Development and Evaluation of Assessment Tools and Management Strategies for Salmon Fisheries in Western Alaska

by

Benjamin A. Staton

A Dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 4, 2019

Keywords: Pacific salmon, phenological forecasts, management strategy evaluation,
mixed-stock fisheries, Bayesian inference

Approved by

Matthew J. Catalano, Chair, Associate Professor of Fisheries Science
Asheber Abebe, Professor of Mathematics and Statistics
Lewis G. Coggins, Jr., Affiliate Professor of Fisheries Science
Conor P. McGowan, Professor of Wildlife Sciences
George T. Flowers, Graduate School Dean

Abstract

The management of natural resources is fraught with difficulties stemming from uncertainty and conflicting objectives. Fisheries for Pacific salmon *Oncorhynchus* spp. in western Alaska are no exception, but rather provide a fantastic example. This area of the world is tremendously remote, making resource monitoring and decision-making challenging. It has widely been proposed that quantitative tools can be useful in aiding decision-making by producing predictions of uncertain states of nature, system responses to harvest, and quantifying likely outcomes of candidate management actions. In this dissertation, several such quantitative tools seeking to serve these purposes are developed and evaluated. Throughout, the Kuskokwim River subsistence salmon fishery (one of the largest in the world) is used as a case study to illustrate the development, evaluation, and application of these tools.

Before investigating three primary research topics in Chapters 2, 3, and 4, an overview of the difficulties faced by practitioners of salmon management is provided in Chapter 1. The management of salmon resources is presented as a three-tiered decision-making hierarchy made up of (1) fundamental objectives stemming from societal values, (2) inter-annual management strategies which define how the objectives are to be attained in the long-term, and (3) finer-scale (*e.g.*, in-season) tactics used to implement the strategy. The important considerations and key uncertainties at play at each level are discussed, as is the role of quantitative tools in decision-making.

Chapter 2 focuses on salmon migration timing, the problems that its inter-annual variability inserts for salmon managers and assessments of run abundance, and the development and evaluation of a forecasting tool that attempts to predict the timing of the run before any fish arrive. Run timing variability is a well-known and pervasive problem: in-season

abundance index data are often consistent with many different run scenarios ranging from small and early to large and late, making decisions about the magnitude of harvestable surplus difficult. Run timing has been widely shown to covary with environmental variables linked to temperature such that early run years tend to coincide with warm years, suggesting that useful predictive relationships may exist. A statistically rigorous approach to forecasting the date of 50% run completion for Kuskokwim River Chinook salmon *O. tshawytscha* was developed based on relationships with temperature-related variables. An objective and intuitive temporal variable selection approach (the sliding climate window algorithm) was employed to determine the time periods for each variable most likely to produce accurate forecasts and the analysis relied heavily on multi-model inference in the face of a high degree of model uncertainty. The rather complex forecasting framework was found to perform no better on average at forecasting the median run date than the most naïve model that used solely the historical average as the forecast. Although the environmental variable forecast did not have value in terms of improving the accuracy of interpretations of an in-season run abundance index, its use did reduce the statistical uncertainty in this index, particularly early in the season. The analysis in this chapter should serve as a useful example for researchers faced with high-dimensional spatio-temporal variable selection problems, particularly with respect to the formal evaluation of the performance of alternative forecasts using retrospective cross-validation techniques.

Chapter 3 takes a broader view of the in-season management problem by simulation-testing a set of four harvest control strategies in a framework known as stochastic management strategy evaluation. A detailed mathematical caricature of the Kuskokwim River salmon system was constructed based on empirical data to serve as the operating model with which to test different decision rules for determining how many days the fishery should be open each week. Strategy performance was assessed relative to four pre-defined management objectives dealing with both conservation and exploitation such that key trade-offs could be identified.

A key research question was regarding strategy complexity and performance: whether more involved and data-intensive feedback strategies should be favored over simpler fixed-schedule strategies. The primary finding revealed by this work was that several strategies ranging from simple to complex can perform essentially equally well at attaining the objectives of salmon management in the Kuskokwim River as defined by the utility functions used. This finding suggests that in-season management may be made more difficult than needed, and that perhaps more focus should be placed on the transparency and defensibility of strategies.

Chapter 4 takes an even broader view still, and addresses the topic of assessment of salmon fisheries that harvest fish from multiple substocks as a mixed-stock. Different substocks within a larger salmon-producing drainage vary in their size and resilience to harvest pressure (*i.e.*, intrinsic productivity), which suggests a trade-off between maximizing harvest and preserving substock biodiversity exists. To incorporate this trade-off into harvest policies, an assessment of the heterogeneity in substock size and productivity must be completed. Salmon populations are often assessed as single stock units using simple regression approaches, however, this chapter makes a compelling case for integrating data from multiple substocks into a single assessment model that represents population dynamics and observation processes at the substock-level using a state-space analytical framework. A range of assessment methods were constructed that varied in their assumptions about process variability and were applied to empirical data from Kuskokwim River Chinook salmon substocks and evaluated in a simulation-estimation context. All state-space models assessed performed substantially better than regression-based approaches at returning accurate and precise estimates of biological reference points at the substock- and mixed-stock level, and by capturing the sources of variability allowed for more rich ecological interpretations which may be useful in future policy analyses.

The dissertation concludes with Chapter 5 which presents further reflection on the utility, performance, and generality of the tools developed in these studies.

Acknowledgments

I am grateful to many people for helping to make my graduate career successful and enjoyable. First, I would like to thank my major professor, Dr. Matt Catalano, for his superb mentorship, providing and allowing me to pursue many challenging opportunities, and granting me freedom and flexibility in how I completed my work. I have had many other mentors as well: Dr. Lew Coggins, Dr. Mike Jones, Dr. Brendan Connors, Steve Fleischman, and Dr. Dan Gwinn have all been instrumental in shaping me as a quantitative scientist and I am immensely fortunate to consider them as such. In completing my academic and applied work, I have had the opportunity to work with many dedicated professionals: Dr. Troy Farmer, Nick Smith, Zach Liller, Janessa Esquible, Gary Decossas, Dr. Bill Bechtol, Aaron Moses, and Ken Stahlnecker; all of whom have made the experience enjoyable. I have received top-notch instruction and advice on my work while at Auburn; specifically, I would like to thank: Drs. Ash Abebe and Stephen Dobson (for their insights on Chapter 2), Dr. Conor McGowan (for his course on structured decision making and useful modeling techniques in that context), Dr. Todd Steury (for serving as an external university reader and for his excellent course introducing a wide range of statistical methods in ecology, which provided me with a solid foundation in linear modeling early in my graduate career) and Dr. Laurie Stevison (for her gentle but thorough introduction to Bash Shell programming, which allowed me to conduct the Chapter 4 analysis on the Alabama High-Performance Computing [HPC] cluster). My work would not be possible without the dedication of the developers of software I used; I extend many thanks to the R Core Development Team for building a comprehensive and intuitive statistical environment, Dr. Martyn Plummer for developing JAGS, and the team at RStudio for their contributions in making working in R pleasurable in all aspects of

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Wild Pacific salmon *Oncorhynchus* spp. represent a fantastic natural resource, which results largely from their unique life history strategy. Pacific salmon exhibit a migratory strategy known as anadromy: adults spawn in freshwater where eggs hatch and juveniles rear for 0, 1, or 2 years. Juveniles then migrate to the ocean where they spend the majority of their lives feeding on abundant prey resources. Once reaching maturity, adults return to their natal streams to spawn and complete the life cycle. The result of this life history strategy is an incredibly productive resource that grows entirely on its own and all but delivers itself to harvesters when the time comes for exploitation.

There is a long history of salmon fishery resource development, exploitation, regulation, and dependence throughout Alaska (Cooley 1963). In many cases, the resource use is dictated by the locality of the system; for example, stocks located near urban areas are often primarily exploited by recreational fishers whereas more remote stocks often constitute commercial and/or subsistence uses. This dissertation discusses the challenges and explores quantitative tools for assessing and informing management of more remote stocks and the fisheries that rely heavily upon them, with particular relevance to the remote areas of western Alaska.

Similar to all exploited natural resources, the management of Pacific salmon fisheries involves making decisions about how to exploit the resource in order to best attain a suite of biological, social, and economic objectives (Walters 1986). These decisions are inherently difficult due to conflicting objectives and uncertainties in system state, system response to management actions, and implementation (Walters and Holling 1990). Put another way, assuming a manager knows exactly what they wish to obtain, getting there is made difficult

by not knowing (for example) how large the harvestable surplus is, how the stock will respond to harvesting, or that their management action will actually result in the expected outcome. Despite these difficulties, a decision must be made (without decision-making there is no management; Hilborn and Walters 1992) and the consequences, whether favorable or undesirable, must be accepted. Thus, it could be argued that the science of monitoring, assessment, and prediction in the context of Pacific salmon fisheries is tasked with informing the relative likelihood of different outcomes conditional on a candidate management action.

The management of Pacific salmon fisheries can be thought of as a hierarchy of (1) guiding objectives, (2) management strategies to attain objectives, and (3) tactics to implement the management strategies (Table 1.1). At the upper level, long-term decisions are made about the objectives of resource exploitation. These long-term objectives are often referred to as fundamental objectives: they are desired endpoints, but do not at all imply how they should be attained. Fundamental objectives often involve notions of sustainability and maintenance of biological diversity and often include social objectives such as maximization and stability of harvest or profit. Already, it is clear that these fundamental objectives are often conflicting. For example, consider the objective of maximizing harvest: in fisheries that harvest multiple stocks (*i.e.*, distinct spawning units), oftentimes maximum harvest may only obtained by overexploiting weak stock components and possibly eroding diversity. As another example, consider the objective of long-term sustainability: in order to ensure that the stock is sustained, some level of harvest fluctuations must be accepted (lower harvests must be allowed when the stock is at low abundance). These conflicting objectives imply that trade-offs exist – all objectives cannot be maximized simultaneously. It is worth noting here that the decisions made at the uppermost level of the management hierarchy are based purely on societal values and salmon stock assessment scientists should play little or no advisory or advocacy roles in making many of these decisions, except to the extent that they

2

are also members of society (Walters and Martell 2004). The Policy for the Management of Sustainable Salmon Fisheries[1] states that the objectives of salmon management in Alaska are

> "...to ensure conservation of salmon and salmon's required marine and aquatic habitats, protection of customary and traditional subsistence uses and other uses, and the sustained economic health of Alaska's fishing communities."

The policy goes on to say that managers should target "...to the extent possible, maximum sustained yield [MSY]."

