaska Department of Fish and Game, Fishery Data Series 08-38, Anchorage.

- Eiler, J. H. 1995. A remote satellite-linked tracking system for studying Pacific salmon with radio telemetry. *Transactions of the American Fisheries Society* 124:184-193.
- Evans, M., N. Hastings, and B. Peacock. 1993. *Statistical distributions*. 2nd ed. Wiley and Sons, New York.
- Fall, J. A., and coauthors. 2007a. Alaska subsistence salmon fisheries 2004 annual report. Alaska Department of Fish and Game, Technical Paper No. 317, Anchorage.
- Fall, J. A., and coauthors. 2007b. Alaska subsistence salmon fisheries 2005 annual report. Alaska Department of Fish and Game, Technical Paper No. 318, Anchorage.
- Fall, J. A., and coauthors. 2009a. Alaska subsistence salmon fisheries 2006 annual report. Alaska Department of Fish and Game, Technical Paper No. 344, Anchorage.
- Fall, J. A., and coauthors. 2009b. Alaska subsistence salmon fisheries 2007 annual report. Alaska Department of Fish and Game, Technical Paper No. 346, Anchorage.
- Fleischman, S. J., M. J. Catalano, R. A. Clark, D. R. Bernard, and Y. Chen. 2013. An age-structured state-space stock-recruit model for Pacific salmon (*Oncorhynchus spp.*). *Canadian Journal of Fisheries and Aquatic Sciences* 70:401-414.
- Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2004. *Bayesian data analysis*. Second ed. CRC/Chapman & Hall, Boca Raton, Florida.
- Griffiths, J. R., and coauthors. 2014. Performance of salmon fishery portfolios across western North America. *Journal of Applied Ecology* 51:1554-1563.
- Hansen, T. R., and B. J. Blain. 2014. Salmon escapement monitoring in the Kuskokwim River, 2013. Alaska Department of Fish and Game, Fishery Data Series No. 14-54, Anchorage.
- Hilborn, R., T. P. Quinn, D. E. Schindler, and D. E. Rogers. 2003. Biocomplexity and fisheries sustainability. *Proceedings of the National Academy of Sciences* 100:6564-6568.
- Keefer, M. L., and coauthors. 2004. Stock-specific migration timing of adult spring–summer Chinook salmon in the Columbia River Basin. *North American Journal of Fisheries Management* 24:1145-1162.
- Keefer, M.L., C.A. Peery, and C.C. Caudill. 2008. Migration timing of Columbia River spring Chinook salmon: effects of temperature, river discharge, and ocean environment. *Transactions of the American Fisheries Society*. 137: 1120-1133.
- Linderman, J.C., Jr. and D.J. Bergstrom. 2009. Kuskokwim management area: salmon escapement, harvest, and management. Pages 541-599 in C.C. Krueger and C.E. Zimmerman, ed. *Pacific salmon: ecology and management of western Alaska's populations*. American Fisheries Society, Symposium 70, Bethesda, Maryland.
- Mundy, P.R. and D.F. Evenson. 2011. Environmental controls of phenology of high-latitude Chinook salmon populations of the Yukon River, North America, with application to fishery management. *ICES Journal of Marine Science*. 68: 1155-1164.
- Plummer, M. 2013. JAGS Version 3.4.0 user manual.
- R Development Core Team. 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Version 3.1.0. Vienna, Austria. <http://www.R-project.org>.
- Reimer, A. M., A. St. Saviour, and S. J. Fleischman. 2016. Stock specific abundance and run timing of Chinook salmon in the Kenai River, 2007–2012. Alaska Department of Fish and Game, Fishery Manuscript No. 16-01, Anchorage.
- Schaberg, K. L., Z. W. Liller, D. B. Molyneaux, B. G. Bue, and L. Stuby. 2012. Estimates of total annual return of Chinook salmon to the Kuskokwim River, 2002-2007. Alaska Department of Fish and Game, Fishery Data Series No. 12-36, Anchorage.
- Schindler, D. E., and coauthors. 2010. Population diversity and the portfolio effect in an exploited species. *Nature* 465:609-612.
- Stewart, R., D. J. Costello, D. B. Molyneaux, and J. M. Thalhauser. 2008. Tatlawiksuk River salmon studies, 2007. Alaska Department of Fish and Game, Fishery Data Series No. 08-59, Anchorage.

- Stuby, L. 2007. Inriver abundance of Chinook salmon in the Kuskokwim River, 2002-2007. Alaska Department of Fish and Game, Fishery Data Series No. 07-93, Anchorage.
- Su, Y. and M. Yajima. 2015. R2jags: Using R to run "JAGS". R package version 0.5.6. <http://CRAN.R-project.org/package=R2jags>.
- Thalhauser, J. M., D. J. Costello, R. Stewart, and D. B. Molyneaux. 2008. George River salmon studies, 2007. Alaska Department of Fish and Game, Fishery Data Series No. 08-63, Anchorage.
- Waples, R. S., and 9 coauthors. 2001. Characterizing diversity in salmon from the Pacific Northwest. *Journal of Fish Biology* 59:1-41.
- Williams, D. L., and C. A. Sheldon. 2010. Kogrukluk River weir salmon studies, 2008. Alaska Department of Fish and Game, Fishery Data Series No. 10-24, Anchorage.
- Wuttig, K. G., and M. J. Evenson. 2001. Inriver abundance, spawning distribution, and migratory timing of Copper River Chinook salmon in 2000. Alaska Department of Fish and Game, Fishery Data Series No. 01-22, Anchorage.



Table 1. Tributaries with the Kuskokwim River Basin that contained at least one final location for a radiotagged Chinook salmon during 2003-2007. The Composite Group and Tributary Group for each tributary. See Figure 1 for the location of each tributary. Tributary Groupings were used as the spatial component in the BART model.

<b>Tributary</b>	<b>Composite Group</b>	<b>Tributary Group</b>
Kuskokwim Mainstem	Mainstem	Kuskokwim Mainstem (MST)
Aniak River	Middle Tributaries	Aniak & Middle Tributaries (AMT)
Owhat River	Middle Tributaries	Aniak & Middle Tributaries (AMT)
Holokuk River	Middle Tributaries	Aniak & Middle Tributaries (AMT)
Sue Creek	Middle Tributaries	Aniak & Middle Tributaries (AMT)
Venah River	Middle Tributaries	Aniak & Middle Tributaries (AMT)
Oskawalik River	Middle Tributaries	Aniak & Middle Tributaries (AMT)
Salmon River (weir)	Middle Tributaries	Salmon River (SAL)
George River (weir)	Middle Tributaries	Geroge River (GEO)
Vreeland River	Holitna Drainage	Holitna River (HLT)
Holitna River	Holitna Drainage	Holitna River (HLT)
Kogruluk River (weir)	Holitna Drainage	Kogruluk River (KGK)
Hoholitna River	Holitna Drainage	Hoholitna River (HOH)
Stony River	Upper Tributaries	Stony River (STN)
Swift River	Upper Tributaries	Swift River (SWF)
Tatlawiksuk River(weir)	Upper Tributaries	Tatlawiksuk River (TLS)
Takotna River (weir)	Upper Tributaries	Takotna River (TAK)
Selatna River	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
Black River	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
Mainstem Headwaters	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
Blackwater River	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
Middle Fork	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
Big River	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
Pitka Fork	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
South Fork	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)
East Fork	Upper Tributaries	Upper Kuskokwim Tributaries (UKT)

Table 2. Tributaries within the Kuskokwim River Basin that contained at least one final location for a radiotagged Chinook salmon during 2003-2007. Tributary Group was used as the spatial component in the BART model. See Figure 1 for the location of each drainage in the Kuskokwim River Basin.

<b>Tributary/Watershed</b>	<b>Tributary Group</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>
Total Tagged		488	381	449	506	340
Downstream of Tagging Site		28	63	39	43	16
Kuskokwim Mainstem	MST	80	66	85	54	27
Aniak River	AMT	74	80	49	102	50
Owhat River	AMT	-	-	-	-	1
Holokuk River	AMT	5	10	7	4	12
Sue Creek	AMT	-	1	-	-	-
Venah River	AMT	1	-	-	-	-
Oskawalik River	AMT	7	2	8	7	8
Salmon River (weir)	SAL	7	5	4	9	8
George River (weir)	GEO	10	9	6	9	10
Vreeland River	HLT	1	-	2	2	-
Holitna River	HLT	86	49	75	98	75
Kogruklu River (weir)	KGK	49	24	49	35	42
Hoholitna River	HOH	45	35	44	36	22
Stony River	STN	7	7	23	39	19
Swift River	SWF	32	17	24	32	25
Tatlawiksuk River (weir)	TLS	15	5	12	7	5
Takotna River (weir)	TAK	2	1	1	0	0
Selatna River	UKT	-	-	-	1	1
Black River	UKT	-	-	1	-	-
Mainstem Headwaters	UKT	16	3	5	11	4
Takotna River (not past weir)	UKT	4	-	2	1	7
Blackwater River	UKT	2	-	2	3	2
Middle Fork	UKT	1	-	4	4	-
Big River	UKT	-	1	1	3	-
Pitka Fork	UKT	14	3	5	5	5
South Fork	UKT	1	-	-	-	1
East Fork	UKT	1	-	1	1	-
Total Final Locations		460	318	410	463	324

Table 3. Total escapement at each of 5 weir projects in the Kuskokwim River Basin during 2003-2007. Total escapement from weir projects is presented as reported in Schaberg et al. (2012) and Hansen and Blain (2014). Total escapement from Hansen and Blain (2014) was used in the base BART model.

Year	Salmon River (Aniak)	George River	Kogrukluk River	Tatlawiksuk River	Takotna River
Schaberg et al. (2012)					
2003	-	4,693	11,771	-	378
2004	-	5,207	19,651	2,833	461
2005	-	3,845	22,000	2,920	499
2006	6,732	4,357	19,414	1,700	539
2007	6,220	4,883	13,029	2,061	418
Hansen and Blain (2014)					
2003	-	-	12,008	-	390
2004	-	5,488	21,819	2,833	461
2005	-	3,845	21,819	2,864	499
2006	-	4,355	20,205	1,700	541
2007	6,220	4,011	-	2,032	412

Table 4. Gillnet catch-per-unit-effort (CPUE) from collections of Chinook salmon for radiotagging in the Kuskokwim River, Alaska during 2003-2007. CPUE data is presented as the number of Chinook salmon captured per 15 minutes of soak time and is partitioned into five weekly time strata used in the BART model.

<b>Year</b>	<b>15 June and earlier</b>	<b>16-22 June</b>	<b>23-29 June</b>	<b>30 June - 6 July</b>	<b>7 July and later</b>
2003	0.40	1.08	1.04	0.80	0.20
2004	0.26	0.98	1.25	0.43	0.24
2005	0.29	0.68	0.88	0.97	0.19
2006	0.07	0.63	0.95	0.94	0.42
2007	0.01	0.84	0.87	0.46	0.13

Table 5. The number of final locations of radiotagged Chinook salmon by tributary group and time strata in the Kuskokwim River Basin during 2003-2007. Spatially and temporally stratified counts of final locations were used in the BART model.

Year	Time Strata	Tributary Group											
		MST	AMT	SAL	GEO	HOH	HLT	KGK	STN	SWF	TLS	UKT	TAK
2003	15 June & earlier	23	6	1	1	6	13	6	1	4	3	14	0
	16 - 22 June	22	22	1	1	16	24	11	1	9	5	8	0
	23 - 29 June	13	30	3	4	7	21	10	2	11	2	11	1
	30 June - 6 July	12	18	1	3	12	15	18	3	3	2	1	1
	7 July & later	10	11	1	1	4	14	4	0	5	3	5	0
2004	15 June & earlier	13	9	0	1	4	1	1	0	3	3	2	1
	16 - 22 June	16	18	3	2	11	12	3	1	8	0	3	0
	23 - 29 June	14	23	0	4	12	18	9	4	3	2	2	0
	30 June - 6 July	4	15	1	0	3	4	9	1	0	0	0	0
	7 July & later	19	28	1	2	5	14	2	1	3	0	0	0
2005	15 June & earlier	22	15	2	1	14	12	14	3	8	2	10	1
	16 - 22 June	17	12	1	2	9	17	8	7	15	6	5	0
	23 - 29 June	8	13	0	1	4	15	11	3	0	2	2	0
	30 June - 6 July	12	12	0	1	11	17	9	6	1	1	2	0
	7 July & later	26	12	1	1	6	16	7	4	0	1	2	0
2006	15 June & earlier	4	1	0	0	2	2	0	2	0	0	12	0
	16 - 22 June	12	9	5	0	5	13	6	4	7	2	7	0
	23 - 29 June	19	57	3	3	14	29	15	13	12	2	7	0
	30 June - 6 July	7	25	1	3	9	28	9	11	10	1	2	0
	7 July & later	12	21	0	3	6	28	5	9	3	2	1	0
2007	15 June & earlier	0	0	0	0	0	1	0	0	0	0	0	0
	16 - 22 June	4	8	1	0	3	7	5	1	5	1	2	0
	23 - 29 June	11	31	3	3	7	29	20	8	10	2	10	0
	30 June - 6 July	5	20	3	4	6	16	10	5	4	0	6	0
	7 July & later	7	12	1	3	6	22	7	5	6	2	2	0

Table 6. Prior distributions for model parameters and alternate priors tested for sensitivity. In the BART mode  $i = 12$  tributary groups and  $j = 5$  time strata.

Text	JAGS	Prior	Alternate Prior(s)
$N$	N	$N \sim \text{"Uniform"} (\infty, \infty)^{\dagger}$	
$q$	q	$q \sim \text{Beta} (1, 1)$	
$\gamma$	D	$\gamma[i,j] \sim \text{Gamma} (1, 0.1)$	$\gamma[i,j] \sim \text{dgamma} (0.1, 0.01)$
$\rho$	prop	$\rho \sim \text{Beta} (1, 1)$	
$\alpha$	alpha	$\alpha[i,j] \sim \text{Gamma} (0.001, 0.001)$	
$\sigma_{\text{CPUE}}$	sigma.index	$\sigma_{\text{CPUE}} \sim 1/V \text{ Gamma} (100, 1)$	$\sigma_{\text{CPUE}} \sim 1/V \text{ Gamma} (25, 1)$ $\sigma_{\text{CPUE}} \sim 1/V \text{ Gamma} (50, 1)$ $\sigma_{\text{CPUE}} \sim 1/V \text{ Gamma} (200, 1)$ $\sigma_{\text{CPUE}} \sim 1/V \text{ Gamma} (400, 1)$

<sup>†</sup> Where "Uniform" is in quotes, a normal distribution with mean 0 and extremely large variance was used. These uninformative normal distributions were designed to be equivalent to uniform distributions over the range of the likelihood. Normal distributions were less likely to cause computational disruptions during MCMC sampling.

Table 7. Total Abundance by time strata for each year (2003-2007) estimated from the base BART model. Values are presented as mean (SD).

<b>Time Strata</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>
15 June & earlier	13,163 (2,495)	15,736 (5,846)	15,209 (5,754)	5,224 (4,354)	691 (1,342)
16 - 22 June	32,140 (2,833)	53,461 (9,225)	32,562 (10,533)	42,063 (4,426)	46,868 (2,715)
23 - 29 June	31,263 (2,518)	66,145 (5,162)	41,815 (6,305)	60,197 (7,024)	48,318 (1,761)
30 June - 6 July	24,791 (924)	25,923 (6,935)	45,246 (9,526)	59,333 (9,314)	28,159 (4,797)
7 July & later	6,933 (160)	14,919 (10,036)	10,156 (10,310)	28,655 (6,104)	8,463 (1,919)
<b>Total</b>	108,290 (17,599)	176,184 (25,967)	144,988 (19,300)	195,472 (27,002)	132,499 (132,498)

Table 8a. Run timing proportions ( $p_{ij}$ ) for Kuskokwim River Chinook salmon at Birch Tree Crossing (rkm 294) for tributary groups in the mainstem, middle and Holitna composite groups during 2003-2007. Values are presented as means (SD).