The second level of the management hierarchy is made up of harvest strategies and policies that guide how the long term objectives are to be obtained. The State of Alaska has selected the fixed escapement policy as the management strategy to obtain the long-term objectives of sustainability and yields that are close to the maximum. These escapement goals are given as ranges that dictate the target number of spawning adults each year; any portion of the stock above the escapement goal is considered surplus (excess biological production) and should be harvested for the benefit of society. Uncertainty at this intermediate level of the management hierarchy (*e.g.*, regarding the optimal escapement goal) is often a result of incomplete understanding of system status and function. For example, in order to determine what the optimal escapement goal should be to obtain MSY, knowledge of stock productivity and carrying capacity are required. These quantities are often derived using spawner-recruit analyses (see Walters and Martell 2004, Ch.7 for an overview), which are inherently uncertain: data are rarely informative about the shape of the true underlying population dynamics relationships (Walters and Hilborn 1976), but instead provide snapshot in time (*e.g.*, 20+ years) of how the population has responded to its environment and harvesting, and are often fraught with measurement errors (Ludwig and Walters 1981). Traditionally, it has been thought that these uncertainties can be reduced by more monitoring and the development of

---

[1] 5 AAC 39.222; a piece of legislation that defines correct salmon management practices by the Alaska Department of Fish and Game. Available at: http://www.adfg.alaska.gov/static/regulations/regprocess/fisheriesboard/pdfs/2016-2017/jointcommittee/5aac39.pdf

rigorous assessment and prediction models to better understand system function. However, it has often been argued that while monitoring and assessment models are obviously important (performance relative to objectives must be measured after all), true understanding of system behavior comes only from experimentation in management (the concept of "active adaptive management"; Walters 1986). A classic example is to assess the maximum productivity of the stock (*i.e.*, in the absence of density dependent mortality), the spawning stock must be forced to small sizes and the resulting distribution of recruitments must be observed (Walters and Hilborn 1976). However, management actions that ensure these observations are made may be undesirable to many managers and stakeholders, considering that exploiting a stock down to these low levels is risky (Walters 1986).

At the lowest level in the management hierarchy, intra-annual (or in-season) decisions are made regarding how to exploit the current year's run according to the rules of the strategy defined in the intermediate decision level. In other words, given a management strategy (*i.e.*, fixed escapement), the manager is still tasked with deciding how to best implement the fishery within a year to ensure the strategy is followed. Salmon runs are notoriously short in duration (*e.g.*, 80% of the the run may pass in a two or three week period), necessitating swift decision-making. As is illustrated in this dissertation, these decisions at the intra-annual level of the management hierarchy are often poorly informed by data which can result in indecisiveness, subjectivity, non-transparency, frustration, and missed opportunities.

This dissertation is partitioned into three primary projects (presented in Chapters 2, 3, and 4), each which expands on the aforementioned difficulties in decision-making and develops and implements quantitative tools intended to help guide managers of Pacific salmon fisheries. Each chapter relies on the Kuskokwim River drainage in western Alaska as a case study, which is characterized by being a large drainage ($>$50,000 km^2), harvests are taken by primarily subsistence users who are nearly all native Alaskans, and the primary species of interest being Chinook salmon *O. tshawytscha*. This system supports one of the largest

subsistence fisheries for Pacific salmon in the world in terms of the total numbers of fish harvested annually. Between 1995 and 2015, the average number of salmon harvested of all species for subsistence purposes in the Kuskokwim drainage was 201,000 (range: 140,000 – 294,000) making up an average of 21% (range: 16 – 28%) of all state-wide subsistence salmon harvests (data reported in annual reports, most recently in Fall *et al.* 2018). The Kuskokwim River contains the largest subsistence fishery for Chinook salmon in the state: an average of 65,000 Chinook salmon were harvested in this region between 1995 and 2015 (range 15,000 – 104,000), composing 48% (range: 34 – 61%) of state-wide subsistence harvests for this species annually (Fall *et al.* 2018). Although this dissertation is relatively narrow in its geographical and biological focus, the concepts and tools discussed, developed, and evaluated have broad generality and will be of interest to practitioners in other systems with similar spatial structures, exploitation characteristics, and/or population dynamics.

Chapter 2 works at the in-season level of the hierarchy to develop and evaluate the performance of a run timing forecast model that can be used to aid in the interpretation of in-season data. It includes an illustration of why uncertainty in run timing makes the interpretation of in-season abundance data difficult and a review of what is known about mechanisms driving variability of Pacific salmon run timing. The overall objective of Chapter 2 is to develop and evaluate the reliability of a run timing forecast model for Kuskokwim River Chinook salmon. A secondary goal of Chapter 2 is to retrospectively assess the utility of having access to the run timing forecast model in terms of reducing uncertainty and bias in run size indices used for in-season harvest management decisions.

Chapter 3 again addresses the lowest level of the management hierarchy (*i.e.*, intra-annual decision-making), but in this case in a more direct sense using an analysis framework known broadly as management strategy evaluation (MSE; *e.g.*, Butterworth 2007; Punt *et al.* 2014). This analysis evaluates a set of harvest control decision rules to identify strategies that perform well at attaining pre-defined objectives (*e.g.*, meeting the escapement goal,

distributing harvest equally across villages and substock components, *etc.*) across a range of biological states (*e.g.*, run size, stock composition, and run timing). The strategies assessed in this chapter fall along a continuum of complexity in their decision rules and the resulting increase in information requirements. While the fixed escapement policy seems simple to execute, actually doing so is made difficult largely due to uncertainty regarding the size of the incoming run. The MSE framework allows testing different rules for decision-making with sparse information, and may elucidate those that are robust to uncertainty. Additionally, there may be a set of decision rules that perform well at limiting harvest in low run size years but by doing so in a "fair way", where the burdens of shortages are not carried primarily by any particular subset of resource users, nor are the harvest burdens borne by a select subset of the substocks spawning within the larger drainage. If a consistent set of rules or triggers could be identified that perform reasonably well at meeting management objectives without precise knowledge of run size or harvestable surplus, it could prove useful to managers and decision-making within the region.

Chapter 4 moves up the hierarchy to the second level and attempts to extend the single stock assessment models currently used in many systems in Alaska to multi-stock assessments. When an aggregate stock is made up of several distinct components, each with their own productivity, it is likely that exploitation at some level (*e.g.*, 50% annually) results in the more productive components being under-exploited while the weaker stocks may be over-exploited. This reality implies a trade-off: to preserve stock diversity, some harvest must be foregone. Before the shape and magnitude of these types of "harvest-biodiversity" trade-offs can be quantified, some understanding of the variation in substock productivity and carrying capacity is required. The multi-stock assessment framework developed in Chapter 4 is tailored to provide this information for these sorts of trade-off analyses and others that require similar information sources. Multi-stock assessments may assume one of several different model structures (*e.g.*, by fitting separate models to the data from each stock or by fitting a single

model to all data simultaneously). In some cases, one approach may be preferable over the other, and a primary objective of Chapter 4 is to evaluate the estimation performance of a range of assessment strategies.

TABLE 1.1: One way of viewing the structure of renewable natural resource (including salmon) management as described in the text, including examples of alternatives and sources of uncertainty at each level.

Examples	Sources of Uncertainty
Fundamental Objectives	
Ensure sustainability	Relative importance of objectives
Maximize harvest	Problem boundaries
Stabilize harvest	
Maximize economic value	
Inter-annual Strategies	
Constant escapement	Stock productivity
Constant exploitation rate	Stock status
Constant catch	Drivers of stock change
Adaptive exploitation	Shape/magnitude of trade-offs
Intra-annual Tactics	
Triggers and thresholds	Harvestable surplus
Time, area, gear restrictions	Uninformative data
Limited participation	Fisher behavior

Chapter 2

Development and Evaluation of a Migration Timing Forecast Model for Kuskokwim River

Chinook Salmon

Note: The majority of this chapter has been published in Staton *et al.* (2017a). Changes included herein since publication include (*a*) an updated analysis with two additional years (2017 and 2018) (*b*) some embellishment on the sliding climate window approach, (*c*) several new tables and figures, and (*d*) revision of all figures to reflect additional data. All qualitative inferences made herein are identical to those published in Staton *et al.* (2017a).

Additionally, a more comprehensive analysis of the utility of the run timing forecast model for interpreting in-season run abundance index data is presented in Staton and Catalano (2019).

Abstract

Annual variation in adult salmon migration timing makes the interpretation of in-season assessment data difficult, contributing to in-season uncertainty about run size. I developed and evaluated a run timing forecast model for the Kuskokwim River Chinook salmon stock, located in western Alaska, intended to aid in reducing this source of uncertainty. An objective and adaptive approach (using model-averaging and a sliding window algorithm to select predictive time periods, both calibrated annually) was adopted to deal with multi-dimensional selection of four climatic variables and was based entirely on predictive performance. Forecast cross-validation was used to evaluate the performance of three forecasting approaches: the null (*i.e.*, intercept only) model, the single model with the lowest mean absolute error, and a model-averaged forecast across 16 nested linear models. As of 2018, the null model had the lowest mean absolute error (2.7 days), although the model-averaged forecast performed as well or better than the null model in the majority of retrospective years and currently has a mean absolute error of 3.2 days. The model-averaged forecast had a consistent mean absolute

error regardless of the type of year (*i.e.*, average or extreme early/late) the forecast was made for, which was not true of the null model. Additionally, the availability of the run timing forecast was not found to increase overall accuracy but did reduce uncertainty of in-season run assessments in relation to the null model, particularly early in the season.

2.1   Introduction

In-season management strategies for Pacific salmon *Oncorhynchus* spp. fisheries rely heavily on indices of in-river abundance (*e.g.*, test fisheries, sonar counts, *etc.*) to inform harvest control rules that attempt to attain the balance of meeting pre-determined escapement objectives while allowing adequate opportunity for harvest (Catalano and Jones 2014). However, because indices of abundance are confounded by the phenology (*i.e.*, timing) of the migration, their interpretation is very difficult in-season. For example, smaller-than-average index values early in the season could be due to either a small run with average timing or a late large run, when interpreted in the context of historical years (Adkison and Cunningham 2015). This ultimately leads to great uncertainty about how much of the incoming run has passed, which is a key piece of information that dictates fishery harvest opportunities. There exists no information in the current year's abundance index to inform the manager if (for example) 25% or 75% of the run has passed on any given day. Yet, depending which is true, the optimal management decision could be vastly different. Thus, in-season assessment typically involves some characterization of the variation in historical run timing to formulate a range of possible run size scenarios that could be representative of the current year's run size. However, given the amount of variation in historical run timing, these scenarios are rarely informative during the majority of the migration, when key harvest decisions are being made because the run scenarios may span all possible run sizes. As a result, the pre-season run size forecast may remain the most precise source of information for much of the season. If it were possible to predict the timing of the incoming run (*e.g.*, earlier- or later-than-average) with

some level of confidence, it could prove valuable for in-season assessment and decision-making by reducing uncertainty in run size predictions.