Year	Time Strata	Tributary Group						
		Mainstem (MST)	Aniak (AMT)	Salmon (SAL)	George (GEO)	Hoholitna (HOH)	Holitna (HLT)	Kogrukluuk (KGK)
2003	15 June & earlier	0.21 (0.03)	0.05 (0.01)	0.12 (0.08)	0.09 (0.06)	0.09 (0.03)	0.10 (0.02)	0.08 (0.03)
	16 - 22 June	0.33 (0.05)	0.27 (0.04)	0.19 (0.11)	0.14 (0.09)	0.38 (0.07)	0.31 (0.05)	0.24 (0.06)
	23 - 29 June	0.20 (0.04)	0.37 (0.05)	0.39 (0.14)	0.38 (0.12)	0.18 (0.05)	0.28 (0.05)	0.22 (0.05)
	30 June - 6 July	0.18 (0.04)	0.23 (0.04)	0.19 (0.11)	0.30 (0.12)	0.29 (0.06)	0.20 (0.04)	0.39 (0.06)
	7 July & later	0.06 (0.01)	0.05 (0.01)	0.08 (0.05)	0.06 (0.04)	0.04 (0.02)	0.07 (0.01)	0.04 (0.01)
2004	15 June & earlier	0.15 (0.03)	0.07 (0.02)	0.07 (0.06)	0.10 (0.07)	0.08 (0.03)	0.02 (0.01)	0.03 (0.02)
	16 - 22 June	0.33 (0.06)	0.26 (0.04)	0.49 (0.15)	0.26 (0.12)	0.35 (0.07)	0.30 (0.06)	0.16 (0.07)
	23 - 29 June	0.31 (0.06)	0.35 (0.05)	0.13 (0.11)	0.46 (0.13)	0.40 (0.08)	0.46 (0.07)	0.43 (0.09)
	30 June - 6 July	0.08 (0.03)	0.19 (0.03)	0.22 (0.13)	0.07 (0.07)	0.10 (0.04)	0.10 (0.04)	0.33 (0.08)
	7 July & later	0.11 (0.02)	0.11 (0.02)	0.07 (0.05)	0.08 (0.05)	0.05 (0.02)	0.10 (0.02)	0.02 (0.01)
2005	15 June & earlier	0.12 (0.02)	0.09 (0.02)	0.18 (0.10)	0.07 (0.04)	0.13 (0.03)	0.06 (0.01)	0.10 (0.02)
	16 - 22 June	0.22 (0.04)	0.17 (0.04)	0.26 (0.15)	0.25 (0.12)	0.19 (0.05)	0.19 (0.04)	0.14 (0.04)
	23 - 29 June	0.22 (0.05)	0.36 (0.06)	0.23 (0.16)	0.31 (0.16)	0.19 (0.07)	0.33 (0.06)	0.40 (0.07)
	30 June - 6 July	0.29 (0.06)	0.30 (0.06)	0.21 (0.16)	0.29 (0.15)	0.42 (0.08)	0.34 (0.06)	0.30 (0.07)
	7 July & later	0.13 (0.02)	0.06 (0.01)	0.10 (0.08)	0.06 (0.04)	0.05 (0.02)	0.07 (0.01)	0.04 (0.01)
2006	15 June & earlier	0.03 (0.01)	0.00 (0.00)	0.02 (0.02)	0.03 (0.03)	0.03 (0.01)	0.01 (0.00)	0.00 (0.00)
	16 - 22 June	0.30 (0.06)	0.11 (0.03)	0.52 (0.13)	0.09 (0.08)	0.19 (0.06)	0.17 (0.04)	0.24 (0.07)
	23 - 29 June	0.29 (0.05)	0.43 (0.04)	0.22 (0.10)	0.26 (0.11)	0.31 (0.07)	0.24 (0.04)	0.31 (0.07)
	30 June - 6 July	0.18 (0.05)	0.29 (0.04)	0.17 (0.10)	0.37 (0.13)	0.32 (0.07)	0.36 (0.05)	0.33 (0.08)
	7 July & later	0.17 (0.04)	0.14 (0.02)	0.05 (0.04)	0.23 (0.10)	0.13 (0.04)	0.20 (0.03)	0.10 (0.04)
2007	15 June & earlier	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.01 (0.01)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	16 - 22 June	0.41 (0.11)	0.31 (0.07)	0.45 (0.16)	0.19 (0.14)	0.39 (0.12)	0.30 (0.07)	0.33 (0.09)
	23 - 29 June	0.34 (0.09)	0.39 (0.06)	0.25 (0.12)	0.31 (0.12)	0.28 (0.09)	0.39 (0.06)	0.41 (0.07)
	30 June - 6 July	0.16 (0.06)	0.23 (0.04)	0.23 (0.11)	0.37 (0.13)	0.23 (0.08)	0.20 (0.04)	0.20 (0.05)
	7 July & later	0.07 (0.02)	0.04 (0.01)	0.03 (0.02)	0.10 (0.05)	0.07 (0.03)	0.09 (0.02)	0.04 (0.01)



Table 8b. Run timing proportions ( $p_{ij}$ ) for Kuskokwim River Chinook salmon at Birch Tree Crossing (rkm 294) for tributary groups in the upper tributaries composite group during 2003-2007. Values are presented as means (SD).

Year	Time Strata	Tributary Group				
		Stony (STN)	Swift (SWF)	Tatlawiksuk (TLS)	Upper Kusko. Tribs. (UKT)	Takotna (TAK)
2003	15 June & earlier	0.11 (0.07)	0.09 (0.04)	0.15 (0.07)	0.25 (0.05)	0.25 (0.05)
	16 - 22 June	0.18 (0.11)	0.31 (0.07)	0.36 (0.11)	0.24 (0.06)	0.24 (0.06)
	23 - 29 June	0.28 (0.13)	0.38 (0.08)	0.18 (0.09)	0.35 (0.07)	0.35 (0.07)
	30 June - 6 July	0.37 (0.13)	0.12 (0.05)	0.18 (0.09)	0.08 (0.04)	0.08 (0.04)
	7 July & later	0.04 (0.04)	0.07 (0.03)	0.10 (0.05)	0.06 (0.02)	0.06 (0.02)
2004	15 June & earlier	0.05 (0.05)	0.12 (0.05)	0.33 (0.13)	0.22 (0.10)	0.22 (0.10)
	16 - 22 June	0.19 (0.11)	0.51 (0.10)	0.12 (0.10)	0.37 (0.13)	0.37 (0.13)
	23 - 29 June	0.51 (0.14)	0.24 (0.10)	0.37 (0.15)	0.29 (0.12)	0.29 (0.12)
	30 June - 6 July	0.16 (0.10)	0.04 (0.04)	0.11 (0.10)	0.08 (0.07)	0.08 (0.07)
	7 July & later	0.05 (0.04)	0.06 (0.03)	0.05 (0.05)	0.03 (0.03)	0.03 (0.03)
2005	15 June & earlier	0.05 (0.02)	0.15 (0.05)	0.10 (0.05)	0.23 (0.07)	0.23 (0.07)
	16 - 22 June	0.25 (0.08)	0.61 (0.09)	0.42 (0.11)	0.25 (0.09)	0.25 (0.09)
	23 - 29 June	0.24 (0.09)	0.07 (0.06)	0.25 (0.11)	0.23 (0.10)	0.23 (0.10)
	30 June - 6 July	0.38 (0.10)	0.13 (0.08)	0.16 (0.09)	0.22 (0.10)	0.22 (0.10)
	7 July & later	0.06 (0.02)	0.01 (0.01)	0.06 (0.04)	0.05 (0.03)	0.05 (0.03)
2006	15 June & earlier	0.02 (0.01)	0.01 (0.01)	0.05 (0.05)	0.18 (0.04)	0.18 (0.04)
	16 - 22 June	0.15 (0.06)	0.27 (0.07)	0.28 (0.12)	0.38 (0.09)	0.38 (0.09)
	23 - 29 June	0.27 (0.06)	0.27 (0.07)	0.24 (0.11)	0.24 (0.07)	0.24 (0.07)
	30 June - 6 July	0.36 (0.07)	0.36 (0.08)	0.18 (0.11)	0.13 (0.07)	0.13 (0.07)
	7 July & later	0.17 (0.05)	0.07 (0.03)	0.22 (0.11)	0.05 (0.03)	0.05 (0.03)
2007	15 June & earlier	0.00 (0.00)	0.00 (0.00)	0.02 (0.02)	0.00 (0.00)	0.00 (0.00)
	16 - 22 June	0.24 (0.13)	0.48 (0.11)	0.39 (0.17)	0.30 (0.12)	0.30 (0.12)
	23 - 29 June	0.40 (0.11)	0.31 (0.08)	0.34 (0.15)	0.40 (0.10)	0.40 (0.10)
	30 June - 6 July	0.25 (0.09)	0.13 (0.05)	0.10 (0.09)	0.24 (0.08)	0.24 (0.08)
	7 July & later	0.08 (0.03)	0.06 (0.02)	0.13 (0.07)	0.03 (0.02)	0.03 (0.02)

Table 9. Total abundance by tributary group in each year estimated with the base BART model. Values are presented as mean (SD).

Year	Tributary Group					
	Mainstem (MST)	Aniak (AMT)	Salmon (SAL)	George (GEO)	Hoholitna (HOH)	Holitna (HLT)
2003	17,041 (3,514)	20,337 (4,077)	2,465 (850)	3,220 (1,040)	10,958 (2,458)	19,154 (3,840)
2004	30,884 (6,341)	43,931 (8,505)	4,826 (1,830)	5,615 (541)	20,545 (4,727)	26,191 (5,614)
2005	24,019 (4,681)	22,894 (4,694)	2,371 (1,015)	3,837 (371)	15,304 (3,377)	28,393 (5,522)
2006	21,781 (4,395)	43,779 (7,631)	5,883 (1,872)	4,436 (423)	15,532 (3,416)	40,553 (7,185)
2007	11,701 (3,589)	27,438 (6,851)	6,086 (598)	4,039 (394)	9,531 (3,099)	25,707 (6,565)
Upper Kuskokwim						
	Kogrukluk (KGK)	Stony (STN)	Swift (SWF)	Tatlawiksuk (TLS)	Tribs. (UKT)	Takotna (TAK)
2003	11,943 (1,195)	2,638 (920)	7,784 (1,894)	3,965 (1,149)	8,389 (1,955)	396 (39)
2004	18,709 (1,753)	6,230 (2,177)	10,648 (2,938)	2,929 (285)	5,206 (1,886)	469 (47)
2005	20,623 (1,905)	9,674 (2,488)	7,791 (1,963)	3,067 (292)	6,512 (1,805)	503 (50)
2006	18,790 (1,752)	16,544 (3,557)	15,268 (3,434)	1,826 (180)	10,540 (2,564)	541 (54)
2007	17,224 (4,719)	7,524 (2,489)	11,985 (3,781)	2,092 (204)	8,757 (2,828)	415 (41)

Table 10a. Stock composition of the run ( $\theta_{ij}$ ) for Kuskokwim River Chinook salmon at Birch Tree Crossing (rkm 294) for tributaries group within each time strata during 2003-2005. See Table 1 for tributary group names and Table 5 for time strata dates. Values are presented as means (SD).

Strata Data: Values are presented as means (SD).												
Strata	MST	AMT	SAL	GEO	HOH	Tributary Group						
						HLT	KGK	STN	SWF	TLS	UKT	TAK
2003												
1	0.27 (0.05)	0.08 (0.03)	0.02 (0.02)	0.02 (0.02)	0.08 (0.03)	0.16 (0.04)	0.08 (0.03)	0.02 (0.02)	0.05 (0.02)	0.05 (0.02)	0.16 (0.04)	0.01 (0.00)
2	0.18 (0.03)	0.18 (0.03)	0.02 (0.01)	0.01 (0.01)	0.13 (0.03)	0.19 (0.03)	0.09 (0.03)	0.01 (0.01)	0.08 (0.02)	0.05 (0.02)	0.06 (0.02)	0 (0)
3	0.11 (0.03)	0.25 (0.04)	0.03 (0.02)	0.04 (0.02)	0.06 (0.02)	0.18 (0.03)	0.09 (0.03)	0.02 (0.01)	0.10 (0.03)	0.02 (0.01)	0.10 (0.03)	0 (0)
4	0.13 (0.03)	0.19 (0.04)	0.02 (0.01)	0.04 (0.02)	0.13 (0.03)	0.16 (0.04)	0.19 (0.04)	0.04 (0.02)	0.04 (0.02)	0.03 (0.02)	0.03 (0.02)	0 (0)
5	0.16 (0.04)	0.17 (0.05)	0.03 (0.02)	0.03 (0.02)	0.07 (0.03)	0.22 (0.05)	0.07 (0.03)	0.02 (0.01)	0.09 (0.03)	0.06 (0.03)	0.08 (0.03)	0 (0)
2004												
1	0.29 (0.06)	0.21 (0.06)	0.02 (0.02)	0.04 (0.03)	0.11 (0.04)	0.04 (0.03)	0.05 (0.03)	0.02 (0.02)	0.08 (0.04)	0.06 (0.03)	0.07 (0.03)	0.01 (0.00)
2	0.19 (0.04)	0.22 (0.04)	0.04 (0.02)	0.03 (0.01)	0.14 (0.04)	0.15 (0.04)	0.06 (0.03)	0.02 (0.02)	0.10 (0.03)	0.01 (0.01)	0.04 (0.02)	0 (0)
3	0.15 (0.03)	0.23 (0.04)	0.01 (0.01)	0.04 (0.01)	0.13 (0.03)	0.19 (0.04)	0.12 (0.03)	0.05 (0.02)	0.04 (0.02)	0.02 (0.01)	0.02 (0.01)	0 (0)
4	0.10 (0.04)	0.32 (0.07)	0.04 (0.03)	0.02 (0.02)	0.08 (0.04)	0.10 (0.04)	0.25 (0.06)	0.04 (0.03)	0.02 (0.02)	0.01 (0.01)	0.02 (0.02)	0 (0)
5	0.23 (0.05)	0.34 (0.05)	0.02 (0.02)	0.03 (0.02)	0.07 (0.03)	0.18 (0.04)	0.04 (0.02)	0.02 (0.02)	0.05 (0.02)	0.01 (0.01)	0.01 (0.01)	0 (0)
2005												
1	0.20 (0.04)	0.14 (0.03)	0.02 (0.02)	0.02 (0.01)	0.13 (0.03)	0.11 (0.03)	0.14 (0.03)	0.03 (0.02)	0.08 (0.02)	0.02 (0.01)	0.10 (0.03)	0.01 (0.00)
2	0.17 (0.04)	0.12 (0.03)	0.02 (0.01)	0.03 (0.02)	0.09 (0.03)	0.17 (0.04)	0.09 (0.03)	0.07 (0.02)	0.15 (0.03)	0.04 (0.01)	0.05 (0.02)	0 (0)
3	0.13 (0.04)	0.20 (0.05)	0.01 (0.01)	0.03 (0.02)	0.07 (0.03)	0.23 (0.05)	0.20 (0.05)	0.06 (0.03)	0.01 (0.01)	0.02 (0.01)	0.04 (0.02)	0 (0)
4	0.16 (0.04)	0.16 (0.04)	0.01 (0.01)	0.03 (0.01)	0.14 (0.04)	0.22 (0.04)	0.14 (0.04)	0.08 (0.03)	0.02 (0.02)	0.01 (0.01)	0.03 (0.02)	0 (0)
5	0.31 (0.05)	0.15 (0.04)	0.02 (0.02)	0.02 (0.02)	0.08 (0.03)	0.2 (0.04)	0.10 (0.03)	0.06 (0.02)	0.01 (0.01)	0.02 (0.01)	0.03 (0.02)	0 (0)

Table 10b. Stock composition of the run ( $\theta_{ij}$ ) for Kuskokwim River Chinook salmon at Birch Tree Crossing (rkm 294) for tributaries group within each time strata during 2006-2007. See Table 1 for tributary group names and Table 5 for time strata dates. Values are presented as mean (SD).

are presented as mean (SD).

Strata	MST	AMT	SAL	GEO	HOH	Tributary Group		STN	SWF	TLS	UKT	TAK
						HLT	KGK					
2006												
1	0.14 (0.06)	0.06 (0.04)	0.03 (0.03)	0.03 (0.03)	0.09 (0.05)	0.09 (0.05)	0.03 (0.03)	0.09 (0.05)	0.03 (0.03)	0.02 (0.02)	0.37 (0.08)	0.02 (0.01)
2	0.16 (0.04)	0.12 (0.04)	0.08 (0.03)	0.01 (0.01)	0.07 (0.03)	0.17 (0.04)	0.11 (0.04)	0.06 (0.03)	0.10 (0.03)	0.01 (0.01)	0.10 (0.03)	0.01 (0.00)
3	0.11 (0.02)	0.31 (0.03)	0.02 (0.01)	0.02 (0.01)	0.08 (0.02)	0.16 (0.03)	0.10 (0.02)	0.08 (0.02)	0.07 (0.02)	0.01 (0.00)	0.04 (0.01)	0 (0)
4	0.07 (0.02)	0.22 (0.04)	0.02 (0.01)	0.03 (0.01)	0.08 (0.03)	0.25 (0.04)	0.11 (0.03)	0.10 (0.03)	0.09 (0.03)	0.01 (0.00)	0.02 (0.01)	0 (0)
5	0.13 (0.03)	0.22 (0.04)	0.01 (0.01)	0.04 (0.02)	0.07 (0.03)	0.29 (0.05)	0.07 (0.03)	0.10 (0.03)	0.04 (0.02)	0.01 (0.01)	0.02 (0.01)	0 (0)
2007												
1	0.08 (0.08)	0.08 (0.08)	0.09 (0.08)	0.08 (0.08)	0.09 (0.08)	0.17 (0.1)	0.08 (0.08)	0.09 (0.08)	0.08 (0.08)	0.08 (0.07)	0.08 (0.07)	0 (0)
2	0.10 (0.04)	0.19 (0.06)	0.06 (0.03)	0.02 (0.01)	0.08 (0.04)	0.17 (0.05)	0.13 (0.05)	0.04 (0.03)	0.13 (0.05)	0.02 (0.01)	0.06 (0.03)	0 (0)
3	0.08 (0.02)	0.22 (0.03)	0.03 (0.01)	0.03 (0.01)	0.06 (0.02)	0.21 (0.03)	0.14 (0.03)	0.06 (0.02)	0.08 (0.02)	0.02 (0.01)	0.07 (0.02)	0 (0)
4	0.07 (0.03)	0.23 (0.04)	0.05 (0.02)	0.05 (0.02)	0.08 (0.03)	0.19 (0.04)	0.12 (0.03)	0.07 (0.03)	0.06 (0.02)	0.01 (0.01)	0.07 (0.03)	0 (0)
5	0.1 (0.03)	0.16 (0.04)	0.03 (0.02)	0.05 (0.02)	0.08 (0.03)	0.27 (0.05)	0.1 (0.03)	0.07 (0.03)	0.08 (0.03)	0.03 (0.02)	0.03 (0.02)	0 (0)

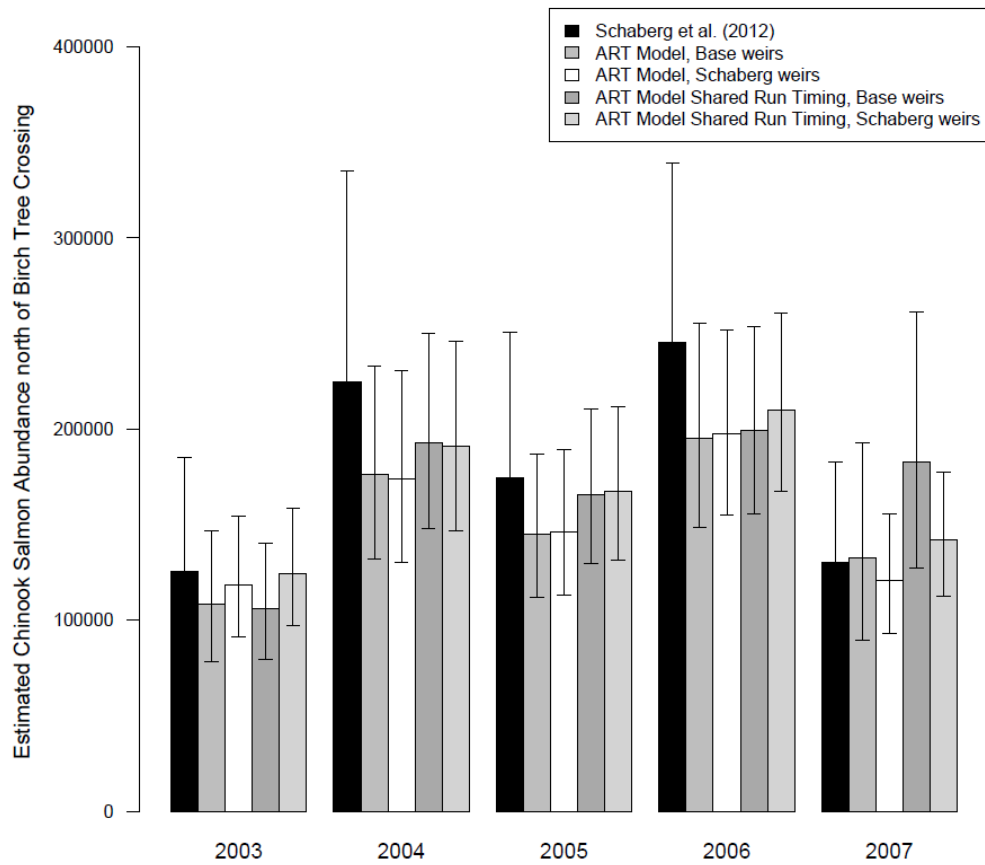


Figure 1. Estimated Chinook Salmon abundance north of Birch Tree Crossing. Estimates from Schaberg et al. (2012) are from a previously conducted mark-recapture study using the same data.

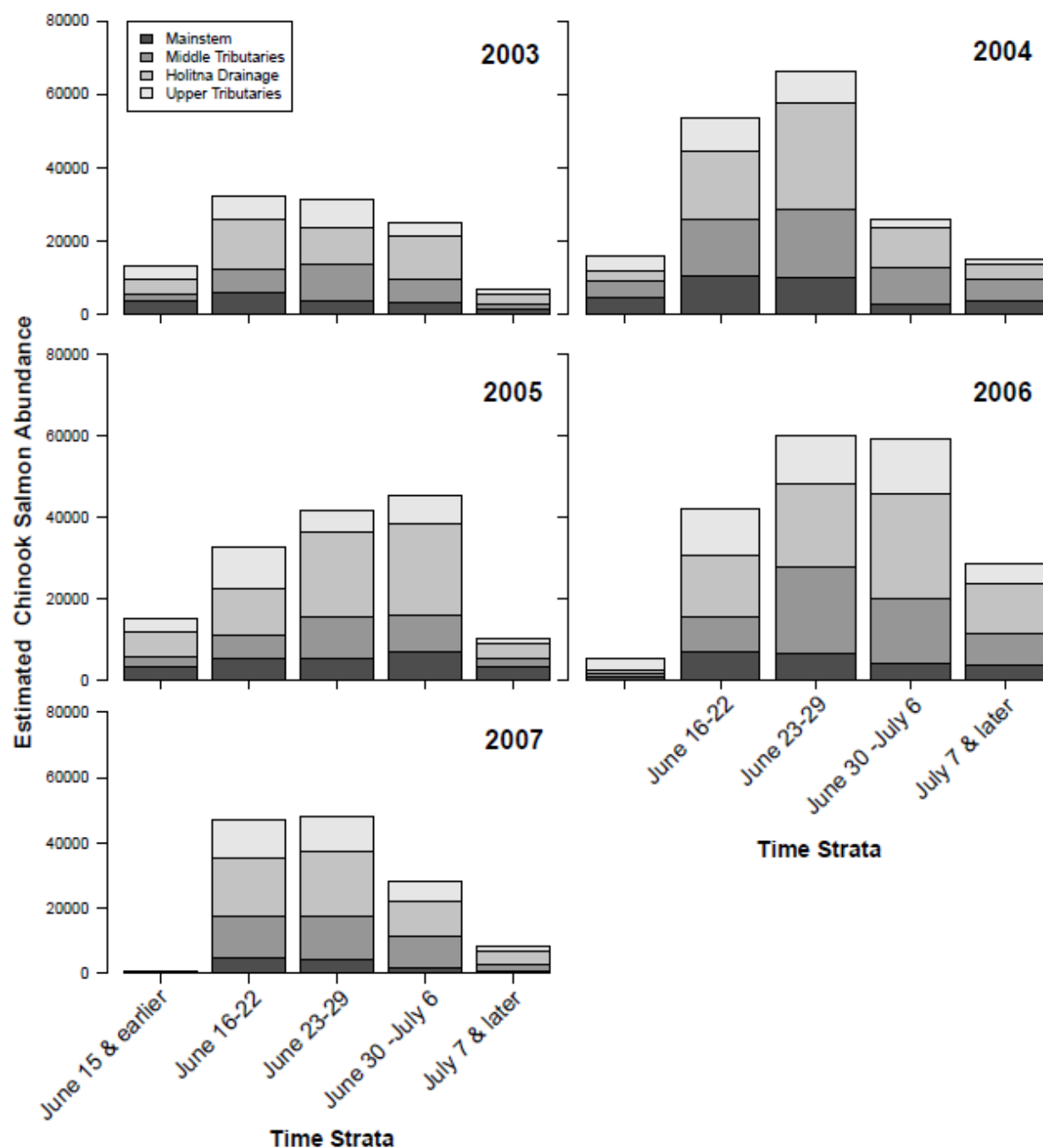


Figure 2. Estimated Chinook Salmon abundance for each of four large, composite tributary groups during 2003-2007 by time strata. All estimates were generated under the Bayesian ART Model using the base weir data with no shared run timing across years. (See Table 2 for tributaries that comprise each group.)

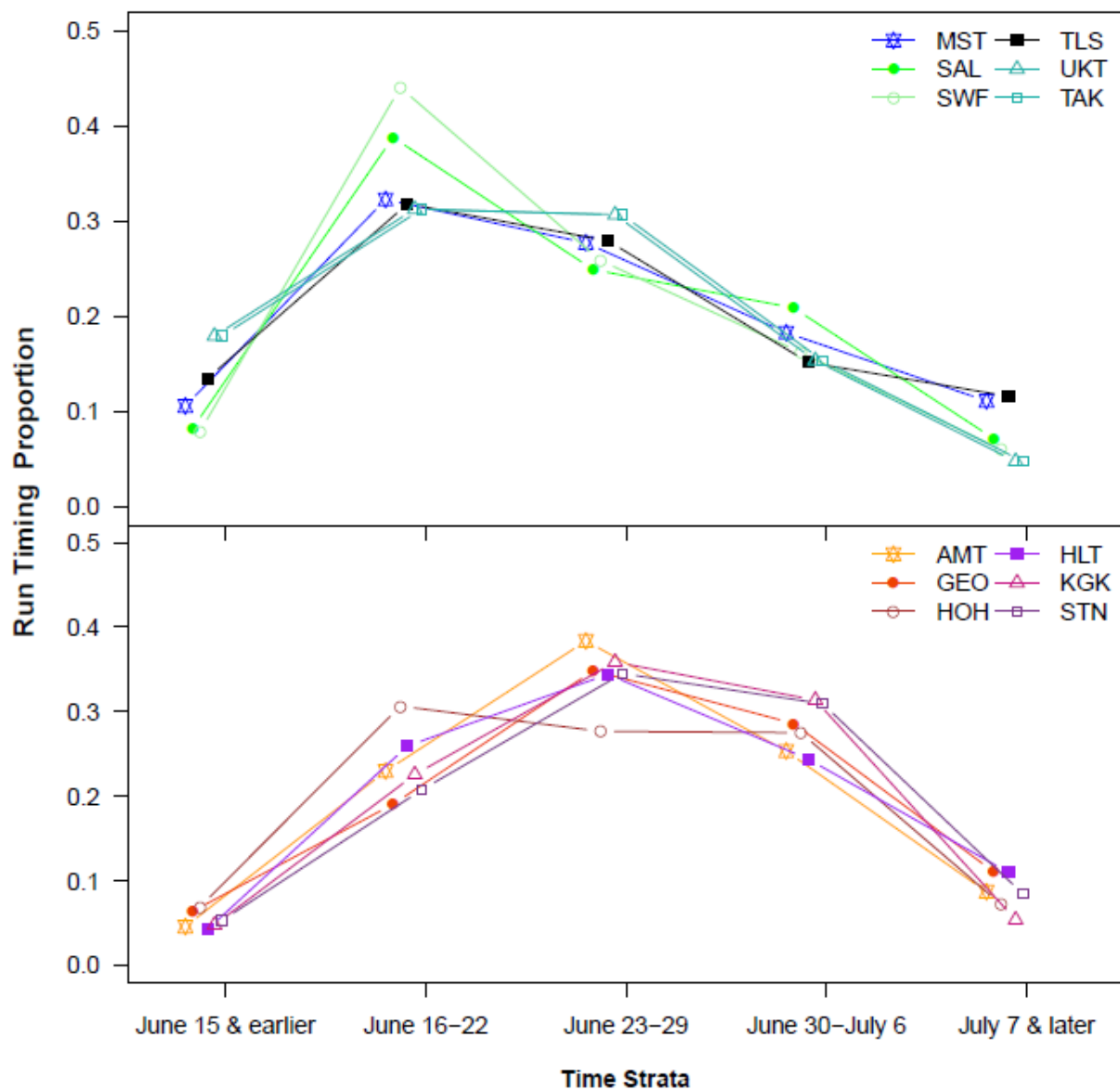


Figure 3. Mean run timing proportions from the base BART model for each of 12 tributary groups across time strata. Run timing proportions for each tributary group in each time strata were averaged across years (2003-2007). See Table 1 for tributary group abbreviations.

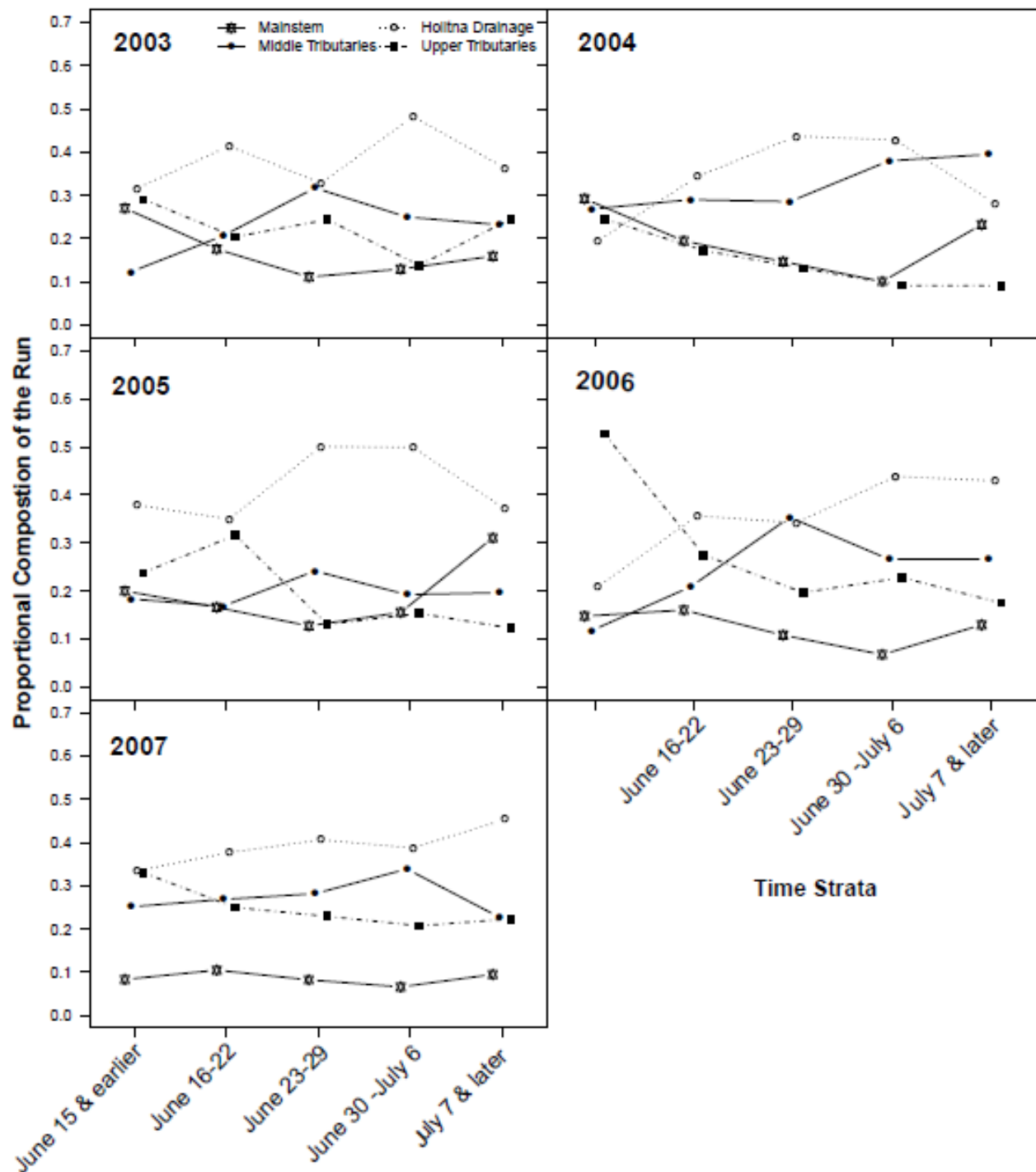


Figure 4. Proportional composition of the run in each time strata during 2003-2007. Proportions are shown for four large composite tributary groups. See Table 1 for information on specific tributaries contained within each group.



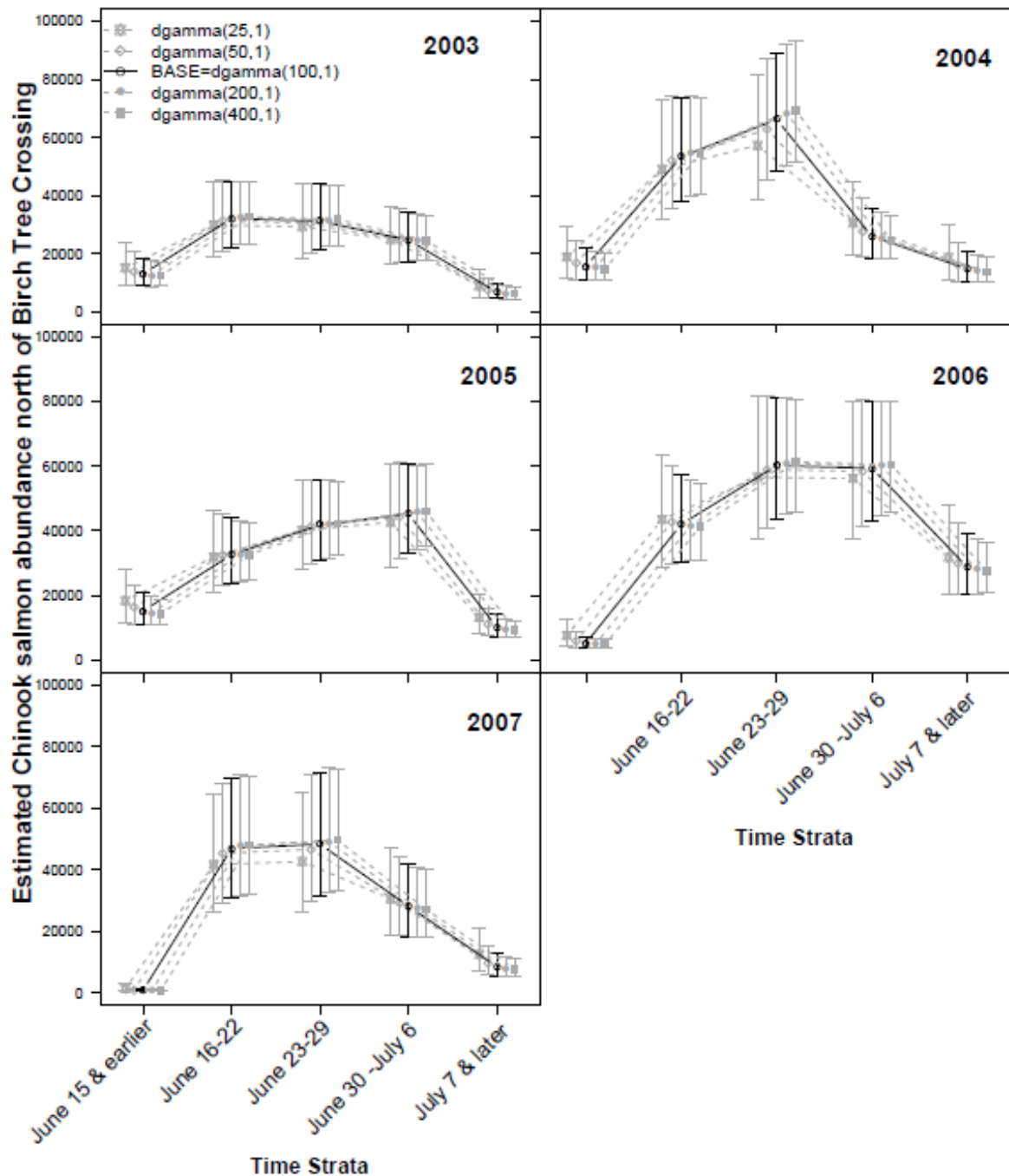


Figure 5. Estimated Chinook salmon abundance north of Birch Tree Crossing in response to differing prior distributions for the variance assigned to the abundance (i.e., gillnet CPUE) index for each time strata during 2003-2007. Individual panels represent years with individual. All estimates were generated under the BART Model using the base weir data (Hansen and Blain 2014) with no shared run timing across years.