While previous research has uncovered several key physiological mechanisms that are involved with natal homing (Hasler and Scholz 1983) and return migrations of adult salmon to freshwater environments (Cooke *et al.* 2008; Cooperman *et al.* 2010; Hinch *et al.* 2012), the exact physiological and behavioral responses of adult salmon to relatively small-scale environmental gradients within estuaries, which are likely the ultimate determinants of freshwater entry timing, are still poorly understood. Despite this uncertainty, several hypotheses have been put forth that are broadly consistent with the observed timing patterns of several species across a large geographic area (*i.e.*, western and southwestern Alaska). Two primary influences have been suggested: genetic (Anderson and Beer 2009; O'Malley *et al.* 2010; Quinn *et al.* 2000) and environmental (Hodgson *et al.* 2006; Keefer *et al.* 2008) mechanisms. Substantial evidence exists to suggest that both genetic and environmental controls are involved in determining migration timing, however it is broadly thought that genetic variation influences substock variation (*i.e.*, different tributary spawning groups within the same major river basin) and environmental variation influences the timing of the aggregate (*i.e.*, basin-wide) run (Anderson and Beer 2009; Keefer *et al.* 2008). This is consistent with the notion that genetically distinct components of the aggregate run behave differently as a result of their life history strategies and/or the characteristics of their specific spawning grounds (*e.g.*, substocks that must travel farther in-river to reach spawning grounds enter freshwater earlier; Clark *et al.* 2015; substocks that spawn in tributaries influenced by warmer lakes enable later spawning; Burger *et al.* 1985) but that certain environmental conditions act on the aggregate run to either hasten or delay freshwater entry in any given year. It has also been suggested that run size may have an influence on migration timing, although empirical support for this claim seems to be lacking. If there were indeed relationships between run timing and run size, these need to be quantified as certain combinations are particularly

troublesome for managers (*e.g.*, small/early runs and large/late runs appear the same early in-season; Adkison and Cunningham 2015).

At the aggregate population scale, which is the focus of this Chapter, it has been observed that migrations occurring in the spring and summer generally occur earlier in years with warmer spring temperatures (Mundy and Evenson 2011; Hodgson *et al.* 2006). Mundy and Evenson (2011) suggested that this pattern may be explained by the stability of the estuarine water column where adult salmon stage in preparation for riverine entry (or alternatively, marine exit). High estuarine water column stability was hypothesized to impede riverine entry through two mechanisms: (1) by presenting an osmotic barrier between freshwater riverine discharge and the saline ocean water which prevents osmotically incompetent individuals from crossing and, (2) by preventing freshwater-competent individuals from receiving olfactory cues essential to the homeward migration. Thus, Mundy and Evenson (2011) hypothesized that years in which the estuarine water column is stable over a longer period of time would be associated with later migration timing. Although water column stability is a difficult variable to measure over large spatial scales, several variables that are known to influence it are available at large scales *via* remote sensing (*e.g.*, satellite measurements). Such variables are sea ice cover which prevents wind-driven mixing, associated local temperature-related variables like land-based air temperature or sea surface temperature (SST), and broader scale indicators such as the Pacific Decadal Oscillation (PDO), an index of temperature anomalies in the northern Pacific Ocean. Observational studies across the North American range of Chinook salmon *O. tshawytscha* have found environmental-run timing correlations that are consistent with this hypothesis (Hodgson *et al.* 2006; Keefer *et al.* 2008; Mundy and Evenson 2011). Even if the water column stability hypothesis is incorrect, observed patterns suggest that environmental variables may be useful in forecasting run timing with some level of accuracy and certainty; ultimately, understanding of the physical, physiological, or

ecological mechanisms mediating run timing variability is of lesser importance than making good harvest and population management decisions.

Thus, several efforts have been made at exploiting these environmental-run timing relationships to develop run timing forecast models for Pacific salmon migrations. Mundy and Evenson (2011) developed a model for Yukon River Chinook salmon that used air temperature, SST, and sea ice cover to predict the day at which the 15th and 50th percentiles of the run would pass a test fishery index location. Their model fitted the observed data well (errors were nearly always within seven days, usually within three days), although out-of-sample predictive ability was not presented. Keefer *et al.* (2008) developed a similar framework for Columbia River spring run Chinook salmon and found run timing relationships with river discharge, river temperature, and ocean condition indices (*e.g.*, PDO). Their best model explained 49% of the variation in median run timing with variation in the environmental variables. Anderson and Beer (2009) continued this work on the Columbia River spring Chinook stock, but added genetic components to their analysis based on the arrival timing of precocious males. Their findings revealed that both environmental variables and substock composition affected overall run timing of the aggregate spring Chinook salmon run in the Columbia River. These advancements have shown that relationships between migration timing and environmental variables exist and may have utility for use in forecasting applications.

The Kuskokwim River, located in western Alaska, is the second largest river system in the state and supports culturally and economically important Chinook salmon fisheries. Chinook salmon return beginning in late May and continue through early August, with the median date of passage occurring between June 14 and July 2. Fisheries within the region harvest salmon in-river during freshwater migrations using primarily drift gillnet gear. The Kuskokwim River salmon fishery has a distinct cultural importance: nearly all inhabitants are native Alaskans belonging to the Yup'ik group and take salmon annually in the summer months for subsistence purposes (Linderman and Bergstrom 2009). While commercial salmon fisheries

operate within the river, these fishers often also participate in subsistence take and revenues from the sale of commercially harvested salmon often contribute directly to participation in subsistence activities (Wolfe and Spaeder 2009). To ensure long-term sustainable harvest, the Chinook salmon fishery is managed with a drainage-wide escapement goal derived from an age-structured state-space spawner-recruit analysis (Hamazaki *et al.* 2012; Staton *et al.* 2017b). To meet these pre-determined escapement goals, in-season management strategies implement time, gear, and area closures based on limited and imprecise information regarding annual run size. The distant locations of the majority of escapement assessment projects makes direct measurement of escapement performance unavailable until late in the season. Thus, the primary sources of run size assessment information are (1) a pre-season run size forecast range (obtained as the previous year's run size estimate $\pm \sim 20\%$) and (2) an in-river drift gillnet test fishery operated in Bethel, AK which has been implemented using consistent methods since 1984 (described in Bue and Lipka 2016). The interpretation of this test fishery index suffers from the same issue of being confounded by run timing described earlier, making management decisions difficult. Without precise in-season indicators of run size, managers must often choose to either trust a pre-season run size forecast for the majority of the season or somehow place weights on the various run timing hypotheses when interpreting in-season data. Both options could lead to the wrong interpretation of the actual run size, which could have serious consequences for the management of the fishery in a given year (*i.e.*, the unwarranted opening or closing the fishery resulting in severe under- or over-escapement). The necessity of more accurate and precise in-season perceptions of run size is particularly evident in years with anticipated low runs, such as in recent years (*i.e.*, since 2010 for Chinook salmon in the Kuskokwim and many other locations state-wide), as this may allow managers to more effectively guard against over-exploitation while still allowing for limited harvest opportunities to support the cultural and subsistence needs of the region.

This chapter presents an analysis that develops and evaluates the performance of a run timing forecast model for Kuskokwim River Chinook salmon. The objectives were to:

(1) quantify historical run timing,

(2) develop a run timing forecast model using environmental variables selected based on out-of-sample predictive performance,

(3) assess the utility of the forecasting model for improving predictions of end-of-season test fishery indices of run size,

(4) determine if there is a relationship between run size and run timing for the Kuskokwim River Chinook salmon stock.

2.2 Methods

2.2.1 Estimates of migration timing

In this analysis, the forecasted quantity that characterized migration timing was the day at which 50% of the run passed an index location (hereafter, D_{50}). To inform this quantity for each year in the analysis, daily catch per effort (CPE) data were used from the Bethel Test Fishery (BTF) operated by the Alaska Department of Fish and Game (ADF&G), which spans the years 1984 – 2018. The raw data were daily CPE beginning on June 1 and ending August 24 each year. The cumulative sum of these daily CPE values within a year follows a sigmoidal pattern reflecting the shape of the incoming salmon run which is characterized by relatively few early migrants, a peak where the majority of the fish are running, and relatively few late migrants. To estimate the median day of passage as a continuous variable, a logistic model was fitted to the cumulative proportion of daily CPE of the form:

$$p_{d,t} = \frac{1}{1 + e^{-h_t(d - D_{50,t})}}, \tag{2.1}$$

where $p_{d,t}$ is the predicted cumulative proportion on day-of-the-year (DOY) d in calendar year t, h_t is the parameter that controls the steepness of the curve (*i.e.*, duration of the run), and $D_{50,t}$ is the day at which 50% of the total annual CPE was caught in year t. Annual estimates of $D_{50,t}$ and h_t were obtained by fitting $p_{d,t}$ to observed daily cumulative proportion by minimizing the sum of squared deviations from the model prediction. Statistical uncertainty from model fitting for these parameter estimates was not considered further in the analysis as the uncertainty was negligible (coefficients of variation for D_{50} were less than 1%). Further, the use of the BTF daily CPE values to infer the location and shape of year-specific logistic timing curves made the assumption that these data provided an accurate representation of daily run strength within a year (*i.e.*, that the influence of weather conditions or harvest on sampling was negligible).

2.2.2 Environmental variables

Environmental variables assessed for forecasting performance were chosen based on three criteria: (1) previously established association with salmon run timing, (2) availability for the Kuskokwim River during the years for which BTF index observations exist (1984 – 2018), and (3) availability for use in a pre-season forecast model (*i.e.*, available no later than June 10 in the year for which the forecasted value would be used). Based on these criteria, four environmental variables were chosen for analysis: SST, percent sea ice cover (SIC), PDO, and land-based air temperature taken in Bethel, AK.