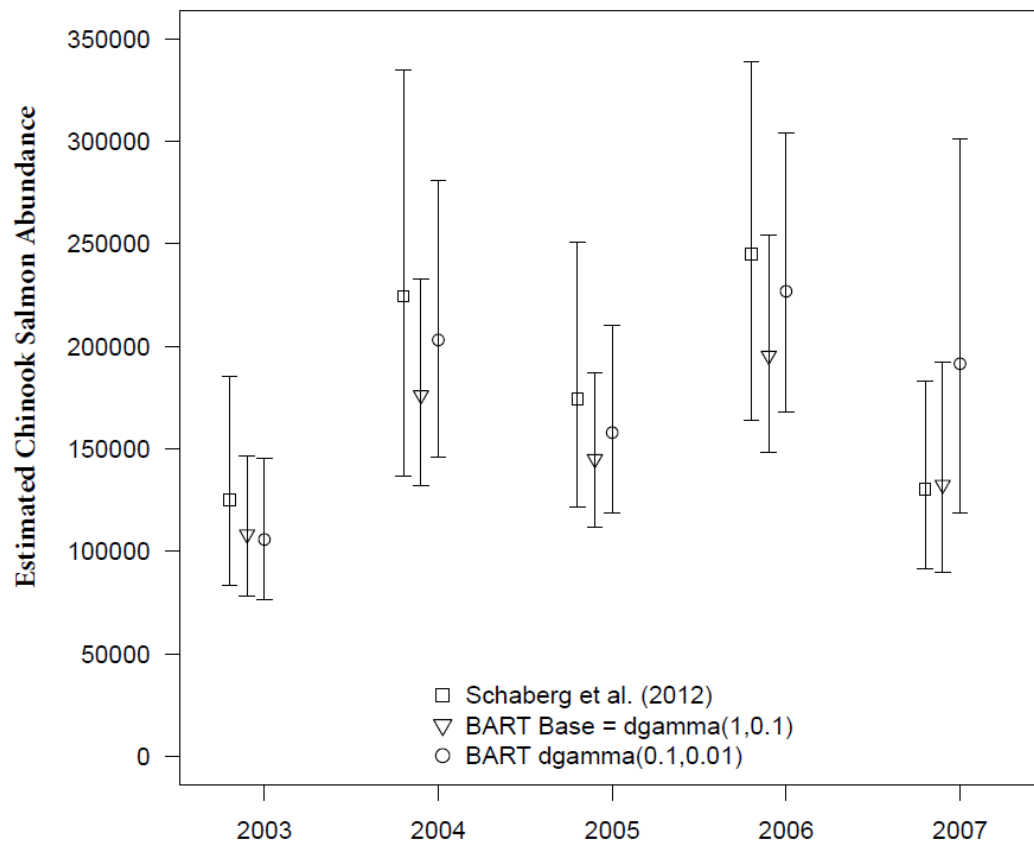


Figure 6. Estimated Chinook salmon abundance north of Birch Tree Crossing in response to changes in the prior Dirichlet distribution used for 2003-2007. All estimates were generated under the BART model using the base weir data with no shared run timing across years.

## Appendix A. Code for base BART model

### #### Import Data ####

#### #import total weir counts from each weir (from Schaberg et al. 2012)

```
log.Ni.hat = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x_matrix/data/2003_weirs_BASE.csv")
```

#### #import x matix for final locations

```
#order is: (1)MST, (2)AMT, (3)SAL, (4)GEO, (5)HOH, (6)HLT, (7)KGK, (8)STN, (9)SWF, (10)TLS, (11)UKT, (12)TAK
```

```
x = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x_matrix/data/2003_x_mat_12_grps_5_strata.csv")
```

#### #import log value for index of abundance (CPUE) in each time strata

```
log.index = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x_matrix/data/2003_log_index_5_strata.csv")
```

#### #providing values of coefficient of variation for abundance at each weir

```
cv.N = c(0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.1, 0.1)
```

#### #total number of chinook successfully tracked (i.e., fish was assigned to one of the 12 tribs. groups) during each time strata

```
n = array(0, dim=c(5, 1))  
n[,1] = c(sum(x[1,]), sum(x[2,]), sum(x[3,]), sum(x[4,]), sum(x[5,]))
```

### #### Specify Bayesian Analysis of Abundance and Run Timing (BART) model ####

```
#Write model
```

```
mod = function(){
```

```
  tau.index ~ dgamma(100,1)  
  sigma.index <- 1 / sqrt(tau.index)
```

```
  N ~ dnorm(0,1.0E-12) %_ % l(1,)   
  q ~ dbeta(1,1)
```

```
  D.sum <- sum(D[1:11,1:5])
```

```
  for(i in 1:10) {  
    N.i[i] <- sum(N.ij[i,1:5])  
    log.Ni.hat[i,1] ~ dnorm(log.Ni[i],tau.logNi[i])  
    log.Ni[i] <- log(N.i[i])  
    tau.logNi[i] <- 1 / log(cv.N[i] * cv.N[i] + 1)
```

```
  for(j in 1:5) {  
    D[i,j] ~ dgamma(1,0.1)  
    pi[i,j] <- D[i,j] / D.sum  
    N.ij[i,j] <- pi[i,j] * N  
    theta[j,i] <- N.ij[i,j] / N.j[j]  
    pi.Ni[i,j] <- N.ij[i,j] / N.i[i]  
  }
```

```

}

for(i in 11:11) {
  N.i[i] <- sum(N.ij[i,1:5])
  log.Ni.hat[i,1] ~ dnorm(log.Ni[i],tau.logNi[i])
  log.Ni[i] <- log(N.i[i])
  tau.logNi[i] <- 1 / log(cv.N[i] * cv.N[i] + 1)
  prop ~ dbeta(1,1)

  for(j in 1:5) {
    D[i,j] ~ dgamma(1,0.1)
    pi[i,j] <- (D[i,j]*prop) / D.sum
    N.ij[i,j] <- pi[i,j] * N
    theta[j,i] <- N.ij[i,j] / N.j[j]
    pi.Ni[i,j] <- N.ij[i,j] / N.i[i]
  }
}

for(i in 12:12) {
  N.i[i] <- sum(N.ij[i,1:5])
  log.Ni.hat[i,1] ~ dnorm(log.Ni[i],tau.logNi[i])
  log.Ni[i] <- log(N.i[i])
  tau.logNi[i] <- 1 / log(cv.N[i] * cv.N[i] + 1)

  for(j in 1:5) {
    pi[i,j] <- (D[i-1,j]*(1-prop)) / D.sum
    N.ij[i,j] <- pi[i,j] * N
    theta[j,i] <- N.ij[i,j] / N.j[j]
    pi.Ni[i,j] <- N.ij[i,j] / N.i[i]
  }
}

for(j in 1:5) {
  N.j[j] <- sum(N.ij[1:12,j])
  log.Nq[j] <- log(N.j[j] * q)
  log.index[j,1] ~ dnorm(log.Nq[j], tau.index)

  x[j,1:12] ~ dmulti(theta[j,1:12],n[j,1])
}

```

#### **#derived quantities**

```

N.MST <- sum(N.i[1])
N.MKT <- sum(N.i[2:4])
N.HLT <- sum(N.i[5:7])
N.UKT <- sum(N.i[8:12])

for (j in 1:5){
  N.MST.j[j] <- sum(N.ij[1,j])

```

```

    N.MKT.j[j] <- sum(N.ij[2:4,j])
    N.HLT.j[j] <- sum(N.ij[5:7,j])
    N.UKT.j[j] <- sum(N.ij[8:12,j])
  }
}

```

```

model.file = "model.txt"
write.model(mod, model.file)

```

### **#bundle data to send to JAGS**

```

win.data=list(log.Ni.hat = log.Ni.hat, x = x, log.index = log.index, cv.N = cv.N, n = n)

```

### **#initial parameters**

```

inits1 = list(N = (200000),
  tau.index = 10,
  q = 0.01,
  D = array(20, dim=c(11, 5)))

```

```

inits2 = list(N = c(150000),
  tau.index = 12,
  q = 0.001,
  D = array(15, dim=c(11, 5)))

```

```

inits=list(inits1,inits2)

```

### **# Parameters to estimate**

```

params <- c("theta", "N.ij", "q", "N.i", "N.j", "pi", "tau.index", "sigma.index", "D", "N", "D.sum",
"log.Ni.hat", "prop", "log.index", "N.MST", "N.MKT", "N.HLT", "N.UKT", "N.MST.j", "N.MKT.j", "N.HLT.j",
"N.UKT.j", "pi.Ni", "log.Nq")

```

### **# MCMC settings**

```

nc = 2 ; ni = 500000 ; nb = 50000 ; nt = 100

```

## Appendix B. Code for base BART model with shared run timing among years

### # Import Data

#### ##### Import weir data #####

#create an empty matrix for final location

log.Ni.hat = array(0, dim=c(12, 5))

# order is: (1)MST, (2)AMT, (3)SAL, (4)GEO, (5)HOH,(6)HLT, (7)KKG, (8)STN, (9)SWF,(10)TLS,  
(11)UKT, (12)TAK

#2003

log.Ni.hat.03 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x\_matrix/data/2003\_weirs\_BASE.csv")

#2004

log.Ni.hat.04 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x\_matrix/data/2004\_weirs\_BASE.csv")

#2005

log.Ni.hat.05 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x\_matrix/data/2005\_weirs\_BASE.csv")

#2006

log.Ni.hat.06 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x\_matrix/data/2006\_weirs\_BASE.csv")

#2007

log.Ni.hat.07 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x\_matrix/data/2007\_weirs\_BASE.csv")

#combining annual log.Ni.hat files into a single matrix

log.Ni.hat[,1] = log.Ni.hat.03[,1]

log.Ni.hat[,2] = log.Ni.hat.04[,1]

log.Ni.hat[,3] = log.Ni.hat.05[,1]

log.Ni.hat[,4] = log.Ni.hat.06[,1]

log.Ni.hat[,5] = log.Ni.hat.07[,1]

#### ##### Import final location data in x-matrix #####

#create an empty matrix for final location

x = array(0, dim=c(5, 12, 5))

# order is: (1)MST, (2)AMT, (3)SAL, (4)GEO, (5)HOH, (6)HLT, (7)KKG, (8)STN, (9)SWF, (10)TLS,  
(11)UKT, (12)TAK

# 2003

x\_03 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects SSART/shared run timing - staton methods - newest x\_matrix/data/2003\_x\_mat\_12\_grps\_5\_strata.csv")

```

# 2004
x_04 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects
SSART/shared run timing - staton methods - newest x_matrix/data/2004_x_mat_12_grps_5_strata.csv")

# 2005
x_05 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects
SSART/shared run timing - staton methods - newest x_matrix/data/2005_x_mat_12_grps_5_strata.csv")

# 2006
x_06 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects
SSART/shared run timing - staton methods - newest x_matrix/data/2006_x_mat_12_grps_5_strata.csv")

# 2007
x_07 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random Effects
SSART/shared run timing - staton methods - newest x_matrix/data/2007_x_mat_12_grps_5_strata.csv")

for (r in 1:5){
  for (c in 1:12){
    x[r,c,1] = x_03[r,c]
    x[r,c,2] = x_04[r,c]
    x[r,c,3] = x_05[r,c]
    x[r,c,4] = x_06[r,c]
    x[r,c,5] = x_07[r,c]
  }
}

#### Import index of abundance data ####
#create an empty matrix for indices of abundance
log.index = array(0, dim=c(5,5))

#log value for index of abundance (CPUE) in each time strata in 2003
log.index_03 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random
Effects SSART/shared run timing - staton methods - newest
x_matrix/data/2003_log_index_5_strata.csv")

#log value for index of abundance (CPUE) in each time strata in 2004
log.index_04 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random
Effects SSART/shared run timing - staton methods - newest
x_matrix/data/2004_log_index_5_strata.csv")

#log value for index of abundance (CPUE) in each time strata in 2005
log.index_05 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random
Effects SSART/shared run timing - staton methods - newest
x_matrix/data/2005_log_index_5_strata.csv")

```

```
#log value for index of abundance (CPUE) in each time strata in 2006
log.index_06 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random
Effects SSART/shared run timing - staton methods - newest
x_matrix/data/2006_log_index_5_strata.csv")
```

```
#log value for index of abundance (CPUE) in each time strata in 2007
log.index_07 = read.csv("C:/Users/farmetm/Documents/AYK SSI Work/R code/SSART Kusko/Random
Effects SSART/shared run timing - staton methods - newest
x_matrix/data/2007_log_index_5_strata.csv")
```

```
for (i in 1:5){
  log.index[i,1] = log.index_03[i,1]
  log.index[i,2] = log.index_04[i,1]
  log.index[i,3] = log.index_05[i,1]
  log.index[i,4] = log.index_06[i,1]
  log.index[i,5] = log.index_07[i,1]
}
```

#### #### create CV data for weirs ####

```
#providing values of coefficient of variation for abundance at each weir
cv.N = c(0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10, 0.10)
```

#### #### create n-matrix - total final locations in each strata ####

```
#create an empty matrix for indices of abundance
n = array(0, dim=c(5, 5))
```

```
#total number of chinook successfully tracked during in each time strata in 2003
n[,1] = c(sum(x[,1]), sum(x[,2]), sum(x[,3]), sum(x[,4]), sum(x[,5]))
```

```
#total number of chinook successfully tracked during in each time strata in 2004
n[,2] = c(sum(x[,2]), sum(x[,3]), sum(x[,4]), sum(x[,5]))
```

```
#total number of chinook successfully tracked during each time strata in 2005
n[,3] = c(sum(x[,3]), sum(x[,4]), sum(x[,5]))
```

```
#total number of chinook successfully tracked during each time strata in 2006
n[,4] = c(sum(x[,4]), sum(x[,5]))
```

```
#total number of chinook successfully tracked during each time strata in 2007
n[,5] = c(sum(x[,5]))
```

#### #### Specify Bayesian Analysis of Abundance and Run Timing (BART) model ####

```
#Write model
mod = function(){
  tau.index ~ dgamma(100,1)
  sigma.index <- 1 / sqrt(tau.index)
```



```

#setting up some constraints for alpha
for (j in 1:5) {
  for (i in 1:11) {
    alpha[i,j] ~ dgamma(0.001,0.001)
  }
}

for(y in 1:5){

  N[y] ~ dnorm(0,1.0E-12) %_ % l(1,)
  q[y] ~ dbeta(1,1)

  for(i in 1:11) {
    for(j in 1:5) {
      D[i,j,y] ~ dgamma(alpha[i,j],0.1)
    }
  }

  D.sum[y] <- sum(D[1:11,1:5,y])

  for(i in 1:10) {
    N.i[i,y] <- sum(N.ij[i,1:5,y])
    log.Ni.hat[i,y] ~ dnorm(log.Ni[i,y],tau.logNi[i,y])
    log.Ni[i,y] <- log(N.i[i,y])
    tau.logNi[i,y] <- 1 / log(cv.N[i] * cv.N[i] + 1)

    for(j in 1:5) {
      pi[i,j,y] <- D[i,j,y] / D.sum[y]
      N.ij[i,j,y] <- pi[i,j,y] * N[y]
      theta[j,i,y] <- N.ij[i,j,y] / N.j[j,y]
      pi.Ni[i,j,y] <- N.ij[i,j,y] / N.i[i,y]
    }
  }

  for(i in 11:11) {
    N.i[i,y] <- sum(N.ij[i,1:5,y])
    log.Ni.hat[i,y] ~ dnorm(log.Ni[i,y],tau.logNi[i,y])
    log.Ni[i,y] <- log(N.i[i,y])
    tau.logNi[i,y] <- 1 / log(cv.N[i] * cv.N[i] + 1)
    prop[y] ~ dbeta(1,1)

    for(j in 1:5) {
      pi[i,j,y] <- (D[i,j,y]*prop[y]) / D.sum[y]
      N.ij[i,j,y] <- pi[i,j,y] * N[y]
      theta[j,i,y] <- N.ij[i,j,y] / N.j[j,y]
      pi.Ni[i,j,y] <- N.ij[i,j,y] / N.i[i,y]
    }
  }
}

```

```

for(i in 12:12) {
  N.i[i,y] <- sum(N.ij[i,1:5,y])
  log.Ni.hat[i,y] ~ dnorm(log.Ni[i,y],tau.logNi[i,y])
  log.Ni[i,y] <- log(N.i[i,y])
  tau.logNi[i,y] <- 1 / log(cv.N[i] * cv.N[i] + 1)

  for(j in 1:5) {
    pi[i,j,y] <- (D[i-1,j,y]*(1-prop[y])) / D.sum[y]
    N.ij[i,j,y] <- pi[i,j,y] * N[y]
    theta[j,i,y] <- N.ij[i,j,y] / N.j[j,y]
    pi.Ni[i,j,y] <- N.ij[i,j,y] / N.i[i,y]
  }
}

for(j in 1:5) {
  N.j[j,y] <- sum(N.ij[1:12,j,y])
  log.Nq[j,y] <- log(N.j[j,y] * q[y])
  log.index[j,y] ~ dnorm(log.Nq[j,y], tau.index)

  x[j,1:12,y] ~ dmulti(theta[j,1:12,y],n[j,y])
}

#derived quantities
for (y in 1:5){
  N.MST[y] <- sum(N.i[1, y])
  N.MKT[y] <- sum(N.i[2:4, y])
  N.HLT[y] <- sum(N.i[5:7, y])
  N.UKT[y] <- sum(N.i[8:12, y])

  for (j in 1:5){
    N.MST.j[j,y] <- sum(N.ij[1,j,y])
    N.MKT.j[j,y] <- sum(N.ij[2:4,j,y])
    N.HLT.j[j,y] <- sum(N.ij[5:7,j,y])
    N.UKT.j[j,y] <- sum(N.ij[8:12,j,y])

    p.MST.j[j,y] <- N.MST.j[j,y]/N.MST[y]
    p.MKT.j[j,y] <- N.MKT.j[j,y]/N.MKT[y]
    p.HLT.j[j,y] <- N.HLT.j[j,y]/N.HLT[y]
    p.UKT.j[j,y] <- N.MST.j[j,y]/N.UKT[y]
  }
}

model.file = "model.txt"
write.model(mod, model.file)