2.2.2.1 PDO data

Data collected for the PDO variable came from one of several indices produced by the National Oceanic and Atmospheric Administration (NOAA) (Mantua *et al.* 2017)[1]. The index is produced by taking the first principal component of monthly SST anomalies in the

---

[1]PDO data: http://research.jisao.washington.edu/pdo/PDO.latest.txt

northern Pacific Ocean, after removing any global trends due to any systematic change over time (Mantua *et al.* 2017). Thus, for each year of the data set, a single monthly value was available for PDO. Previous studies have found PDO values prior to the initiation of the run have predictive value for Chinook salmon populations (Beer 2007; Keefer *et al.* 2008).

2.2.2.2 Bethel air temperature data

Air temperature data for Bethel, AK were accessed from the Alaska Climate Research Center[2]. These data were available as daily means for each day of each year in the 1984 – 2018 data set. Mundy and Evenson (2011) found relationships with land-based air temperature and Chinook salmon run timing, which is why it was included in this analysis.

2.2.2.3 SST and SIC

SST and SIC data were accessed from the NOAA Optimum Interpolation SST V2 High Resolution Dataset (Reynolds *et al.* 2007)[3]. These data were available as daily means for any 0.25° by 0.25° latitude by longitude grid cell on the globe. To limit the search, only grid cells within Kuskokwim Bay were selected for analysis (Figure 2.1) as that is the area where Chinook salmon bound for the Kuskokwim River likely aggregate prior to riverine entry. The area with grid cells ranged from 58.5° N to 60° N by 164.25° W to 162° W, which resulted in a total of 54 0.25° latitude by 0.25° longitude grid cells. For SST, four grid cells fell partially over land (resulting in 50 grid cells with daily data) and for SIC, five grid cells were partially over land (49 grid cells with daily data). "Empty" grid cells were excluded and the remaining grid cells were used for prediction. Previous analyses have used a simple average over a wide spatial area (*e.g.*, Mundy and Evenson 2011) to obtain a single value for SST or SIC each year. However, this is somewhat arbitrary and does not account for the possibility of certain areas having stronger timing signals than others or that the areas with stronger signals may

---

[2]Alaska air temperature data: http://akclimate.org/acis_data
[3]Global gridded SST and SIC: http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html

change over time. Thus, the gridded spatial structure of these variables was retained and the treatment of this structure in the forecast analysis is discussed below in Section 2.2.5.

2.2.3 Forecast model

To produce a forecast of run timing, relationships between historical observed pairs of the environmental variables each year and $D_{50,t}$ must be quantified. The normal linear regression framework was used to obtain these historical relationships:

$$D_{50,t} = \beta_0 + \beta_j x_{t,j} + \cdots, + \beta_n x_{t,J} + \varepsilon_t,$$
$$\varepsilon_t \overset{\text{iid}}{\sim} \text{N}(0, \sigma^2) \tag{2.2}$$

where $D_{50,t}$ is the observed run timing value in year t, $x_{t,j}$ is the observed value of covariate j (of which there are J included in the model), β_0 and β_j are coefficients linking the observed values of $D_{50,t}$ with $x_{t,j}$, ε_t are random residual effects that explain deviations of observed $D_{50,t}$ from the fitted value and have constant variance equal to σ^2. No time series patterns were visually present in the residuals and ARIMA models suggested no serial autocorrelation, which supports the use of the temporal independence assumption.

There are many such regression models that could be used to produce a run timing forecast (*i.e.*, $\hat{D}_{50,t+1}$). This is because: (1) there are four variables (PDO, air temperature, SST, and SIC) that could be included, (2) each variable is temporally structured, *i.e.*, there are daily or monthly values for each variable, and (3) two variables (SST and SIC) are spatially structured, *i.e.*, there are different values for each day and year for different areas of Kuskokwim Bay (Figure 2.1). Item (1) deals with the specific values of j and J whereas items (2) and (3) deal with what values of x_t for a given variable j should take on.

18

2.2.4 Selection of predictive time periods

To fit the regression model given by (2.2), a single value for each $x_{t,j}$ was required. The covariate data were temporally structured, however, indicating that some selection of which time periods to use to populate $x_{t,j}$ was needed. Oftentimes the average over an arbitrary time period, such as daily values in the month of February, is used based on *a priori* assumptions of the behavior of important factors (van de Pol *et al.* 2016). While this approach is simple to implement and explain, it is possible that a better (*i.e.*, reliably more accurate predictor) time window exists but was not considered. Furthermore, the importance of various time windows may change over time and the arbitrary selection of a single window does not allow for such changes to be detected. To avoid these issues, a rigorous temporal selection process, known as the sliding climate window algorithm (SCWA; van de Pol *et al.* 2016), was implemented to determine the best predictive time period for each variable considered in the forecast model. To find the most reliable temporal window for prediction, the SCWA evaluates all possible windows (subject to certain restrictions) over which to average for use as the predictor variable in the forecast model. Although van de Pol *et al.* (2016) developed, discussed, applied, and simulation-tested the SCWA, no mathematical representation of the algorithm was presented. To provide more clarity on this approach to temporal variable selection, the following section provides the details of the SCWA.

2.2.4.1 The SCWA

A "window" in this context is hereafter defined as a block of consecutive days in some portion of the year with starting day-of-the-year (DOY) denoted by D_F and ending day equal to D_L. The daily values within each evaluated window (including D_F and D_L) were averaged for the $x_{t,j}$ value to be used in a linear regression framework. As input constraints, the SCWA used in this analysis required: (1) the start DOY of the first window to be evaluated (D_0), (2) the end DOY of the last window to be evaluated (D_n), and (3) the minimum window

size of a candidate window ($\Delta_{D,min}$). The algorithm started with the earliest and smallest possible time window: $D_F = D_0 = 1$ through $D_L = D_0 + \Delta_{D,min} - 1 = 5$. The performance of this window when used to obtain $x_{t,j}$ was evaluated for predictive performance (see Section 2.2.4.2 below) and the result was stored for comparison to other candidate windows. For the next window, D_F would remain at D_0, but D_L would be incremented by 1 day ($\ell = 1$). Thus, the endpoints of all candidate windows with $D_F = D_0$ can be generalized as:

$$[D_0, D_0 + \Delta_{D,min} - 1 + \ell], \tag{2.3}$$

for each $\ell = 0, 1, \cdots, n - 1$, where $n = D_L - D_F + 1$. For all windows, including those with $D_F = D_0$, this generalizes to:

$$[D_0 + f, D_0 + f + \Delta_{D,min} - 1 + \ell], \tag{2.4}$$

for each $f = 0, 1, \cdots, n - \Delta_{D,min}$ and $\ell = 0, 1, \cdots, n - \Delta_{D,min} - f$. Windows with $f > n - \Delta_{D,min}$ would contain fewer than $\Delta_{D,min}$ days and are thus prohibited according to the constraints. After evaluating all windows, the single window with the best predictive performance was used to obtain the forecast predictor variable for that data source (*i.e.*, PDO *versus* air temperature).

As an example, consider the following inputs: $D_0 = 1$ (*i.e.*, January 1), $D_n = 31$ (*i.e.*, January 31), and $\Delta_{D,min} = 5$. The SCWA would start with the window encompassing January 1 – January 5, then do January 1 – 6, January 1 – 7, *etc.*, January 1 – 31. Next, it would exclude January 1 from consideration and evaluate all windows starting with January 2, starting with January 2 – January 6. When it completes the one window starting with January 27, it must stop because windows starting later than January 27 would result in windows shorter than five days.

The values of D_0 and D_L for the four covariates are shown in Table 2.2. The setting for $\Delta_{D,min}$ for air temperature, SST, and SIC was five days. Note that because PDO was available in monthly values only, each month was treated as "day" in the algorithm described above and $\Delta_{D,min}$ was set to one.

2.2.4.2 Forecast cross-validation

A metric was needed to objectively measure the performance of the many evaluated windows. A time series forecast cross-validation procedure was used, which is an out-of-sample technique for data that are collected through time (Arlot and Celisse 2010). The procedure operated by producing a forecasted value of D_{50} for year $t+1$ trained based on all data $x_{t,j}$ available from years $1, \cdots, t$. It then continued for all $t = m, \cdots, n-1$, where m is the minimum number of years necessary to fit the model (set at $m = 10$ in all cases) and n is the number of years of available data. Then, absolute forecast error was calculated based on all forecasted years as $|D_{50,t+1} - \hat{D}_{50,t+1}|$, and yearly forecast errors were averaged to obtain mean absolute error ($\overline{\text{AE}}$) which was used as the metric of model performance in window selection. The window with the lowest $\overline{\text{AE}}$ was selected as the optimal window to average over for prediction. The forecasting cross-validation procedure was used as opposed to other out-of-sample validation procedures, such as k-fold or leave-one-out methods, because the data were collected through time and the forecast model would never need to predict (for example) year 2010 from years 1984 – 2009 and 2011 – 2018, but rather it would always need to predict year $t+1$ from all previously collected data.

When forecasting $D_{50,t+1}$ from training data from $1, \cdots, t$, a single optimal climate window was selected for each variable and that window was used to estimate coefficients based on training data and obtain the environmental variable value for prediction in year $t+1$ to forecast $D_{50,t+1}$. When a new year of data was added to the training data (such as in the retrospective forecast analysis; Section 2.2.8), the optimal window for each variable was

re-assessed using the algorithm again. For PDO and Bethel air temperature (which had no spatial structure), the SCWA was used to select the range of monthly (for PDO) or daily (for Bethel air temperature) values to include in the predictive climate window for each year in the analysis. For SST and SIC, the SCWA was used on each grid cell separately. The result was 50 unique grid cell-specific windows for SST and 49 windows for SIC for each year of the analysis. The treatment of this spatial structure in the forecast analysis is discussed below in Section 2.2.5.

2.2.5 Evaluated forecast models

Linear regression (2.2) was used to assess the forecast performance of each of the variables described above, both in isolation of and in combination with other variables. All possible subsets were evaluated (excluding interactive effects) for predictive ability through time, resulting in a total of 16 models ranging from the null (β_0 only; $J = 0$) model to the full model (all $J = 4$ variables included as additive predictors).

For the spatially explicit variables (*i.e.*, SST and SIC), a more complex treatment was required to prevent all grid cell values from being used as predictors in a single model. To handle the spatial structure, grid cell-specific regression models were fitted, then model-averaging (Burnham and Anderson 2002) based on Akaike's Information Criterion (AIC, Akaike 1974) was used to obtain a single forecast D_{50} for each year. Under this approach, each grid cell g received an AIC score:

$$\text{AIC}_{c,g} = n \log \left(\hat{\sigma}_g^2 \right) + 2K + \frac{2K(K+1)}{n - K - 1}, \tag{2.5}$$

where n is the number of data points used in each model, $\hat{\sigma}_g$ is the estimate of the residual standard deviation under grid g, and K is the number of model parameters. The corrected version of AIC (AIC_{c}) is recommended in cases where the ratio of n to K is small (Burnham

and Anderson 2002). Then, each grid cell received a $\Delta\text{AIC}_{\text{c}}$ score, representing its relative performance in comparison to the best grid cell:

$$\Delta_g = \text{AIC}_{\text{c},g} - \text{AIC}_{\text{c},min}, \tag{2.6}$$

where $\text{AIC}_{\text{c},min}$ is the minimum AIC_{c} across all grid cells. Model (grid cell) weights were then calculated as:

$$w_g = \frac{e^{-0.5\Delta_g}}{\sum_j^G e^{-0.5\Delta_j}}, \tag{2.7}$$

where G is the number of grid cells. Grid cell-averaged predictions were then obtained as:

$$\hat{y}_{t+1} = \sum_g^G w_g \hat{y}_{g,t+1}, \tag{2.8}$$

where $\hat{y}_{g,t+1}$ is the forecasted value of D_{50} for grid cell g.

2.2.6 Forecast uncertainty

In addition to forecast accuracy, forecast precision is also of great importance as it is an index of the confidence with which the forecast should be interpreted. For models that did not require AIC_{c} model-averaging across grid cells, the following estimator was used to produce a forecast standard error (SE):

$$\text{SE} = \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i^n (x_i - \bar{x})^2}}, \tag{2.9}$$

where n is the number of years the model was fitted to, x is the value of the predictor variable used for forecasting, and $\bar{x}$ is the mean of all predictor values excluding the new value used for forecasting. For models that used AIC_{c} model-averaging (*i.e.*, those including SST and SIC), the following estimator was used to produce prediction SE:

$$\text{SE} = \sum_g^G w_g \sqrt{\text{SE}_g^2 + (\hat{y}_{g,t+1} - \hat{y}_{t+1})^2}, \tag{2.10}$$

where SE_g is the prediction SE from grid cell g calculated using (2.9). This estimator of unconditional sampling standard error accounts for uncertainty within each model and the uncertainty due to model selection (Burnham and Anderson 2002). Prediction intervals were calculated using the point estimate of prediction, the prediction SE, and appropriate quantiles from the corresponding t-distribution.

2.2.7 Forecast model selection

Given 16 forecast models, it is impossible to know which may produce the best forecast for the current year. Thus, three methods to obtain a forecast for D_{50} were evaluated: (1) the null (*i.e.*, intercept only) model, (2) the single model with the lowest forecast cross-validation score as of the last year, and (3) model-averaging across the ensemble of 16 forecast models based on AIC_c scores. According to Burnham and Anderson (2002), model-averaging should perform better than a single "best model" at prediction when there is a high degree of uncertainty about which model is best. This procedure was performed using (2.5) – (2.10), by substituting the prediction, prediction SE, and K for forecast model i, in place of grid g. Prediction intervals based on model-averaged predictions and prediction SE present somewhat of a problem when the different models contributing to the average contain differing degrees of freedom as it is unclear how many standard errors the prediction limits should lie from the mean prediction. Thus, the estimator suggested by Burnham and Anderson (2002) of the "adjusted SE" (ASE) was used:

$$ASE = \sum_i^{16} w_i \sqrt{\left(\frac{t_{df_i,1-\alpha/2}}{z_{1-\alpha/2}}\right)^2 \text{SE}_i^2 + (\hat{y}_{i,t+1} - \hat{y}_{t+1})^2}, \tag{2.11}$$

24

where $t_{df,i,1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the t-distribution with degrees of freedom equal to that of model i and $z_{1-\alpha/2}$ is the corresponding quantile of the z-distribution (*i.e.*, standard normal). The confidence level $\alpha = 0.05$ was used in all cases.

2.2.8 Retrospective forecast analysis

The analysis was conducted in a retrospective forecast framework starting in 1994, which allowed the models to be trained to 10 years of observed data before being used for forecasting. All data after 1994 were ignored, optimal windows were selected for each of the four variables (and all grids for SST and SIC), all 16 models were fitted, a D_{50} forecast was made for 1995 using the three approaches described in Section 2.2.7, and each was evaluated for predictive accuracy. This process was repeated annually until the present (*i.e.*, out-of-sample predictions made for $1995 - 2018$), which allowed for the calculation of $\overline{\text{AE}}$ through time as if the forecast model would have been available beginning in spring 1995. In addition to $\overline{\text{AE}}$, median absolute error ($\widetilde{\text{AE}}$) was calculated to validate prediction accuracy of estimates by ignoring the effect of outlying poor predictions.

2.2.9 Value of forecast to run size assessments

It is important to remember that the purpose of producing a run timing forecast is to aid in the interpretation of in-season indices of run size such as test fisheries. To evaluate the utility of having access to the run timing forecast model, the accuracy and precision of an imperfect abundance index for the Kuskokwim River were compared when informed using D_{50} forecasts from the model-averaged and the null forecast models. The abundance index is denoted by EOS_t, and is the end-of-season cumulative CPE observed in the BTF in year t. Under the assumption of constant catchability, EOS_t should be proportional to total abundance, with deviations introduced by sampling noise. In-season predictions of EOS_t were made for each year t, model i, and day d in the season with:

$$\widehat{\text{EOS}}_{d,t,i} = \frac{\text{CCPE}_{d,t}}{\hat{p}_{d,t,i}}, \tag{2.12}$$

where $\text{CCPE}_{d,t}$ is the cumulative CPE caught at the BTF through day d in forecasting year t $\left(\text{CCPE}_d = \sum_{j=1}^{d} \text{CPE}_j\right)$, $\hat{p}_{d,t,i}$ is the predicted cumulative proportion of the run that had passed the BTF location on day d in year t from model i (*i.e.*, model-averaged *versus* null forecast model) obtained by inserting the forecasted value of D_{50} into the logistic function (2.1). Uncertainty in the run timing forecast model was propagated to $\text{EOS}_{d,t,i}$ using a parametric Monte Carlo procedure (Bolker 2008). Independent normal random samples for $D_{50,t}$ and h_t were sampled to obtain $p_{d,t,i}$ for use in (2.12). The distribution for generating samples of D_{50} differed based on the mean and SE of the forecasted value of D_{50} according to either the null or the model-averaged model. Estimates of $D_{50,t}$ and h_t seem to be independent for the Kuskokwim River Chinook stock, so the method of drawing samples of h_t was identical for the comparison and involved sampling from a normal distribution with mean and standard deviation equal to the estimated quantities from all years before year t to ensure the consistency of out-of-sample predictions. A sufficiently large number of Monte Carlo samples were drawn (10,000) for each evaluated day and year. Prediction uncertainty was quantified using the coefficient of variation (calculated as median/sample standard deviation across all Monte Carlo samples for year t and day d). Accuracy was assessed using mean absolute percent error (MAPE) and mean percent error (MPE), where the point estimate used was the median of all Monte Carlo samples of $\text{EOS}_{d,t,i}$. Prediction performance measures were compared between the null and the model-averaged forecast model on June 15, June 30, July 15, and July 30 each year a forecast was available (1995 – 2018), which represent the historical average 21%, 79%, 95%, and 99% points of the Chinook salmon run in the Kuskokwim River. Using the null model to obtain $\hat{p}_{d,t}$ is one of several methods that could be done to produce the predictions $\text{EOS}_{d,t}$ in the absence of an environmental forecast variable

model for D_{50}. This method was used (as opposed to other methods, like simply taking the average p_d across all years to populate (2.12) so the comparison involved the same assumption of symmetry in the sigmoidal timing curve, which has the ability to affect the accuracy of $EOS_{d,t,i}$ predictions.

2.2.10   Investigation of a run timing *versus* run size relationship

To test the hypothesis that run timing is related to run size (*e.g.*, small runs are typically early, or *vice versa*), two models were investigated for their predictive performance using the forecast cross-validation criteria: the null model and a model that included run size as a predictive covariate in place of the environmental variables. Run size was obtained from a maximum likelihood run reconstruction model that compiles all assessment information (*i.e.*, 20 escapement count indices, harvest estimates, drainage-wide mark-recapture estimates, *etc.*) to estimate the run size that makes the collected data most likely to have been observed (Bue *et al.* 2012; Liller *et al.* 2018). The forecast absolute errors in each year were then compared using a two-tailed paired t-test with $\alpha = 0.05$.

2.3   Results

2.3.1   Estimates of run timing

There was a considerable amount of inter-annual variability in D_{50}, with a range of 17 days and sample standard deviation of 3.62 days over the 35 years with run timing data from the BTF (Figure 2.2; Table 2.1). Based on D_{50} alone, the earliest run on record was in 1996 when D_{50} occurred on DOY 166 (June 14) and the latest run was in 1985 when D_{50} was attained on DOY 183 (July 2). The average D_{50} was 173.75 across all years available (which rounded down is June 22 in a normal year and June 21 in a leap year).

The logistic curve fitted the daily cumulative CPE proportions well in all years of the BTF data set (Table 2.1), as indicated by an average residual standard error estimate of 0.022, with a maximum estimate of 0.038 in 1992. The majority (95%) of all residuals from all years fell between -0.056 and 0.044. Parameter estimates were quite precise, with D_{50} having a smaller average coefficient of variation than h, (0.07% and 2.07%, respectively). Given this negligible degree of parameter uncertainty, it was ignored throughout the rest of the analysis.

2.3.2 Variable-specific relationships

Looking at each of the environmental variables in isolation of all others, it is clear that there is a distinct relationship between temperature-related environmental variables and Kuskokwim River Chinook salmon migration timing (Figure 2.3). For illustration purposes, the figures for the two gridded variables (SST and SIC) were produced by taking an average across all grid cells weighted by the AIC_c weight for each grid cell. Air temperature, PDO, and SST all had negative relationships with D_{50}, whereas SIC had a positive relationship (Table 2.3). All single variable relationships were significant at the 0.01 level, but the estimated residual standard error was approximately 3 days for each model (Table 2.3). R^2 values were generally low (range: 0.19 for air temperature and 0.33 for SIC; Table 2.3).