```

**#bundle data to send to JAGS**

```
win.data=list(log.Ni.hat = log.Ni.hat, x = x, log.index = log.index, cv.N = cv.N, n = n)
```

**#initial parameters**

```
inits1 = list(N = c(200000, 200000, 200000, 200000, 200000),
             tau.index = 10,
             alpha = array(0.01, dim=c(11,5)),
             q = c(0.01, 0.01, 0.01, 0.01, 0.01),
             D = array(20, dim=c(11, 5, 5)))
inits2 = list(N = c(150000, 150000, 150000, 150000, 150000),
             tau.index = 12,
             alpha = array(0.05, dim=c(11,5)),
             q = c(0.001, 0.001, 0.001, 0.001, 0.001),
             D = array(15, dim=c(11, 5, 5)))
```

```
inits=list(inits1,inits2)
```

**# Parameters to estimate**

```
params <- c("theta", "N.ij", "q", "N.i", "N.j", "pi", "tau.index", "sigma.index", "D", "alpha", "N", "D.sum",
"log.Ni.hat", "prop", "log.index", "N.MST", "N.MKT", "N.HLT", "N.UKT", "N.MST.j", "N.MKT.j", "N.HLT.j",
"N.UKT.j", "p.MST.j", "p.MKT.j", "p.HLT.j", "p.UKT.j", "pi.Ni", "log.Nq")
```

**# MCMC settings**

```
nc = 2 ; ni = 1000000 ; nb = 50000 ; nt = 200
```

## CHAPTER 5

### ASSOCIATIONS BETWEEN ALASKAN CHINOOK SALMON POPULATION CHARACTERISTICS AND DRAINAGE-WIDE HABITAT VARIABLES

#### ABSTRACT

Management strategies for Alaska Chinook salmon rely on biological reference points (e.g.,  $S_{MSY}$ ) to set harvest targets which are traditionally derived using spawner-recruit analysis. However, not all stocks in Alaska are sampled intensively enough to allow for derivation of reference points through spawner-recruit analyses, yet their management must still be based on stock-specific reference points. Habitat-based methods have been developed to predict  $S_{MSY}$  for stocks without adequate spawner-recruit information based on a subset of stocks that do have this information. These models form a linear predictive relationship between  $S_{MSY}$  and certain habitat characteristics (e.g., drainage area) that can be used for more data-limited stocks. Here I build on this work by developing a hierarchical modeling approach for Alaskan Chinook salmon stocks that incorporates the uncertainty in stock-specific  $S_{MSY}$  and applying a suite of model selection techniques to select habitat variables that have predictive credibility. Findings showed that hierarchical inclusion of uncertainty in  $S_{MSY}$  did not remove the relationship between drainage area and  $S_{MSY}$ , showing that this relationship may still be used even when including this additional source of variation. Three of the four variable selection techniques consistently agreed on the best variables. The deviance information criterion performed poorly at distinguishing models and potential explanations are discussed. Drainage area as the sole predictor was unanimously chosen by the best model under the three remaining model selection approaches.

## INTRODUCTION

Management strategies for Chinook salmon (*Oncorhynchus tshawytscha*) stocks rely on estimates of biological reference points (e.g.  $S_{MSY}$ ) to set harvest targets. Ideally, these reference points are derived using spawner-recruit analysis on a time series of spawner and return counts which provides information regarding productivity, carrying capacity, and strength of compensation of the stock. However, many Chinook salmon stocks are not sampled intensively enough for reliable estimation of biological reference points via spawner-recruit analysis. In these more data-limited situations, habitat-based assessment approaches have been used. These methods use habitat variables that are related to salmon productivity and capacity to build models using estimates from stocks with adequate spawner-recruit data. The coefficient estimates from this model can then be used for prediction of reference points for stocks that have inadequate spawner-recruit data.

One habitat variable that has shown promise in these models is the area of the drainage for each stock. Parken et al. (2006) developed such a model using stocks from Oregon, Washington, British Columbia, and Alaska. However, a predictive model has not yet been developed for exclusively Alaskan stocks (i.e., those that spawn in rivers that originate in Alaska), which will be the focus of this chapter. Although several of the stocks used in the present analysis were also used by Parken et al (2006), the data for these stocks have since been updated with more current information and the individual spawner-recruit analyses for each stock used a more rigorous Bayesian state-space approach (Catalano 2012).

The methodology for the modeling approach presented here differs in two primary ways from that of Parken et al. (2006). First, Parken et al. (2006) treated the reference points (i.e., the response variable) as known quantities, when they are really derived from parameters estimated from a spawner-recruit analysis. Spawner-recruit relationships are inherently noisy and as a result, their parameters and the derived reference points (equations 2.09-2.011, Chapter 2, this Thesis) contain substantial uncertainty. Treating the reference points as known quantities ignores this uncertainty and may lead to conclusions that would not be valid were this uncertainty included (Brooks and Debora 2015). For this reason, a hierarchical modeling approach is implemented here to allow for incorporation of the uncertainty in the reference points that the model is fit to. The model is cast in the Bayesian mode of inference which allows for an intuitive framework to include this additional source of uncertainty. Second, model uncertainty will be introduced into the habitat-based modeling approach by examining additional habitat variables. These additional variables are hypothesized to influence salmon productivity and capacity and are available via remote sensing techniques.

The Parken et al. (2006) model is not the only predictive habitat-based model for Chinook salmon. Liermann et al. (2010) developed an integrated hierarchical model that included an intrinsic link between watershed area and spawner-recruit parameters in the model fitting. Thus, the Liermann et al. (2010) model used not only the stock-specific spawner-recruit data but also habitat information to estimate the spawner-recruit relationships for stocks used in their model. This sharing of information between stocks was intended to improve predictive performance by reducing uncertainty in the predictive relationship. Although the analyses presented in this chapter do not share information between stocks when fitting the spawner-recruit relationships, as in the Liermann et al. (2010) analysis, the regression coefficients of the habitat model are estimated from all stocks.

Model selection is an area of active research in ecological modeling as a means to variable selection (Hooten and Hobbs 2015; Anderson et al. 2008). Of particular interest are methods for model selection for analyses in the Bayesian mode of inference as no one unified approach performs well for all models (Hooten and Hobbs 2015). Much of this disagreement revolves around the requirement of the prior in Bayesian analyses because objective methods for choosing appropriate priors do not exist. This deficiency remains one of the most prevalent criticisms of the Bayesian framework, as two investigators can obtain different results based on prior beliefs. The prior makes information theoretic approaches to

Bayesian model selection particularly troublesome, as there are no truly free parameters in the Bayesian context due to the information content in the prior, however uninformative it may be (Hooten and Hobbs 2015). Thus, four different approaches are applied here to selecting habitat variables based on their predictive abilities while accounting for uncertainty in the response variable: variable indicator selection, the deviance information criterion (DIC), the Akaike information criterion (AIC) using randomized posterior samples, and leave-one-out cross validation. So in addition to investigating the relationship between various habitat variables with  $S_{MSY}$ , this chapter also presents a comparison of various approaches to model selection. The objectives of the analyses presented in this chapter are to (1) develop a habitat-based predictive model for data-poor Alaskan Chinook salmon stocks that explicitly incorporates uncertainty in the estimates that it is fit to by (2) evaluating the agreement in the best predictor variables between a variety of model selection techniques.

## METHODS

### *Study Area and Species*

This study uses information from Chinook salmon stocks across the state of Alaska (Figure 4.01). Chinook salmon are an important species for commercial, subsistence, and recreational fisheries across the state. In the past decade, many Chinook salmon stocks across the state have seen declines in productivity, making adequate treatment of uncertainty in management quantities of paramount importance. Catalano (2012) conducted a meta-analysis of 15 Chinook salmon stocks across the state of Alaska by developing Bayesian state-space spawner-recruit models for each stock that gave marginal posterior distributions for stock-specific  $S_{MSY}$ . The summaries of these posterior distributions are used to fit the regression models in this chapter.

The stocks used in this analysis cover a wide range in terms of maximum productivity (2.08 – 8.95 recruits per spawner), stock abundance (8,800 – 244,000 average recruits), and drainage area (562 – 286,000 km<sup>2</sup>). Geographically, these stocks are spread all throughout the state (Figure 4.01) including southeastern Alaska (Stikine, Taku, and Alsek), Kodiak Island (Karluk), western Alaska (Goodnews, Kuskokwim, and Unalakleet), and the interior (Chena+Salcha; two rivers, assessed as one stock).

### *Selected Habitat Variables*

Habitat variables were selected based on a hypothesized effect on salmon productivity or stream carrying capacity. Both productivity and capacity impact the  $S_{MSY}$  reference point, so any habitat variable that may impact these two quantities should be included in the analysis. Three habitat variables were selected that each have biological justification for inclusion in the model, as outlined in the following sections. The value of these variables for each stock were accessed using the Riverscape Analysis Project (RAP) online database (Whited et al. 2012; <http://rap.ntsg.umn.edu>). Four of the 15 stocks were outside of the RAP study area, and thus were excluded from this analysis. All habitat variables were standardized (scaled and centered) such that no one data point fell too far from zero, as recommended by Kéry (2010).

The area (km<sup>2</sup>) of the drainage (watershed) of a particular river was hypothesized to influence the value of  $S_{MSY}$  because a larger area results in more spawning and rearing grounds for salmon, which directly impacts the capacity of a system. Drainage area has been shown to have predictive power for  $S_{MSY}$  in both stream-type (Alaskan stocks) and ocean-type (more southern stocks) Chinook salmon (Parken et al. 2006). This variable was log-transformed to linearize the relationship with  $\log(S_{MSY})$ , following Parken et al. (2006).

Tributaries are the areas within the drainage in which adult salmon typically spawn and juvenile salmon spend much of their lives before migrating to sea. Therefore, it stands to reason that drainages that have more tributaries should produce more salmon. This variable was expressed as per km<sup>2</sup> to remove collinearity with area; larger drainages have more tributaries. Tributary node density was not

log-transformed as the range was so small that a log-transformation had no effect in linearizing the relationship with  $\log(S_{\text{MSY}})$ .

The shifting habitat mosaic hypothesis predicts that productivity is maximized at high levels of stream complexity (Standford et al. 2005). Channel nodes are an indicator of this complexity as they represent the amount of braiding in the stream channel and floodplains and so was selected for analysis. This variable was standardized by area ( $\text{km}^2$ ) for the same reason as tributary nodes. Channel node density was not log-transformed as the relationship with  $\log(S_{\text{MSY}})$  appeared linear.

### Model Structure

Consider the normal linear model where the response variable,  $y_i$ , is the natural logarithm of  $S_{\text{MSY}}$  for stock  $i$  and  $\theta_{0:n}$  are coefficients that describe the linear relationship between habitat variable  $x_{1:n}$  and  $y_i$  for stock  $i$ . Residual variation in this relationship (variation in  $y$  not explained by the habitat variables  $x_{1:n}$ ) is assumed to be independently and identically normally distributed with a mean of zero and some standard deviation  $\sigma$  that is freely estimated.

To incorporate uncertainty in  $y_i$  (due to estimation and process error in the spawner-recruit relationship and measurement error in the data collection) an additional source of variation was required, which made the model hierarchical in nature (i.e., having multiple sources of statistical uncertainty). First, a model prediction was made based on the underlying linear process that was assumed to give rise to the mean  $\log(S_{\text{MSY}})$  for each stock:

$$\mu_i = \theta_0 + \theta_1 x_{1,i} + \dots + \theta_n x_{n,i} \quad (3.01)$$

Equation 3.01 constitutes the linear predictor portion of the model, but it is obvious that there is noise in the relationship due to factors other than the selected habitat variables. This additional source of variation is commonly referred to as process error. Thus, a stock-specific latent state  $\log(S_{\text{MSY}})$  was sampled from a normal distribution that represents the true  $\log(S_{\text{MSY}})$  after accounting for this process variation around the deterministic relationship:

$$\psi_i \sim N(\mu_i, \tau_\psi) \quad (3.04)$$

where  $\psi_i$  is the latent state for stock  $i$  as produced by the habitat relationship and unexplained noise and had some estimated precision (inverse of the variance) parameter  $\tau_\psi$  that was common across all stocks. The response variable  $y_i$  was related to the latent  $\psi_i$  via another stochastic node:

$$y_i \sim N(\psi_i, \tau_i) \quad (3.05)$$

where  $y_i$  is the posterior mean  $\log(S_{\text{MSY}})$  and  $\tau_i$  is the inverse of the posterior variance for stock  $i$ , as presented by Catalano (2012).  $\tau_i$  can be interpreted as an assumed measurement error on the true state, and its inclusion allowed for the incorporation of uncertainty in the estimated  $S_{\text{MSY}}$  from the analysis presented by Catalano (2012). It is this additional source of uncertainty in the response variable (measurement as well as process variation) that is not included in a typical multiple regression.

### Habitat Variable Selection

Four different methods were used for selecting which variables should be included in the final predictive habitat-based model. A brief overview of each method is outlined below.

#### Indicator Variable Selection (Parameter Inclusion Probability)

This method used binary indicator parameters that included or excluded habitat variables in each Markov Chain Monte Carlo (MCMC) iteration to select variables that were justified in the model. The approach sampled an indicator parameter  $\omega$  from an uninformative Bernoulli distribution for each slope coefficient in the process model:

$$\omega_{1:n} \sim \text{Bernoulli}(0.5) \quad (3.06)$$

These indicators were then included in the linear predictor in equation 3.03

$$\mu_i = \theta_0 + \omega_1 \theta_1 x_{1,i} + \dots + \omega_n \theta_n x_{n,i} \quad (3.07)$$

For each MCMC iteration, each coefficient was included ( $\omega = 1$ ) or excluded ( $\omega = 0$ ). The posterior mean of  $\omega_n$  can be interpreted as the probability that  $\beta_n x_n$  should be included in the model. This is a technique for incorporating model uncertainty into the model fitting. The combination of coefficients being included or excluded in each MCMC iteration results in a total of eight different models. Thus, a model probability can be calculated by taking the proportion of MCMC iterations the respective coefficients were included in the same MCMC iteration.

If a variable is excluded from the linear predictor in an iteration ( $\omega = 0$ ), its coefficient is still sampled on that iteration. As noted by Hooten and Hobbs (2015), independent diffuse priors in this setting can cause computational problems in the MCMC sampling. For example, if the coefficient is excluded in one iteration, and the sampler draws a value of the corresponding coefficient with a highly diffuse prior, it is possible for the MCMC to get “lost”. Essentially, the sampled value can be so far from the majority of the posterior density that the MCMC algorithm cannot return to that density by sampling a more likely value and the indicator parameter is never included again. This issue can be addressed in at least two ways: (1) through the use of informative priors that prevent the MCMC sampler from getting “lost” by making extreme values so unlikely that they are never sampled or (2) the use of reversible jump MCMC (RJMCMC), where if  $\omega = 0$  for a given iteration, the corresponding coefficient is not sampled that iteration. Since a general RJMCMC algorithm has not yet been developed for use in JAGS (or OpenBUGS), informative priors were chosen to remedy this issue (Table 1).

#### *Deviance Information Criterion (DIC)*

Information theoretic approaches have gained substantial support in model selection techniques and ecological analysis (Anderson et al. 2008). However, as mentioned previously, the information contained in the prior makes information theoretic approaches to Bayesian model selection troublesome (Spiegelhalter et al. 2014). However, one proposed solution is the deviance information criterion (DIC; Spiegelhalter et al. 2002). DIC is the Bayesian analog to the Akaike information criterion (AIC) for models fit in the frequentist mode of inference, in that it uses a measure of statistical fit and penalizes it by the number of parameters in the model (Plummer 2008). DIC is calculated as:

$$DIC = \widehat{D} + p_D \quad (3.08)$$

where  $\widehat{D}$  is the posterior expectation for model deviance calculated at the posterior means of all unknown quantities (Hooten and Hobbs 2015).  $p_D$  is the effective number of parameters, calculated as  $\bar{D}$  (mean posterior deviance) minus  $\widehat{D}$ .

#### *Randomization with AIC Model Selection*

Another way to incorporate posterior uncertainty in  $S_{MSY}$  into the model fitting and selection is to randomize and sample from these posterior samples, fit each model to each randomized sample, and conduct AIC model selection using each sample. The distribution of the model selections across all randomized samples can then serve as the basis for inference on justified variables. The posterior distributions of  $S_{MSY}$  for each stock contain approximately 10,000 samples each (less than 10,000 after removing the very rare negative  $S_{MSY}$ ). These samples were randomized and all valid samples were used for this exercise. Models for this randomization component were fit using ordinary least-square methods with the normal linear model instead of Bayesian methods. Inference regarding which variables were most important was made using three criteria: (1) frequency of placement of each variable in the best model, (2) the distribution of the position of each model in the model selection across all samples, and (3) the distribution of model weight for each model across all samples. Although the term AIC is used to describe this method of model selection, the information theoretic calculation that was used was the version corrected for small sample sizes ( $AIC_C$ ; Anderson et al. 2008).