2.3.3 Selected climate windows

It was difficult to generalize on the climate windows selected for each variable based on forecast cross-validation performance, because the selected windows changed with each new year of data and SST and SIC had windows for each grid cell; however, some noteworthy patterns arose. First, the best window for PDO was consistently the value for the month of May for each year the forecasts were produced (not shown). Second, selected windows for air temperature fluctuated from year to year to some extent: all short and mid-May windows

were selected, then shifted to long windows spanning February to late-May beginning in the early 2000s (Figure 2.4*a*). Third, selected windows through time were substantially more variable for most grid cells for SST and SIC than air temperature, although many grid cells remained relatively constant or became more "focused" as more years of data were added (Figures 2.4*b,c*). In general, chosen windows for SST began in early- to mid-May and ended in late-May (Figure 2.4*b*) whereas windows starting in early-April and ending in mid- to late-April were predominately chosen for SIC (Figure 2.4*c*). The selected climate windows in southern-most grid cells appeared more stable for SST (Figure 2.4 panels *b*3 and *b*4), whereas climate windows in northern grid cells appeared more stable for SIC (Figure 2.4 panels *c*1 and *c*2; stable in the sense that the optimal windows changed less as new years were added to the training data). Finally, the best predictive windows tended to become systematically earlier in the season as the retrospective analysis progressed.

2.3.4 Forecast performance

Of the three investigated forecast methods (null model, model with lowest forecast cross-validation error up to the forecasting year, and AIC_c model-averaging), the null model had the lowest $\overline{AE}$ from 1995 to 2018 (2.7 days; Figure 2.5). AIC_c model-averaging performed the same as using the single model with the lowest cross-validation score (both had $\overline{AE} = 3.2$ days; Figure 2.5). However, these patterns were not consistent across the entire time series. For the period of 1996 to 2008, the model-averaged forecast had a lower $\overline{AE}$ than the null model, and for the period of 2009 to 2015, the model-averaged forecast had approximately the same or lower $\overline{AE}$ scores (Figure 2.6). It was due in a large part to an extreme error in 2016 (Figure 2.5) that the model-averaged forecast had a higher $\overline{AE}$ than the null model. A similar case occurred in 2015, but not 2014 when the run was truly very early (Figure 2.5). Expressing prediction error in terms of median absolute error ($\widetilde{AE}$) resulted in lower average errors (null = 2.1, single best = 3.1, and model-averaged = 2.4), indicating that

extreme prediction errors (*i.e.*, outliers) influenced the value of $\overline{\text{AE}}$ for the model-averaged forecast and the null model the most – $\overline{\text{AE}}$ and $\widetilde{\text{AE}}$ were nearly identical for the approach that used the single model with the lowest CV score at the time (Figure 2.5). Additionally, by comparing the width of the prediction intervals in Figure 2.5 across forecasting approaches, it was clear that model-averaging substantially reduced prediction uncertainty (SE) in relation to the null and single best model approaches.

To compare performance in average *versus* extreme years among forecasting approaches, $\overline{\text{AE}}$ was further calculated in a more specific way: based on how similar or dissimilar the included years were to the mean observed run timing across all years. As would be expected, the null model performed well when years with D_{50} within only ± 1 days of the average were in included in the calculation of $\overline{\text{AE}}$ (Figure 2.7*a*), but its accuracy became increasingly worse as years with more extreme realized D_{50} values were included in the calculation (increasing x-axis values in Figure 2.7*a*). The two environmental variable forecast approaches (model-averaging or the single "best" model in each year) performed nearly the same across this continuum and neither $\overline{\text{AE}}$ score was sensitive to the overall similarity or dissimilarity the included years had with average run timing (Figure 2.7*a*). The lower panel shows the relative frequency with which these various scenarios occurred, indicating how much information each scenario contributed to the overall $\overline{\text{AE}}$. On the other hand, the null model only performed as well as the model-averaged forecast and single best model when years with $D_{50} \pm 0.5$ days outside of the mean were considered (Figure 2.7*b*). As only more extreme years were considered in $\overline{\text{AE}}$ (increasing x-axis values in Figure 2.7*b*), the null model rapidly performed worse and the model-averaged forecast remained relatively insensitive to the degree of extremity of the D_{50} value it was forecasting relative to the mean (Figure 2.7*b*).

2.3.5 Value to in-season run size assessments

When the model-averaged and null model forecasts for $D_{50,t}$ were used to retrospectively assess the potential for the run timing forecast model to improve in-season run assessment based on daily cumulative BTF CPE, it was evident that the range of possible $\widehat{\mathrm{EOS}}_{d,t}$ was substantially smaller when the model-averaged forecast was used as opposed to the null forecast. This is evident by the average daily coefficients of variation in the predictions made for $\widehat{\mathrm{EOS}}_{d,t}$ on the two evaluated dates in June: 71% *versus* 133% on June 15 and 14% *versus* 22% on June 30 (Table 2.4). The reduction in uncertainty of $\widehat{\mathrm{EOS}}_{d,t}$ predictions in the first evaluated day is of importance as it is the time where many harvest management decisions are being made. Conversely, negligible improvements in the accuracy of $\widehat{\mathrm{EOS}}_{d,t}$ predictions were found for the model-averaged forecast model in comparison to the null model. On June 15, MAPE when using $p_{d,t}$ informed by the model-averaged $D_{50,t}$ forecast was 43% as opposed to 47% for the null model, indicating that the model-averaged forecast model was not effective at reducing the average magnitude in $\widehat{\mathrm{EOS}}_{d,t}$ prediction errors. Additionally, the model-averaged forecast was ineffective at reducing the positive bias in these predictions from the null model as indicated by the MPE for both models being approximately 10% (Table 2.4).

A visual example of prediction accuracy and uncertainty from two recent years is provided in Figure 2.8. The upper panels show the time series of $\widehat{\mathrm{EOS}}_d$ when the null and model-averaged forecast models were used to inform the location of the logistic cumulative timing curve in 2013. The horizontal line shows the observed value of EOS_t. 2013 is an example of when the null model would have been preferable to use (in terms of accuracy) and 2014 shows a case when the model-averaged forecast would have performed better. According to Figure 2.5, 2014 was one of the earliest runs on record, which explains using D_{50} informed by the null model lead to over estimates of $\widehat{\mathrm{EOS}}_d$ for much of the season that year.

2.3.6 Run timing *versus* run size relationship

There appeared to be little evidence to lend support for the hypothesis that run timing and run size are related for the Kuskokwim River Chinook salmon stock. Based on the visual depiction of the relationship (Figure 2.9), there appeared to be a weak negative pattern. However, the fitted regression model suggested the effect of run size on run timing was weak: on average, D_{50} occurred 1.15 (95% confidence limits; -2.58 – 0.27) days earlier for each 100,000 fish increase in run size, which was not significantly different than no effect of run size on run timing ($p = 0.11, R^2 = 0.08, \hat{\sigma} = 3.6$). Additionally, based on forecast cross-validation, the model that included run size did not perform better at prediction than the null model. On average, the model that included run size model resulted in an estimated absolute forecast error of 0.3 (95% confidence limits; 0.47 – 1.07) days larger than that of the null model ($p = 0.42$).

2.4 Discussion

The environmental relationships with run timing I detected for the Kuskokwim River Chinook salmon stock are consistent with patterns found elsewhere in the region (*e.g.*, Mundy and Evenson 2011; Hodgson *et al.* 2006). Specifically, I found that warmer years were typically associated with earlier-than-average runs as were years with lower-than-average SIC. These findings are consistent with the water column stability hypothesis suggested by Mundy and Evenson (2011). The amount of unexplained variation in the Kuskokwim model appears to be comparable with the Yukon River Chinook salmon stock as well (Mundy and Evenson 2011). Using the Kuskokwim River relationships shown in Figure 2.3, the correlation with D_{50} was -0.52, -0.57, and 0.59 for air temperature, SST, and SIC, respectively. For the Yukon River Chinook salmon stock, Mundy and Evenson (2011) found correlations of -0.59, -0.72, and 0.66 for the same variables but measured at different spatial and temporal scales and with

approximately 10 more years of data included (Table 2 in Mundy and Evenson 2011). These similar correlations indicate the signals given by environmental variables are of relatively equal strength between these two systems.

Given the presence of the environmental relationships, it is somewhat surprising that the null model forecast performed better on average than did the model-averaged forecast. This could, potentially, be due to the fact that a variety of biological (size; *e.g.*, Bromaghin 2005, and morphology; Hamon *et al.* 2000) and abiotic factors (temperature; *e.g.*, Salinger and Anderson 2006, river discharge; *e.g.*, Keefer *et al.* 2004, and migration distance; *e.g.*, Eiler *et al.* 2015) may affect migration rate (and subsequently, encounter probability) and catchability, introducing additional variability in my run timing estimates. Future research that accounts for these effects on encounter probability or catchability could offer improved predictions of run timing. Regardless of the underlying drivers, the overall prevalence of years with average run timing likely led to the enhanced performance of the null model.

Although the null model performed better in the long-term average (*i.e.*, lower $\overline{\text{AE}}$ as of 2018), there are reasons a manager may still justifiably prefer the model-averaged forecast. First, the difference in $\overline{\text{AE}}$ between the model-averaged forecast and the null model was 0.5 days, which is small relative to the amount of annual variation in run timing (a 17 day range for D_{50} over 35 years). Second, the model-averaged forecast performed equally well in terms of forecast accuracy regardless of the type of run timing it was used to forecast (*i.e.*, prediction error equal in extreme early/late and average years; Figure 2.7b). In contrast, the null model only performed comparably well in years with run timing within ±3 days from average and error increased precipitously in more extreme years. Third, the 95% prediction intervals from the null model seemed too wide as 100% of the observations fell within the intervals, whereas 92% of the observations fell within the prediction intervals from the model-averaged forecast (which is closer to the ideal coverage, *i.e.*, 95%). Prediction uncertainty was lower under the

model-averaged forecast than the null model, which could ultimately lead to fewer run timing scenarios being considered early in the season.

As for the value of having access to a run timing forecast to in-season run size assessments, I showed that the model-averaged forecast offered no greater performance in terms of accuracy than the null model. However, the key difference between approaches was the reduced uncertainty in EOS_t predictions when using the model-averaged forecast due to the exclusion of extreme early or late runs which lead to extreme low and high EOS_t predictions early in the season. The null model was forced to always consider these scenarios, resulting in greater uncertainty (coefficient of variation) in EOS_t predictions, particularly between June 15 and June 30 when key management decisions are made. Due to the large amount of uncertainty under the null model (which is essentially the currently used method), EOS_t predictions go largely ignored for much of the season and the pre-season run size forecast is trusted instead. If the environmental variable forecast model were to be used, it likely would provide managers with more information when making decisions. It should be emphasized that I did not evaluate the ability to assess actual run size herein, only an index of run size (EOS_t). Though because EOS_t can be expressed as a function of actual run size (using assumptions about catchability), more precise predictions of EOS_t should presumably result in more precise predictions of actual run size.

The sliding window algorithm was an objective, adaptive, and data-driven tool for temporal selection of environmental predictors and therefore may be more appropriate than choosing a time window based on *a priori* assumptions, particularly in forecasting applications. The algorithm relied on the data and predictive performance to select the window used for the next year's forecast. This framework allowed for predictor variables to change adaptively as a more accurate window became apparent. This quality of the sliding window algorithm makes it intuitive and potentially preferable in the face of a changing climate. However, the algorithm was computationally intensive for the retrospective analysis. The R code to select

windows for all variables/grids for the procedure took approximately 1.5 days to complete on a desktop computer with a 3.60 GHz processor with four cores and 32 GB of RAM. Each year took approximately 3.5 hours to complete (depending on the number of years the forecast cross-validation procedure was conducted on). The flexibility of the approach also hinders the ability to typify a year as "warm" or "cold", as these criteria may change when an additional year of data is included and a new window is selected. An additional drawback of the sliding window algorithm is that it may be difficult to explain to managers and stakeholders, which may lead to confusion and distrust in the method.

Model-averaging across grid cells for SST and SIC was also an objective, adaptive, and data-driven, (though computationally intensive) solution for dealing with the spatial nature of predicting run timing from these two variables. Of course, it would be possible to average all daily values across grid cells each year and perform the sliding window algorithm on these means. However, this would ignore the fact that some grid cells inherently have a stronger timing signal and would likely insert more variation into the predictive relationship. Additionally, the model had the flexibility to place more weight on different grid cells as more years of data were added, again adding to the flexibility of our overall approach which may be preferable in the face of a changing climate. However, the inherent complexity of including the spatial structure again makes it more difficult to typify a year as "warm" or "cold" as there are many values each year for SIC and SST and the strength of the run timing signal given by each grid cell varies.

These two complexities to my analysis (sliding window selection and model-averaging across spatial grid cells) made the interpretation of effect sizes and selected windows difficult in a biologically meaningful way because the best windows and spatial grid cells could change from year to year. I generally see this flexibility based on predictive performance as more important for this particular analysis than biological inference on research topics like determining the most influential variable on run timing variation or determining the

migration route of Chinook salmon through Kuskokwim Bay. These examples remain exciting research questions for the future, however I see them as secondary objectives of importance. My focus was on the prediction of a single critical quantity, D_{50}, which could aid in-season decision-making. A separate issue confounding biological interpretation is that each variable had some direct or indirect link to temperature, suggesting that there is strong potential for multicollinearity among predictor variables. It is well known that correlated predictor variables can result in biased coefficient estimates and variance inflation (Neter *et al.* 1996). This was one reason coefficient estimates were only presented in the single-predictor case (Table 2.3), as I caution against their interpretation in this particular case. However, my focus was entirely on predictive ability, which is generally known to be unaffected by multicollinearity (Graham 2003).

An important caveat of this analysis is that I assumed a negligible influence of downstream harvest on the ability of the BTF to index the true Chinook salmon run timing. This assumption was likely violated in some years but the magnitude of the impact is unknown. From 2008 to 2017, the approximate average exploitation rate by only villages downstream of the BTF index was 20% (*versus* 38% for villages across the whole drainage), thus there is the potential for a downstream harvest bias on perceived run timing (*i.e.*, approximately a fifth of the run can be removed before it is sampled by the BTF). Moderate to high exploitation rates would not necessarily bias the BTF index if the timing of the harvest was similar to the run. However, in the Kuskokwim, subsistence harvest has historically focused on the early portion of the run (Hamazaki 2008). When coupled with moderate exploitation rates (*i.e.*, 20%), this early nature of the fishery is likely to have resulted in detected timing curves that were biased late (due to early fish being removed before they are sampled by the BTF) to some unknown degree. Historical and future interpretation of the BTF is further complicated by the operation of the fishery, such as a recent regulatory measure which mandates that no directed Chinook fishery may begin on or before 11 June. I suspect that

the magnitude of the bias in the index due to the timing of downstream harvest would be small and would not likely affect the general conclusions of this analysis (although residual variation in environmental-run timing relationships would likely be lower if accounted for, which could improve forecast performance). I suggest that future studies should attempt to develop methods that remove harvest effects from the BTF index and other similar indices and assess the magnitude of potential bias. Even if harvest bias could be removed from the historical index values, addressing bias in the test fishery index would be unfeasible during the season because spatially and temporally explicit harvest data are often unavailable until the season has concluded, and the data regarding the temporal distribution are fragmentary.

It was unsurprising that no meaningful relationship exists between estimated run size and run timing. Given that small/early and large/late runs are problematic for in-season management (Adkison and Cunningham 2015), I see the lack of a relationship as beneficial to the management effort. In other words, a small run is no more likely to be early than it is to be late, and the same is true of large runs. For managers, this means that although these small/early or large/late scenarios have occurred in the past, they need not be particularly worried about them due to an overwhelming prevalence over other run size/run timing scenarios.

There is evidence to suggest that, on a population demographic scale, substock structure and relative stock composition may influence the run timing of the aggregate. For example, Clark *et al.* (2015) showed that Chinook salmon that travel farther in the drainage to spawn (*i.e.*, headwaters) enter the main-stem earlier in the season. This point is supported in the Kuskokwim River based on ADF&G radio telemetry data (Smith and Liller 2017a,b; Stuby 2007), which showed that the date at which 50% of headwaters fish were tagged occurred as many as 10 or 11 days earlier than tagging of fish bound for middle river and lower river tributaries. Thus, it is more appropriate to view the timing curve detected by the BTF index as a mixture distribution made up of several distinct substocks, each entering at different

37

times. The cumulative effect of this is one curve that looks logistic likely because the various substocks overlap to a large extent. However, it is not difficult to see that if in some years the headwaters substocks made up a greater proportion of the aggregate stock than the lower and middle river substocks, the timing curve of the aggregate would be earlier than if other stocks had a greater contribution. Using genetic techniques, Anderson and Beer (2009) found that variations in the relative abundances of the populations composing the spring Chinook salmon run in the Columbia River, USA, explained 62% of the variation in annual run timing. This is a source of variation that was not accounted for in this analysis for at least two reasons. First, the resolution to divide the aggregate curve into its substock components is not currently available: Kuskokwim River Chinook telemetry studies were conducted from $2003 - 2007$ and $2015 - 2017$ and the aggregate timing curve does not deviate enough from the smooth logistic curve to separate the different substock components. Second, information on the relative contribution both in the past and in the forecast year would be necessary to include this complexity in the run timing forecast. This detailed level of substock information is not available for the Kuskokwim River. The telemetry data can shed some light on these issues, but they are confounded by factors like harvest timing (some stocks may be harvested preferentially purely due to the timing of the fishery, which does not mirror that of the aggregate run; Hamazaki 2008) or the potential of tagging stock components in some proportion other than their true contribution.

Methods exist to incorporate run timing forecasts from analyses like these into in-season assessment and management efforts. My predictions of $\widehat{\text{EOS}}_{t,d}$, which is an index of run size, could be used to predict the total end-of-season run size on each day using a regression model that relates historical reconstructed total run abundance and observed EOS_t. These in-season run predictions could be used to update pre-season run size forecasts with in-season data using methods such as inverse variance weighting (*e.g.*, Walters and Buckingham 1975) or Bayesian inference (*e.g.*, Fried and Hilborn 1988). Information updating may be preferable in cases

when the pre-season run forecast is biased, because it would allow for the perception of run size to pull away from the forecast when in-season data suggest it is highly unlikely. As I have shown here, uncertainty in $\widehat{\text{EOS}}_{t,d}$ predictions is a function of the precision in the anticipated proportion of the run completed-to-date (2.12). My analysis suggests that incorporating run timing forecasts into estimates of $p_{d,t}$ (and thus EOS_t) may provide managers with more certainty regarding interpretation of in-season abundance indices, which would facilitate updating of pre-season forecasts with data from the run.

Since completion of the run timing forecast model, I have had the opportunity to more thoroughly test the performance of using it to estimate run size in-season in a Bayesian framework (Staton and Catalano 2019). We found that the inclusion of information from the timing forecast model had no influence on the predicted run size, and thus that method had the same performance as a method that simply expected years with average run timing. This finding was a result of the large amount of sampling variability in the BTF: even when the EOS_t value is completely observed in year t, posterior abundance predictions are still highly uncertain (coefficient of variation of 20%) as a result of a scattered historical relationship. The main conclusion here is that although the migration timing forecast had utility for reducing uncertainty of predictions for $\widehat{\text{EOS}}_t$, EOS_t appears to be a generally poor predictor of run abundance. This indicates that if a more reliable index (such as a sonar project) with a tighter relationship with the total run size were to become available, the run timing forecast model developed here may have greater utility.

TABLE 2.1: Parameter estimates (mean with standard error in parentheses) from logistic curves given by (2.1) fitted to each year separately. $D_{50,t}$ is expressed as the day-of-the-year, for reference, day 174 is June 22 in a leap year and June 23 in a normal year.

Year	$D_{50,t}$	h_t
1984	174.8 (0.2)	0.155 (0.004)
1985	183.5 (0.1)	0.254 (0.006)
1986	173.2 (0.1)	0.207 (0.006)
1987	172.7 (0.1)	0.17 (0.003)
1988	172.1 (0.2)	0.164 (0.004)
1989	173.7 (0.1)	0.203 (0.004)
1990	175.9 (0.1)	0.169 (0.003)
1991	175.8 (0.1)	0.187 (0.003)
1992	173.3 (0.2)	0.153 (0.005)
1993	168.3 (0.1)	0.218 (0.003)
1994	169.8 (0.1)	0.186 (0.004)
1995	172.4 (0.1)	0.193 (0.002)
1996	166.3 (0.1)	0.212 (0.004)
1997	170.6 (0.1)	0.261 (0.008)
1998	175.1 (0.1)	0.199 (0.003)
1999	180.9 (0.2)	0.127 (0.003)
2000	171.5 (0.2)	0.166 (0.005)
2001	174.1 (0.1)	0.192 (0.004)
2002	170 (0.2)	0.174 (0.006)
2003	168.9 (0.2)	0.166 (0.004)
2004	174 (0.2)	0.173 (0.004)
2005	173.6 (0.1)	0.164 (0.004)
2006	175.4 (0.1)	0.194 (0.004)
2007	177.8 (0.1)	0.185 (0.003)
2008	176.1 (0.1)	0.191 (0.003)
2009	173.1 (0.1)	0.228 (0.003)
2010	173.1 (0.2)	0.186 (0.006)
2011	173.9 (0.1)	0.158 (0.002)
2012	178.5 (0.1)	0.217 (0.005)
2013	173.5 (0.1)	0.217 (0.005)
2014	166.6 (0.1)	0.166 (0.003)
2015	174.6 (0.2)	0.117 (0.002)
2016	174.1 (0.1)	0.125 (0.001)
2017	178.5 (0.1)	0.159 (0.003)
2018	175.5 (0.1)	0.167 (0.002)

TABLE 2.2: The input constraints used in the SCWA (Section 2.2.4.1) for selecting the best time period of prediction for each covariate. Note that only monthly variables were available for PDO.