#### *Leave-One-Out Cross Validation*



The fourth and final model selection technique used in this analysis is referred to as “*k*-fold cross-validation”. The premise of this method is to test the capability of a model to predict datasets that are not included in the model fitting (i.e., the model is unaware of them). *K* datasets were sequentially left out of the model fitting, then compared to their corresponding model prediction to calculate the predictive loss of that model (Hooten and Hobbs 2015). In the case where *n* is small, like in the scenario presented in this chapter, one data point can be left out at a time (hence, leave-one-out cross-validation). Hooten and Hobbs (2015) present a cross-validation score calculation for a single model:

$$\sum_{k=1}^K \log \frac{\sum_{t=1}^T P(y_k | y_{-k}, \vartheta_t)}{T} \quad (3.09)$$

where  $y_k$  is the measured value of the data point left out in the  $k^{\text{th}}$  run of the model out of *K* total runs,  $y_{-k}$  is the model prediction for that data point given by the parameters  $\vartheta$  on the  $t^{\text{th}}$  MCMC iteration fit with  $y_k$  left out and *T* is the number of saved MCMC samples. Since the models assume a normal error structure, it is appropriate to use a sum-of-squared residual in place of the likelihood calculation in the numerator of equation 3.09. The altered cross-validation score used in this analysis is:

$$\sum_{k=1}^K \frac{\sum_{t=1}^T (y_k - y_{-k})^2}{T} \quad (3.10)$$

This process was then repeated for all subsets of the full model and the scores were compared across models, where the lowest score indicates the least predictive loss (i.e., accurately predicts data points not included in the model) and the best model for prediction.

### Computation

All analyses were conducted using the R statistical environment (R Core Development Team; 2014). The indicator variable selection and cross validation methods were conducted using JAGS (Just Another Gibbs Sampler; Plummer 2013) using the R package “rjags” (Plummer 2014). JAGS is an implementation of the BUGS (Bayesian Inference Using Gibbs Sampling) language, which is a flexible and intuitive platform for specifying and fitting Bayesian hierarchical models. The DIC approach was conducted using the “dic.samples” function in the “rjags” package (Plummer 2014). Marginal posterior distributions were summarized as the mean and 95% credible intervals (i.e., 2.5 and 97.5 percentiles). Randomization procedures were conducted using original R code to randomly select individual posterior samples of  $S_{\text{MSY}}$  for each stock as the response variable and the R package “MuMIn” (i.e., multi-model inference; Barton 2015) to conduct AIC model selection for all eight models for each randomized sample.

## RESULTS

Convergence of all models fit with MCMC methods was confirmed by a Gelman-Rubin statistic for all estimated parameters of <1.1 (Gelman et al. 2004).

### Indicator Variable Selection

In terms of which variables should be included in the model, drainage area received the most posterior support for inclusion in the model, followed by channel node density, and tributary node density received the least support (Table 2). Barberi and Berger (2004) suggested that variables with indicator parameter values of greater than 0.5 contain justifiable predictive information. Following this threshold, drainage area was the only habitat variable that can be reliably used for prediction.

There was little supporting evidence for the inclusion of more than one predictor variable (Table 4.03). The model that carried the most probability was the drainage area-only model, followed by the channel node density-only model. The intercept-only model carried more probability than any of the other five models, which suggested that the mean  $S_{\text{MSY}}$  across all stocks was a better predictor of stock-specific  $S_{\text{MSY}}$  than the combinations of the habitat variables in these five lower models. Tributary node

density appeared in the worst four models, which indicated that it was a very poor predictor of  $S_{MSY}$  for the stocks included in this analysis. These findings were further supported by the estimated effect size for each parameter and whether its credible interval included zero (Table 4). These effects were calculated from the posterior distribution of each parameter when the parameter was included to prevent the coefficient samples that did not influence the likelihood from affecting the posterior mean. Drainage area had the largest effect and it was the only parameter with a 95% credible interval that did not overlap zero.

Since informative priors were used to aid the MCMC sampler, the influence of prior information was investigated graphically (Figure 4.02). It is clear that the seemingly informative prior (Table 1) was indeed flat over the range of the posterior density for each coefficient. All posterior samples where  $\omega = 0$  were removed for this plot to assess the influence of the prior on the posterior since these were the only samples where the data influenced the posterior draw.

Coefficients were combined using model averaging according to how often they were included in the model fitting and were used to make a predictive diagnostic plot (Figure 4.03). This is an “observed versus predicted” plot where points are pairs of model predictions (x-axis) and observed data points (y-axis). The dotted line is the 1:1 line where a model prediction is the exact same as the data. For stocks with low  $S_{MSY}$ , the model performed well at predicting their value, but the stock with observed  $S_{MSY}$  of approximately 80,000 was not predicted well by the model-averaged coefficients. This abnormal stock is the Kuskokwim River in western Alaska, and it seems that it requires more escaping fish to attain maximum sustainable yield than would be predicted by the habitat variables included in this model.

### ***Deviance Information Criterion***

The DIC analysis showed that each of the models had  $p_D$  within one parameter of one another (Table 3). In general, models with the same number of predictor variables had the same number of effective parameters (e.g., the area-only and the channel node density-only model had  $p_D$  of 9.62 and 9.63, respectively; Table 3). However, this was not always the case as the intercept-only model that contained no coefficients for habitat predictor variables and the full model that contained coefficients for all three habitat variables resulted in essentially the same  $p_D$  (10.02 versus 10.03, respectively; Table 3). This showed that in some cases the sampler was able to adequately estimate the number of free parameters, but that in other cases the number of free parameters changed in an unpredictable manner.

### ***Randomization with AIC Model Selection***

Drainage area and channel node density were supported by this randomization analysis, whereas tributary node density showed very little support, as it was in all four of the worst models (Tables 2, 3; Figure 4). Drainage area and channel node density were present in the best model in 71% and 27% of the randomized samples, respectively (Table 2). Tributary node density was present in the best model in only 2% of the samples.

### ***Leave-One-Out Cross Validation***

In terms of predictive ability for data points left out of the model fitting, the drainage area-only model exhibited the lowest predictive loss and the additive area and channel node density had the next lowest loss (Table 3). The intercept-only model placed fifth in predictive loss, which indicated that there are combinations of variables that perform more poorly at out-of-sample prediction than simply the mean of the included stocks. This intercept-only model had approximately 50% more loss than the area-only model. The worst two models, the full model and the additive channel node density and tributary node density model, each had 78% more predictive loss than the best model.

## DISCUSSION

Overall, the different model selection techniques agreed very well with regard to which variables have predictive credibility. All methods except DIC selected drainage area as the best predictor of  $S_{MSY}$ , followed by channel node density. Tributary node density gained very little support under three out of the four methods (DIC disagreed), perhaps due to the lack of resolution in this variable (all stocks had approximately 0.1 tributaries/km<sup>2</sup>). Additionally, the same three out of the four methods favored selecting only one predictor for the best model, and this single predictor was drainage area. This finding lends more credibility to the Parken et al. (2006) model that considered only drainage area. DIC showed little to no separation between models, making inference based its findings very difficult.

The model weight column of Table 3 (randomization method) is analogous in interpretation to the model probabilities shown in the first column (variable indicator method) because the AIC model weight is defined as the probability that the model is the best, among those selected for analysis. This definition is equivalent to the model probability presented in Table 3. Upon comparison of these two findings, it is clear that the two methods give very similar results. All models carried approximately the same weight under both approaches, with the primary difference being that the intercept-only model placed third best under the Bayesian indicator selection and fourth best in the randomization.

Based on this consistent agreement with the randomization approach, it seems that the Bayesian indicator variable approach is an adequate technique for model selection for this hierarchical model. However, the indicator variable technique only returned these results when the specific prior in Table 1 was used for the coefficients. When the prior was made more diffuse in preliminary runs of the analysis, the indicator variables tended to select the intercept-only model as the best, and with high certainty (approximately 0.8 probability). This was likely due to the MCMC sampler “getting lost” as previously mentioned. So while the Bayesian model selection technique did return very similar inference to the more frequentist method, its findings were contingent on the prior specification, which warrants some caution. As there is no objective way to decide how informative of a prior one should use to correct the “getting lost” problem, it is possible that two investigators would come to very different conclusions using these same data. This discrepancy is potentially due to two linked mechanisms: (1) the small sample size and (2) the hierarchical incorporation of uncertainty in the response variable. Including the uncertainty in the response variable required more parameters and placed a high demand of the small number of stocks used. There were 11 stocks used to fit the variable indicator model, and each had a mean  $S_{MSY}$  and an inverse variance that were provided to the model as data, which lead to a total of 22 pieces of information used by the model (ignoring priors). If the effective number of parameters returned by the DIC analyses would be trusted, then the model had approximately 10 free parameters, which means that there were approximately two pieces of information for every free parameter. This information-to-parameter ratio puts this model on the fringes of over-parameterization. If it were possible to include more data points, it is possible that a more diffuse prior specification could be used on the coefficients.

The hierarchical inclusion of posterior uncertainty in  $\log(S_{MSY})$  from the Catalano (2012) analysis was effective at increasing the uncertainty in the area- $\log(S_{MSY})$  relationship. When a Bayesian hierarchical model with unstandardized area only and uninformative priors on the coefficients (mean = 0, precision = 1E-8) was compared to a frequentist analysis that treated the  $\log(S_{MSY})$  as known quantities, the estimates showed more uncertainty under the Bayesian analysis. The estimate of the area effect was 0.35 (0.08-0.61) under the Bayesian analysis and 0.35 (0.10-0.60) under the frequentist analysis. While these estimates and intervals are very similar, the residual standard error was greater under the Bayesian analysis (0.79) than under the frequentist analysis (0.72). This greater uncertainty in the area- $\log(S_{MSY})$  relationship is due to the inclusion of uncertainty in  $\log(S_{MSY})$ , however one could argue that the difference is not large enough to discredit treating  $\log(S_{MSY})$  as known quantities under a frequentist analysis.

DIC proved ineffective at distinguishing between models, and thus inference based on its results should be done so carefully. The full model included three more coefficients than the intercept-only model, however the DIC model selection indicated that there was no difference in the effective number of parameters between these models. This was an unexpected finding, but one can speculate as to why this was the case. The model used a linear predictor to make the mean prediction for each stock based on habitat variables ( $\mu_i$ ), then sampled a latent state variable ( $\psi_i$ ) from a distribution with the linear prediction ( $\mu_i$ ) as the mean and a common process error ( $\tau_\psi$ ). These latent states were then fit to the mean  $\log(S_{\text{MSY}})$  and their variances from Catalano (2012) in the likelihood. When there were predictor variables included in the model, there was information about the values of the latent states contained in the linear predictor. Conversely, when only the intercept was used as the linear predictor, the only information about the latent states was the mean  $\log(S_{\text{MSY}})$  across all stocks. In the latter case, the latent states were more “free” as they were less informed by habitat data and thus could have resulted in a similar  $p_D$  as the full model. While this explanation is purely speculation, one could see how this might play out within the model fitting. Regardless of the cause the lack of model separation in the DIC analysis, its nature renders inference based on this method regarding the best predictor variables highly uncertain and essentially impossible.

Furthermore, Plummer (2008) noted that DIC is only useful for models in which the ratio of data points to parameters is large. If this is not the case, then more complex models will be underpenalized (Plummer 2008). Although Plummer (2008) never defined what “large” means for this ratio, certainly the ratio in the present analysis was small. There was a total of 11 stocks used in this analysis (Figure 4.01) and the model used latent variables for each stock, plus coefficients for the habitat predictors, and a common precision parameter for the latent variable sampling distributions. This is another potential cause for the DIC model selection results lacking differentiation between models.

It is worth noting that while the results of these model selection exercises gave credence to Parken et al.’s (2006) choice of drainage area as the sole predictor of  $S_{\text{MSY}}$ , the coefficients from the present analysis are not directly comparable to this previous study. Parken et al. (2006) censored drainage area based on accessible habitat: that area of the drainage that feeds the stream below natural and man-made barriers. Drainage area used in the present study was not censored using this criterion. Thus, one unit of area in the Parken et al. (2006) approach can support more spawners than one unit of area in the present study. This statement agrees with the coefficient estimates from both studies: Parken et al. (2006) reported an effect size of  $0.69/\text{km}^2$  whereas a model using only unstandardized drainage area under the present study estimated an effect size of  $0.35/\text{km}^2$ . It could be argued that not censoring area based on accessible habitat makes the model more straight-forward to implement for managers, but it also may introduce an additional source of variation into the model.

## REFERENCES

- Anderson, D.R., K.P. Burnham, W.L. Thompson. 2008. Null hypothesis testing: problems, prevalence, and an alternative. *Journal of Wildlife Management*. 64(4): 912-923.
- Barberi, M.M, and J.O. Berger. 2004. Optimal predictive model selection. *The Annals of Statistics*. 32(3): 870-897.
- Barton, K. 2015. MuMIn: Multi-Model Inference. R package version 1.13.4. <http://CRAN.R-project.org/package=MuMIn>.
- Catalano, M. J. 2012. Comparative analysis of patterns in productivity for AYK and selected non-AYK Chinook salmon stocks. Workshop Background Paper (unpublished working paper) *In* AYK SSI Chinook salmon Synthesis Workshop, May 2-3, 2012. Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative, Anchorage, AK.
- Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2004. *Bayesian Data Analysis*. Second ed. CRC/Chapman & Hall, Boca Raton, Florida.
- Kery, M. 2010. *Introduction to WinBUGS for ecologists*. Academic Press, Burlington, Massachusetts.
- R Development Core Team. 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Version 3.1.0. Vienna, Austria. <http://www.R-project.org>.
- Plummer, M. 2014. rjags: Bayesian graphical models using MCMC. R package version 3-14. <http://CRAN.R-project.org/package=rjags>.
- Plummer, M. 2013. JAGS Version 3.4.0 user manual.
- Plummer, M. 2008. Penalized loss functions for Bayesian model comparison. *Biostatistics*. 9(3): 523-539.
- Parken, C. K., R. E. McNicol, and J. R. Irvine. 2006. Habitat-based methods to estimate escapement goals for data limited Chinook salmon stocks in British Columbia, 2004. Department of Fisheries and Oceans, Canadian Scientific Advisory Secretariat Research Document 2006/083.
- Spiegelhalter, D.J., N.G. Best, B.P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*. 64(4): 583-689.
- Spiegelhalter, D.J., N.G. Best, B.P. Carlin, and A. van der Linde. 2014. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B* 76(3): 485-493.
- Stanford, J.A., M.S. Lorang, and F.R. Hauer. 2005. The shifting habitat mosaic of river ecosystems. *Proceedings of the International Society of Limnology*. 29(1): 123-136.
- Thomas, A., B. O'Hara, U. Ligges, and S. Sturtz. 2006. Making BUGS open. *R News*. 6(1): 12-17.
- Whited, D.C., J.S. Kimball, J.A. Lucotch, N.K. Maumenee, H. Wu, S.D. Chilcote, and J.A. Stanford. 2012. A riverscape analysis tool developed to assist wild salmon conservation across the north Pacific Rim. *Fisheries*. 37(7): 305-314.

Table 1. Prior distributions for all unknown parameters in the Bayesian hierarchical regression models.  $\omega_{1:3}$  were only included in the variable indicator selection method. Note that the dispersion parameter of the normal distribution is the precision (as opposed to variance) which is the inverse of variance.

Parameter	Prior Distribution
$\theta_{0:3}$	$N(0, 0.1)$
$\tau_\psi$	$\text{Gamma}(0.01, 0.01)$
$\omega_{1:3}$	$\text{Bernoulli}(0.5)$

Table 2. Parameter inclusion probability from the indicator variable selection method (labeled Bayes) and the proportion of randomized samples that resulted in each variable being selected in the best model (labeled AIC). A = log(area), C = channel node density, T = tributary node density.

Variable	(Bayes) Inclusion Parameter ( $\omega$ )	(AIC) Proportion In best model
A	0.56	0.71
C	0.36	0.27
T	0.15	0.02

Table 3. Results from the four model selection approaches. Models are ordered by the strength of support under the variable indicator selection method (first column). Bolded values are the selected best model under each approach. The values in the parentheses in the Randomize w/AIC column are the standard deviation of the placement and weight across all randomized posterior samples.

Model	Variable Ind.	DIC		Randomize w/AIC		Cross-Validation	
	Model Probability	pD	DIC	Model Position	Model Weight	Score	% Change from Best Model
A	0.37	9.62	11.77	1.48 (0.78)	0.37 (0.14)	9.23	0%
C	0.19	9.63	11.52	2.33 (1.02)	0.23 (0.12)	12.17	32%
O	0.16	10.01	12.61	4.03 (1.42)	0.11 (0.10)	13.75	49%
AC	0.13	9.59	11.64	3.27 (1.37)	0.16 (0.10)	12.01	30%
T	0.06	10.06	12.25	4.80 (1.33)	0.07 (0.07)	14.40	56%
AT	0.04	9.88	12.19	5.90 (1.11)	0.03 (0.01)	13.13	42%
CT	0.03	9.62	11.62	6.29 (0.96)	0.03 (0.02)	16.40	78%
ACT	0.02	10.03	12.31	7.91 (0.45)	0.01 (0.01)	16.39	78%

Table 4 Coefficient estimates from the variable indicator selection method after samples where  $\omega = 0$ . Values in parentheses indicate the 95% credible interval of these thinned posteriors.

Variable	Mean (95% CI)
A	0.69 (0.07, 1.27)
C	0.56 (-0.07, 1.19)
T	0.28 (-0.62, 1.11)



Figure 1. Distribution of stocks included in this analysis.



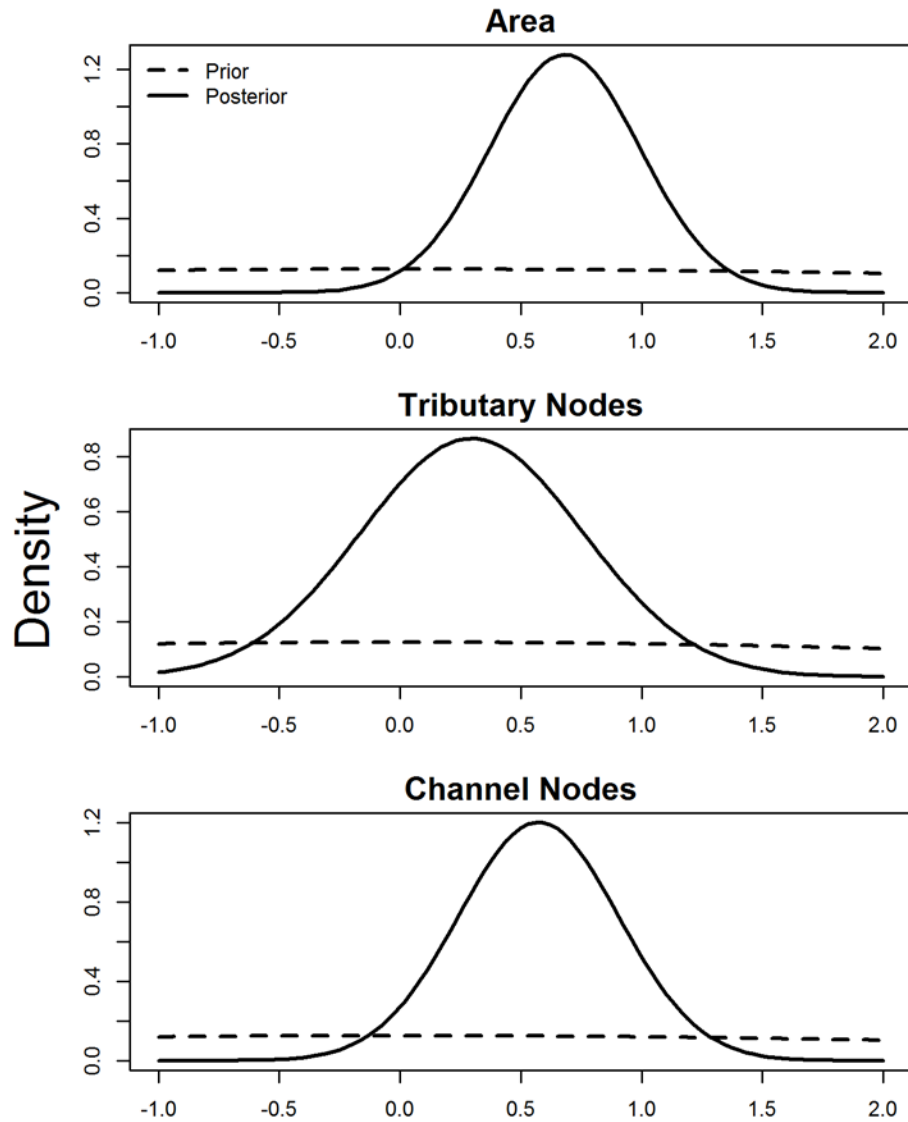


Figure 2. Posterior distribution for each slope coefficient after iterations where the corresponding indicator variable was turned off (solid line) were excluded and the prior shown in Table 1 (dashed line).

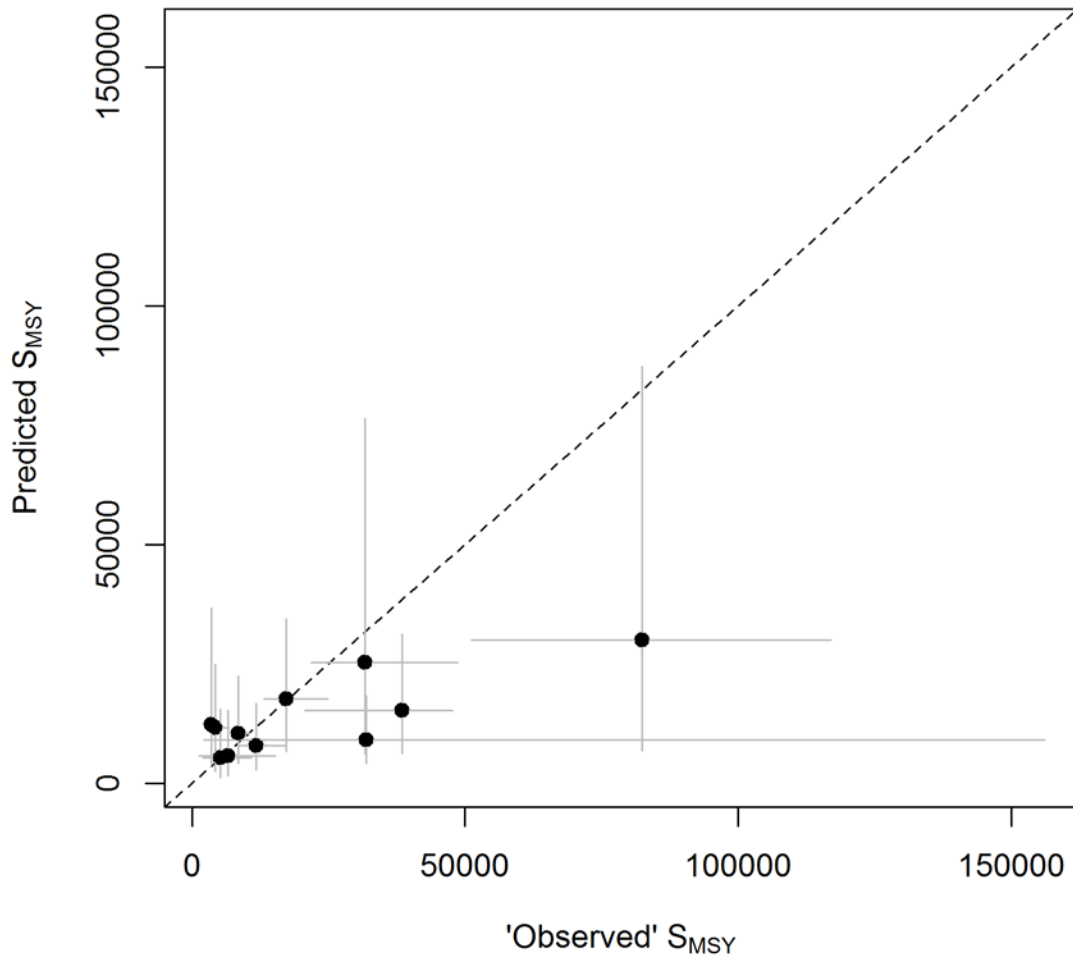


Figure 3. Observed vs. predicted plot for each of the 11 stocks included in the variable indicator selection method. Observed data points are the posterior means presented in Catalano (2012) and predictions were made using model-averaged coefficients from the variable indicator selection. Error bars in the vertical direction are 95% Bayesian credible intervals from the Catalano (2012) analysis and credible intervals for the model predicted  $S_{MSY}$  from the variable indicator model.

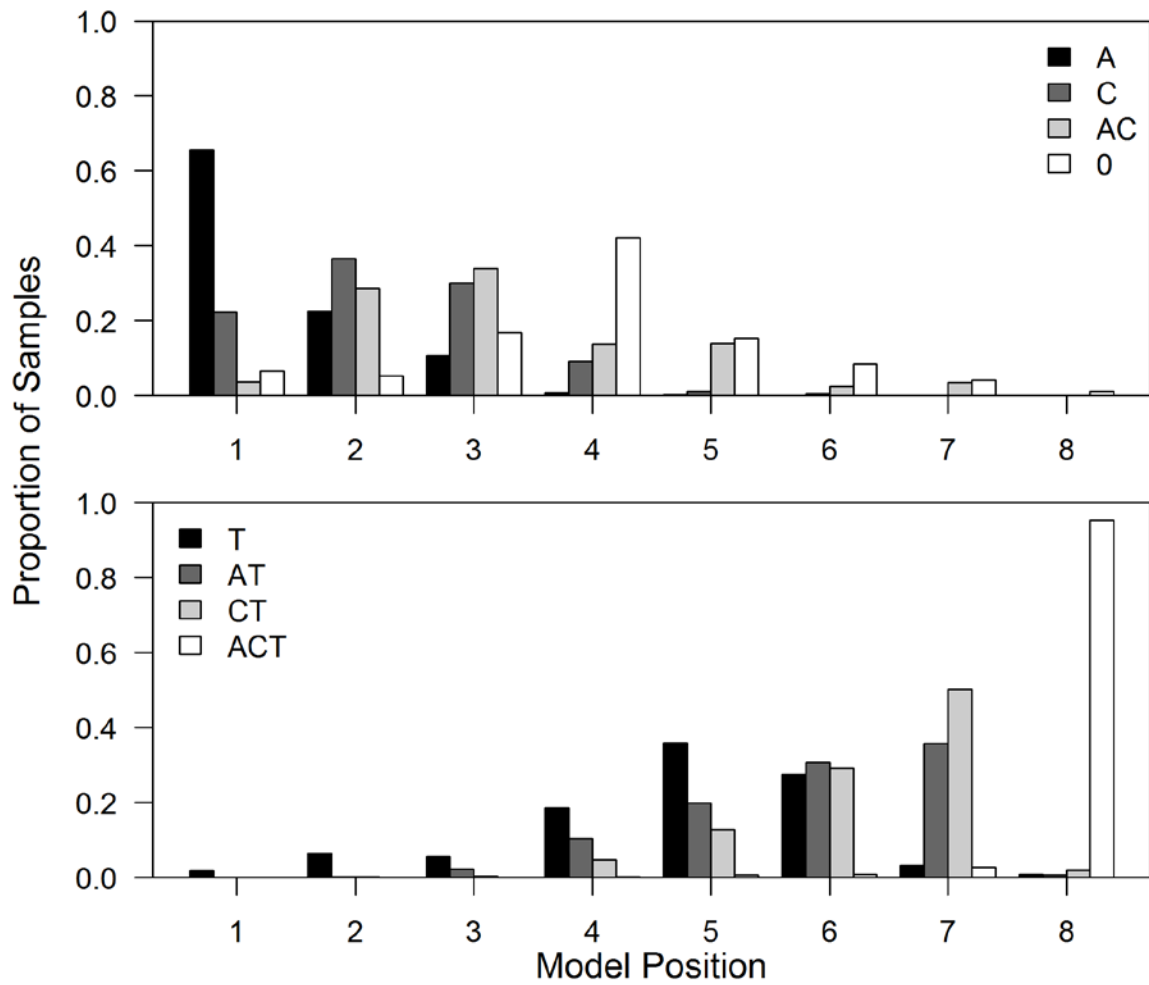


Figure 4. Distributions of model selection outcomes from the randomization procedures. Bar height represents the proportion of randomized posterior samples that each model showed up in each model placement. Divided into two panels for ease of presentation.

**APPENDIX A: DATA**

Table A1: Habitat variables before standardizing.

<b>Stock</b>	<b>A</b>	<b>T</b>	<b>C</b>
Alsek	10.31497	0.012856	0.007819
Anchor	6.331502	0.010676	0.007117
Chena+Salcha	9.304283	0.013382	0.019663
Deshka	7.407328	0.012743	0.027306
Goodnews	7.882315	0.012453	0.053962
Karluk	6.43935	0.011182	0.015974
Kuskokwim	11.6786	0.013218	0.067099
Stikine	10.82146	0.013238	0.044405
Taku	9.777868	0.013039	0.049606
Unalakleet	8.574329	0.013034	0.022667
Yukon	12.56375	0.012629	0.034927

Table A2:  $\log(S_{MSY})$  posterior summaries; used in DIC and indicator variable selection methods (note: SD was transformed to be an inverse variance for JAGS)

<b>Stock</b>	<b>Mean</b>	<b>SD</b>
Alsek	8.309908	0.277004
Anchor	8.411187	0.455808
Chena+Salcha	9.038798	0.107132
Deshka	9.351489	0.17418
Goodnews	8.168945	0.183109
Karluk	8.409064	0.617513
Kuskokwim	11.29503	0.257912
Stikine	9.743034	0.163446
Taku	10.23734	0.250491
Unalakleet	9.107516	1.104626
Yukon	10.33829	0.210793

## APPENDIX B. CODE FOR VARIABLE SELECTION TECHNIQUES

### # PACKAGES

```
library(R2OpenBUGS)
library(rjags)
source("C:/Users/bas0041/Desktop/run_functions_source.R")
main.dir=getwd()
```

### # RESPONSE VARIABLES: REFERENCE POINTS

```
col = c("mean", "sd", "median", "lower95", "upper95")
log.S.msy = read.csv(paste(main.dir, "log_Smsy.csv", sep="/"), row.names=1)
colnames(log.S.msy)=col
```

### # PREDICTOR VARIABLES: HABITAT VARIABLES

```
habitat = read.csv(paste(main.dir, "habitat_vars.csv", sep="/"), row.names=1)
```

### #pullout only stocks we can use

```
no.na = which(!is.na(habitat$WS_A_KM2) & !is.na(habitat$MCND_NUM))
tribs.nat = habitat$TND_NUM/habitat$WS_A_KM2
chans.nat = habitat$MCND_NUM/habitat$WS_A_KM2
```

### # habitat variables

```
habitat.vars = data.frame(area = log(habitat$WS_A_KM2), tribs = tribs.nat, chans = chans.nat)
habitat.vars = habitat.vars[no.na,]
rownames(habitat.vars) = rownames(habitat)[no.na]
```

### # reference points: use posterior means

```
log.S.msy.mean = log.S.msy[no.na, 1]
log.S.msy.tau = 1/(log.S.msy[no.na, 2])^2
```

```
dat = list(
  #dependent variables
  obs.mu = log.S.msy.mean,
  obs.tau = log.S.msy.tau,
  n.stocks = length(log.S.msy.mean),
```

```
  #independent variables
  area = as.numeric(scale(habitat.vars$area)),
  trib_nodes = as.numeric(scale(habitat.vars$tribs)),
  chan_nodes = as.numeric(scale(habitat.vars$chans)))
```

```
windows(record=T)
for(i in 4:6){
  plot(dat$obs.mu~dat[[i]], main=names(dat)[i])
}
```

### ### SPECIFY MODEL ###

```
mod=function(){
```

```

#PRIORS
for(i in 1:3){
  beta[i] ~ dnorm(0, 0.1)
  w[i] ~ dbern(0.5)
}
int ~ dnorm(0, 0.1)
tau.proc ~ dgamma(0.01, 0.01)

for(i in 1:n.stocks){
  obs.mu[i] ~ dnorm(ref.proc[i], obs.tau[i])
  ref.proc[i] ~ dnorm(pred.ref[i], tau.proc)

  pred.ref[i] <- int + w[1]*beta[1]*area[i] + w[2]*beta[2]*trib_nodes[i] + w[3]*beta[3]*chan_nodes[i]

  #derived quantities
  resid.pred[i] <- obs.mu[i]-pred.ref[i]
  resid.proc[i] <- obs.mu[i]-ref.proc[i]
}
sigma.proc <- 1/sqrt(tau.proc)
}

# WRITE MODEL TO A TXT FILE
model.file = "param_inclus.txt"
write.model(mod,paste(main.dir, model.file, sep="/"))

### PARAMETERS TO MONITOR ###
parameters = c(
  #parameters
  "beta", "w", "tau.proc", "ref.proc", "int",
  #calculated quantities
  "sigma.proc", "pred.ref", "resid.pred", "resid.proc")

# RUN FREQUENTIST MODEL TO GET INITIAL VALUES FOR PARAMETERS
freq.mod = lm(obs.mu~area+trib_nodes+chan_nodes, data = dat)
full = summary(freq.mod)
freq.coef = coef(full)[,1]
freq.sigma = full$sigma

# COMPILE INITS
inits1 = list(int=freq.coef[1], beta=freq.coef[2:4], w=rep(1, 3), tau.proc=1/freq.sigma^2)
inits2 =list(int=freq.coef[1]*runif(1), beta=freq.coef[2:4]*runif(1), w=rep(0, 3),
tau.proc=(1/freq.sigma^2)*runif(1))

inits = list(inits1, inits2)

# MCMC SPECIFICATIONS
nc = 2
ni = 100000

```

```

nb = 50000
nt = 20
ni/nt

### RUN JAGS ###

# INITITATE THE MODEL (ADAPTING PHASE)
jmod = jags.model(file=paste(main.dir, model.file, sep="/"), data=dat, n.chains=nc, inits=inits,
n.adapt=1000)

# BURNIN
update(jmod, n.iter=nb, by=1, progress.bar='text')

# SAMPLE THE POSTERIOR
post = coda.samples(jmod, parameters, n.iter=ni, thin=nt)

#### LOOK AT OUTPUT ####
gelman.diag(post, multivariate=F)

# COERCE TO MATRIX AND PULL OUT POSTERIOR OF INTEREST
mypost = as.matrix(post)
w = mypost[, c("w[1]", "w[2]", "w[3]")]
beta = mypost[,c("beta[1]", "beta[2]", "beta[3]")]

# remove w == 0 from betas
beta1 = beta[w[,1] == 1 ,1]
beta2 = beta[w[,2] == 1 ,2]
beta3 = beta[w[,3] == 1 ,3]

summ = function(x){c(mean = mean(x), quantile(x, c(0.025, 0.975))))}

eff = rbind(summ(beta1), summ(beta3), summ(beta2))
rownames(eff) = c("A", "C", "T")
#write.csv(eff, "effectsizes.csv")

### MODEL PROBABILITITES ###
models = as.matrix(table(paste(w[,1],w[,2], w[,3], sep="")))/dim(w)[[1]]
modprobs = as.matrix(as.matrix(models[order(models[,1],decreasing = T),]))
round(modprobs, 2)

# MODEL AVERAGED COEFFNS
parmas = beta*w
results = t(rbind(mean=colMeans(parmas),
sd=apply(parmas,2,sd),
apply(parmas,2,quantile,c(0.025,0.5, 0.975)),
inclusion_probs=colMeans(w)))

```

**##### END PARAMETER INCLUSION PROBS #####**

```
#####  
#  
##### RANDOMIZATION OF POSTERIOR SAMPLES W/AIC SELECTION  
#####  
#####  
#
```

#### **# PACKAGES**

```
library(MuMIn)  
library(pbapply)
```

```
main.dir=getwd()
```

#### **# READ IN THE POSTERIOR SAMPLES**

```
post.samp=read.csv("log.S.msy.post.csv")
```

#### **# PREDICTOR VARIABLES: HABITAT VARIABLES**

```
habitat = read.csv(paste(main.dir, "habitat_vars.csv", sep="/"), row.names=1)
```

#### **#pullout only stocks we can use**

```
no.na = which(!is.na(habitat$WS_A_KM2) & !is.na(habitat$MCND_NUM))# &  
rownames(habitat)!="kuskokwim")
```

```
tribs.nat = habitat$TND_NUM/habitat$WS_A_KM2  
chans.nat = habitat$MCND_NUM/habitat$WS_A_KM2
```

#### **#habitat variables**

```
habitat.vars = data.frame(area = log(habitat$WS_A_KM2),tribs = tribs.nat, chans = chans.nat)  
habitat.vars = habitat.vars[no.na,]  
rownames(habitat.vars) = rownames(habitat)[no.na]
```

```
use.stocks=rownames(habitat.vars)  
use.stocks[length(use.stocks)]="yukon"
```

#### **#pull out only those stocks' posteriors**

```
post.samp=post.samp[,use.stocks]  
colnames(post.samp)
```

#### **# RANDOMIZE THE POSTERIOR FOR EACH STOCK**

```
n.samp=nrow(post.samp)  
post.samp=as.matrix(post.samp)  
sub.samp=apply(post.samp, 2, function(x) sample(x, n.samp, replace=F))  
models = c("null", "area", "chans", "area.chans", "tribs", "area.tribs", "chans.tribs", "full")
```



```
n.pulls = n.samp  
do.plot = F
```

#### **# CONTAINERS FOR RANDOMIZED RESULTS**

```
boot.results = list()  
null.pos = numeric(n.pulls)  
null.wt = numeric(n.pulls)  
area.pos = numeric(n.pulls)  
area.wt = numeric(n.pulls)  
chans.pos = numeric(n.pulls)  
chans.wt = numeric(n.pulls)  
area.chans.pos = numeric(n.pulls)  
area.chans.wt = numeric(n.pulls)  
tribs.pos = numeric(n.pulls)  
tribs.wt = numeric(n.pulls)  
area.tribs.pos = numeric(n.pulls)  
area.tribs.wt = numeric(n.pulls)  
chans.tribs.pos = numeric(n.pulls)  
chans.tribs.wt = numeric(n.pulls)  
full.pos = numeric(n.pulls)  
full.wt = numeric(n.pulls)
```

#### **# INITIATE THE FITTING TO POSTERIOR SAMPLES**

```
pb = startpb(0, n.pulls)  
for(i in 1:n.pulls){  
  Sys.sleep(0.1)  
  
  # make a new data set  
  dat = data.frame(area = as.numeric(scale(habitat.vars$area)),  
                    tribs = as.numeric(scale(habitat.vars$tribs)),  
                    chans = as.numeric(scale(habitat.vars$chans)), smsy = sub.samp[i,])  
  
  if(i == 1 & do.plot == T) windows(record = T)  
  if(do.plot == T) {  
    par(mfrow = c(2,2), oma = c(0, 1, 2, 0), mar = c(2,2,2,2))  
    plot(smsy ~ area, data = dat, main = "Area", xlab = "", ylab = "")  
    plot(smsy ~ chans, data = dat, main = "Chans", xlab = "", ylab = "")  
    plot(smsy ~ tribs, data = dat, main = "Tribs", xlab = "", ylab = "")  
    mtext(side = 3, outer = 2, as.character(i), cex = 2, line = 0, font = 2)  
    plot.new()  
  }  
  
  # fit all models  
  fit = lm(smsy ~ area + chans + tribs, data = dat, na.action = na.fail)  
  
  dd = dredge(fit)
```

#### **#combine important model selection**

```
boot.results[[i]] = data.frame(model = models[as.numeric(rownames(dd))], weight = dd$weight)
```

```

null.pos[i] = which(boot.results[[i]]$model == "null")
null.wt[i] = boot.results[[i]]$weight[null.pos[i]]

area.pos[i] = which(boot.results[[i]]$model == "area")
area.wt[i] = boot.results[[i]]$weight[area.pos[i]]

chans.pos[i] = which(boot.results[[i]]$model == "chans")
chans.wt[i] = boot.results[[i]]$weight[chans.pos[i]]

area.chans.pos[i] = which(boot.results[[i]]$model == "area.chans")
area.chans.wt[i] = boot.results[[i]]$weight[area.chans.pos[i]]

tribs.pos[i] = which(boot.results[[i]]$model == "tribs")
tribs.wt[i] = boot.results[[i]]$weight[tribs.pos[i]]

area.tribs.pos[i] = which(boot.results[[i]]$model == "area.tribs")
area.tribs.wt[i] = boot.results[[i]]$weight[area.tribs.pos[i]]

chans.tribs.pos[i] = which(boot.results[[i]]$model == "chans.tribs")
chans.tribs.wt[i] = boot.results[[i]]$weight[chans.tribs.pos[i]]

full.pos[i] = which(boot.results[[i]]$model == "full")
full.wt[i] = boot.results[[i]]$weight[full.pos[i]]
setpb(pb,i)
}

# close the progress bar
closepb(pb)

# COMBINE RAW OUTPUT
positions = data.frame(null = null.pos, area = area.pos, chans = chans.pos, area.chans = area.chans.pos,
  tribs = tribs.pos, area.tribs = area.tribs.pos, chans.tribs = chans.tribs.pos, full = full.pos)
weights = round(data.frame(null = null.wt, area = area.wt, chans = chans.wt, area.chans = area.chans.wt,
  tribs = tribs.wt, area.tribs = area.tribs.wt, chans.tribs = chans.tribs.wt, full = full.wt),2)

# SUMMARIZE RAW OUTPUT
positions.mean = sort(apply(positions, 2, mean))
positions.sd = apply(positions, 2, sd)[names(positions.mean)]

weights.mean = sort(apply(weights, 2, mean), decreasing = T)
weights.sd = apply(weights, 2, sd)[names(weights.mean)]

short.models = c("0", "A", "C", "AC", "T", "AT", "CT", "ACT")

names(short.models) = models

# PLOT MEAN MODEL POSTION AND WEIGHT

```

```

windows()
par(mfrow = c(2,1), mar = c(3, 3, 2, 2))
mp.1 = barplot(positions.mean, names.arg = short.models[names(positions.mean)], main = "Mean
Model Position", col = "white",
               ylim = c(0, max(positions.mean + positions.sd) + 1), las = 1)
arrows(mp.1, positions.mean - positions.sd, mp.1, positions.mean + positions.sd, code = 3, length = 0.05,
       angle = 90)
axis(side = 1, at = mp.1, labels = rep("", length(models)))
box()

mp.2 = barplot(weights.mean, names.arg = short.models[names(weights.mean)], main = "Mean Model
Weight", col = "white",
               ylim = c(0, max(weights.mean + weights.sd) + 0.1), las = 1)
axis(side = 1, at = mp.2, labels = rep("", length(models)))
arrows(mp.2, weights.mean - weights.sd, mp.2, weights.mean + weights.sd, code = 3, length = 0.05,
       angle = 90)
box()

```

#### **# PULL OUT DISTRIBUTION OF MODEL PLACEMENT**

```

pos.count = apply(positions, 2, function(x) table(x))
count.mat = matrix(NA, nrow = 8, ncol = 8)
colnames(count.mat) = as.character(1:8)
rownames(count.mat) = models

for(i in 1:length(models)){
  count.mat[models[i], as.character(names(pos.count[[i]]))] = pos.count[[i]]
}

```

```

count.mat[is.na(count.mat)] = 0
prop.mat = count.mat/n.samp
best = prop.mat[c("area", "chans", "area.chans", "null"),]
worst = prop.mat[c("tribs", "area.tribs", "chans.tribs", "full"),]

```

#### **# PLOT DISTRIBUTIONS OF MODEL PLACEMENT**

```

windows()
par(mfrow = c(2,1), mar = c(2,2,3,2), oma = c(3,3,0,0))
barplot(best, beside = T, col = c("black", "grey40", "grey80", "white"), cex.lab = 1.5, cex.main = 2, ylim =
c(0, 1), las = 1, main = "4 Best Models", xlab = "", ylab = "Frequency")
axis(side = 1, at = seq(3, 40, by = 5), labels = rep("", 8))
legend("topright", legend = short.models[rownames(best)], fill = c("black", "grey40", "grey80", "white"),
      bty = "n")

barplot(worst, beside = T, col = c("black", "grey40", "grey80", "white"), cex.lab = 1.5, cex.main = 2, ylim =
c(0, 1), las = 1, main = "4 Worst Models", xlab = "", ylab = "")
axis(side = 1, at = seq(3, 40, by = 5), labels = rep("", 8))
legend("topleft", legend = short.models[rownames(worst)], fill = c("black", "grey40", "grey80", "white"),
      bty = "n")

```

```

mtext(side = 1, outer = T, "Model Position", cex = 1.5, line = 0.6)
mtext(side = 2, outer = T, "Proportion of Samples", cex = 1.5, line = 1)

```

#### **# NUMBER OF TIMES EACH VARIABLE SHOWED UP IN THE BEST MODEL**

```

area = sum(count.mat[c("area", "area.chans", "area.tribs", "full"), 1])
chans = sum(count.mat[c("chans", "area.chans", "chans.tribs", "full"), 1])
tribs = sum(count.mat[c("tribs", "area.tribs", "chans.tribs", "full"), 1])
null = sum(count.mat["null",1])
tot = sum(area, chans, tribs)
param.inclus = round(t(data.frame(area = area/tot, chans = chans/tot, tribs = tribs/tot)),2)

```

#### **##### END RANDOMIZATION CODE #####**

```

#####
#
##### LEAVE-ONE-OUT CROSS VALIDATION
#####
#

```

#### **##### LOAD PACKAGES AND MY FUNCTIONS #####**

```

library(R2OpenBUGS)
library(rjags)
source("C:/Users/bas0041/Desktop/run_functions_source.R")

main.dir=getwd()

```

#### **##### GET RAW DATA READ IN AND FORMATED #####**

##### **# RESPONSE VARIABLES: REFERENCE POINTS**

```
col = c("mean", "sd", "median", "lower95", "upper95")
```

##### **# log of ref. posteriors with the negatives removed**

```
log.S.msy = read.csv(paste(main.dir, "log_Smsy.csv", sep="/"), row.names=1); colnames(log.S.msy)=col
```

##### **# PREDICTOR VARIABLES: HABITAT VARIABLES**

```
habitat = read.csv(paste(main.dir, "habitat_vars.csv", sep="/"), row.names=1)
```

#### **#pullout only stocks we can use**

```
no.na = which(!is.na(habitat$WS_A_KM2) & !is.na(habitat$MCND_NUM))# &  
rownames(habitat)!="kuskokwim")
```

```
tribs.nat = habitat$TND_NUM/habitat$WS_A_KM2  
chans.nat = habitat$MCND_NUM/habitat$WS_A_KM2
```

#### **#habitat variables**

```
habitat.vars = data.frame(area = log(habitat$WS_A_KM2),tribs = tribs.nat, chans = chans.nat)  
habitat.vars = habitat.vars[no.na,]  
rownames(habitat.vars) = rownames(habitat)[no.na]
```

#### **#reference points: use posterior means**

```
log.S.msy.mean = log.S.msy[no.na, 1]  
log.S.msy.tau = 1/(log.S.msy[no.na, 2])^2
```

```
area = as.numeric(scale(habitat.vars$area))  
trib = as.numeric(scale(habitat.vars$tribs))  
chan = as.numeric(scale(habitat.vars$chans))
```

#### **##### SPECIFY MODEL #####**

```
mod=function(){  
  #PRIORS  
  for(i in 1:nvars){  
    beta[i] ~ dnorm(0, 0.1)  
  }  
  #int ~ dnorm(0, 0.1)  
  tau.proc ~ dgamma(0.01, 0.01)  
  
  for(i in 1:n.stocks){  
    obs.mu[i] ~ dnorm(ref.proc[i], obs.tau[i])  
    ref.proc[i] ~ dnorm(pred.ref[i], tau.proc)  
  
    pred.ref[i] <- sum(beta * X[i,])  
  
    #derived quantities  
    resid.pred[i] <- obs.mu[i] - pred.ref[i]  
    resid.proc[i] <- obs.mu[i] - ref.proc[i]  
  }  
  sigma.proc <- 1/sqrt(tau.proc)  
}
```

```
model.file = "model.txt"  
write.model(mod,paste(main.dir, model.file, sep="/"))
```

#### **##### PARAMETERS TO MONITOR #####**

```

parameters = c(
  #parameters
  "beta", "tau.proc", "ref.proc",
  #calculated quantities
  "sigma.proc", "pred.ref", "resid.pred", "resid.proc")

##### MCMC SPECIFICATIONS #####
nc = 2
ni = 100000
nb = 50000
nt = 20
ni/nt
n.saved = (ni * nc)/nt

##### MAKING DATA FOR EACH MODEL AND FOR EACH LEAVE ONE OUT RUN #####
# habitat data
K = 11
n.stocks = 10
int = rep(1,n.stocks)
hab = list()
for(k in 1:K){
  hab[[k]] = list(mod1 = cbind(int, area[-k]), mod2 = cbind(int, trib[-k]), mod3 = cbind(int, chan[-k]),
    mod4 = cbind(int, area[-k], trib[-k]), mod5 = cbind(int, area[-k], chan[-k]),
    mod6 = cbind(int, trib[-k], chan[-k]), mod7 = cbind(int, area[-k], trib[-k], chan[-k]),
    mod8 = cbind(int))
}

# reference points
ref = list()
for(k in 1:K){
  ref[[k]] = list(obs.mu = log.S.msy.mean[-k], obs.tau = log.S.msy.tau[-k])
}

# number of predictors for each model
n.vars = c(2,2,2,3,3,3,3,4,1)

##### RUN ALL 8 MODELS OVER ALL 11 COMBINATIONS OF THE LEAVE-ONE-OUT #####
n.models = 8
beta = list()
tau.proc = list()
starttime = Sys.time()
# i models
# k datasets
for(i in 1:n.models){
  beta[[i]] = list()
  tau.proc[[i]] = list()
  for(k in 1:K){
    ### SET UP DATA ###

```

```

dat = list(
  obs.mu = ref[[k]]$obs.mu,
  obs.tau = ref[[k]]$obs.tau,
  n.stocks = n.stocks,
  nvars = n.vars[i],

  X = matrix(unlist(hab[[k]][i]), nrow = 10, ncol = n.vars[i])
)

### RUN JAGS ###
print(paste("Model", as.character(i), "running with stock", as.character(k), "removed"))
jmod = jags.model(file=paste(main.dir, model.file, sep="/"), data=dat, n.chains=nc, inits=NULL,
n.adapt=1000, quiet = T)
update(jmod, n.iter=nb, by=1, progress.bar='text')
post = coda.samples(jmod, parameters, n.iter=ni, thin=nt)

### PULL OUT OUTPUT ###
if(i == 8) beta.name = "beta" else beta.name = "beta["
beta[[i]][[k]] = get.post(post, beta.name, do.post = T)$posterior
tau.proc[[i]][[k]] = get.post(post, "tau.proc", do.post = T)$posterior
}
}
Sys.time() - starttime

##### CALCULATE THE SCORE FROM HOOTEN AND HOBBS (2015) #####
# get the habitat variables with all data
n.stocks.full = 11
int.full = rep(1,n.stocks.full)
hab.full = list(mod1 = cbind(int.full, area), mod2 = cbind(int.full, trib), mod3 = cbind(int.full, chan),
  mod4 = cbind(int.full, area, trib), mod5 = cbind(int.full, area, chan),
  mod6 = cbind(int.full, trib, chan), mod7 = cbind(int.full, area, trib, chan),
  mod8 = cbind(int.full))

# get the reference points with all data
ref.full = log.S.msy.mean

# calculate model predicted S.msy for the left out stock under each model
y.minus.k = list()
for(i in 1:n.models){
  y.minus.k[[i]] = list()
  for(k in 1:K){
    y.minus.k[[i]][[k]] = numeric(n.saved)
    for(t in 1:n.saved){
      if(any(i == 1:7)){
        y.minus.k[[i]][[k]][t] = sum(beta[[i]][[k]][t,] * hab.full[[i]][k,])
      }
    }
    if(i == 8){

```

```

        y.minus.k[[i]][[k]][t] = sum(beta[[i]][[k]][t] * hab.full[[i]][k,])
    }
}
}
}

```

**# use sum of squared residual as the measure of statistical "closeness"**

```

score.i.k = list()
for(i in 1:n.models){
  score.i.k[[i]] = numeric(11)
  for(k in 1:K){
    score.i.k[[i]][k] = sum((ref.full[k] - y.minus.k[[i]][[k]])^2)/n.saved
  }
}

```

**##### INFERENCE #####**

```

scores = unlist(lapply(score.i.k, sum))
model.names = c("A", "T", "C", "AT", "AC", "CT", "ACT", "0")
names(scores) = model.names

```

```

sorted = round(sort(scores),2)
best = min(sorted)

```

```

pct.chng = round((sorted - best)/best,2)

```

**##### END LEAVE-ONE-OUT CROSS VALIDATION CODE #####**