	D_0			D_n		
Variable	DOY	Non-Leap Year	Leap Year	DOY	Non-Leap Year	Leap Year
AIR	1	Jan. 1	Jan. 1	151	May 31	May 30
PDO	–	Jan.	–	–	Jun.	–
SST	92	Apr. 2	Apr. 1	151	May 31	May 30
SIC	50	Feb. 19	Feb. 19	130	May 10	May 9

TABLE 2.3: Estimates and statistics of the effects of each of the four single-variable forecast models fitted with all D_{50} and environmental data through 2018.

Variable	$\hat{\beta}_0$	$\hat{\beta}_1$	t	R^2	$\hat{\sigma}$	F
AIR	171.37	-0.23	-2.88	0.19	3.26	8.83[*]
PDO	174.95	-1.87	-3.53	0.25	3.13	12.57[*]
SST	179.01	-1.53	-4.03	0.31	3.02	15.92[**]
SIC	170.22	11.67	4.23	0.33	2.96	17.81[**]

Significance Codes
[*] P < 0.01
[**] P < 0.001

TABLE 2.4: Retrospective accuracy and uncertainty of predictions of the end-of-season cumulative CPE at the Bethel Test Fishery ($\widehat{\mathrm{EOS}}_{d,t,i}$) as informed using two methods of obtaining estimates of the fraction of the run complete ($p_{d,t}$) as described in the text (Section 2.2.9). Accuracy is expressed as the mean percent absolute percent error (MAPE) and mean percent error (MPE) and uncertainty in the prediction made each day and year is expressed using the coefficient of variation (CV).

	MAPE		MPE		CV	
Date	NULL	Mod. Avg.	NULL	Mod. Avg.	NULL	Mod. Avg.
6/15	47%	43%	8%	7%	133%	71%
6/30	10%	13%	-4%	-1%	22%	14%
7/15	3%	3%	-3%	-3%	2%	2%
7/30	1%	1%	-1%	-1%	0%	0%
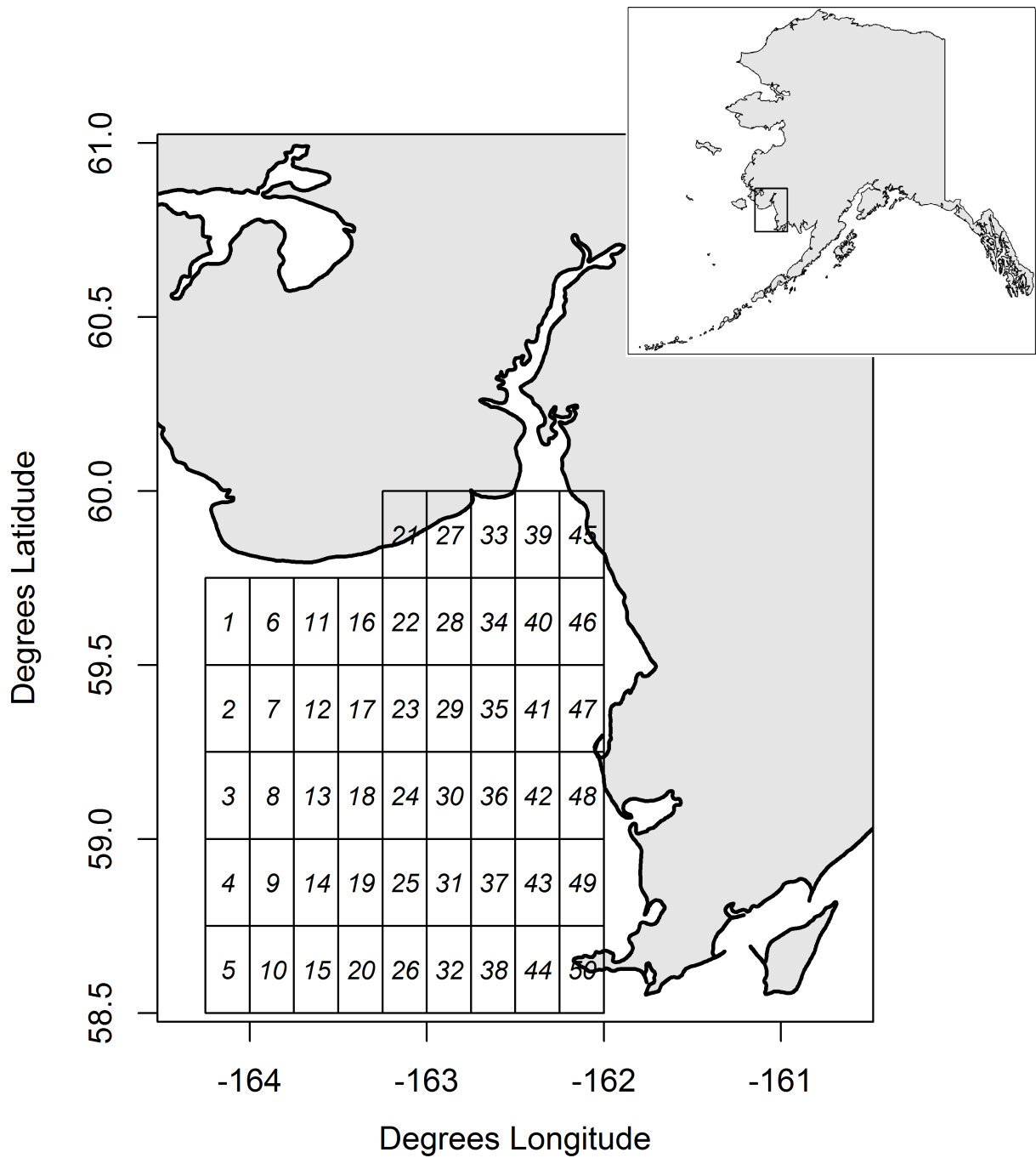
FIGURE 2.1: Map of Kuskokwim Bay where Chinook salmon likely stage for transition to freshwater. Shows grid cells from which daily SST values were used. Daily SIC values came from the same grid cells, though excluding grid cell 45 due to missing values.
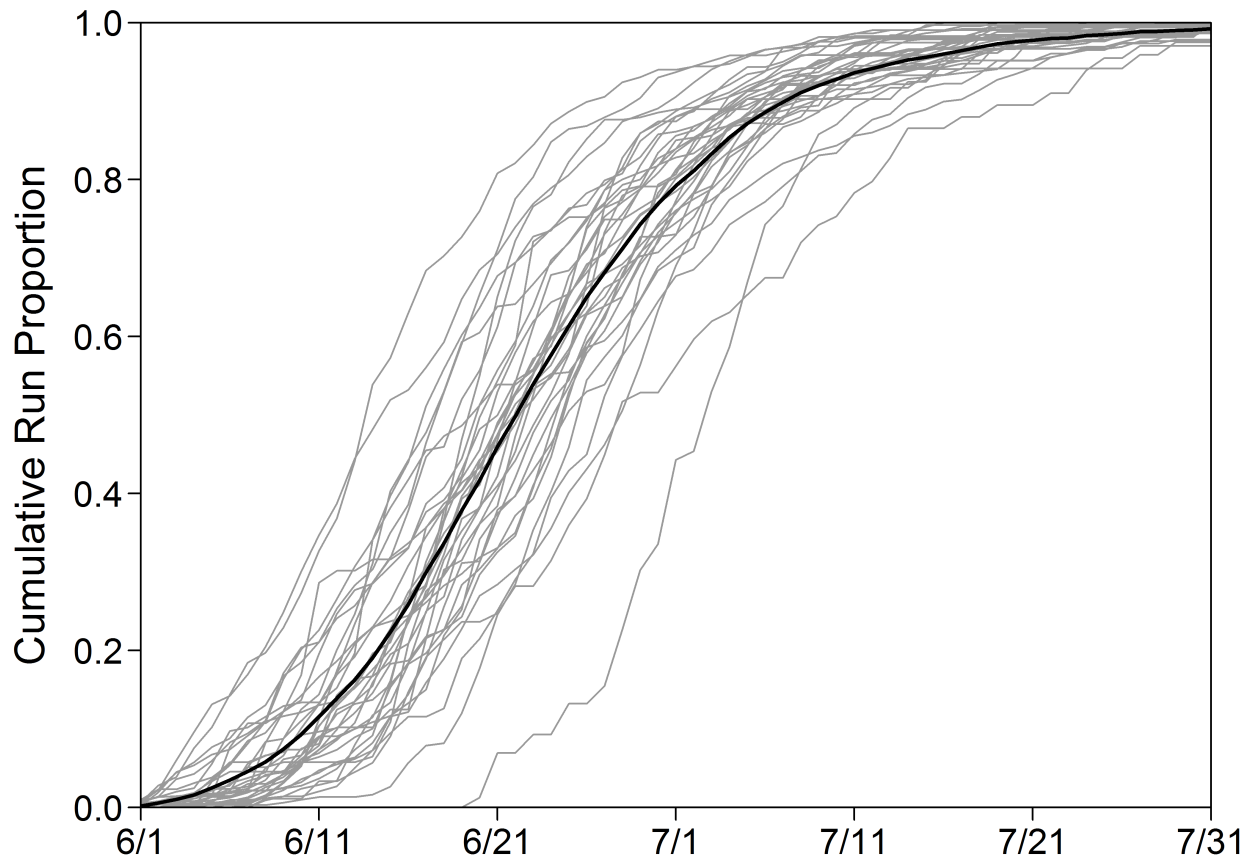
FIGURE 2.2: Shape and variability of run timing patterns of the Kuskokwim River Chinook salmon stock as sampled by the Bethel Test Fishery, 1984 – 2018. Each grey curve represents a year standardized by the total end-of-season cumulative CPE and the black line represents the average value across years on each day of the season.
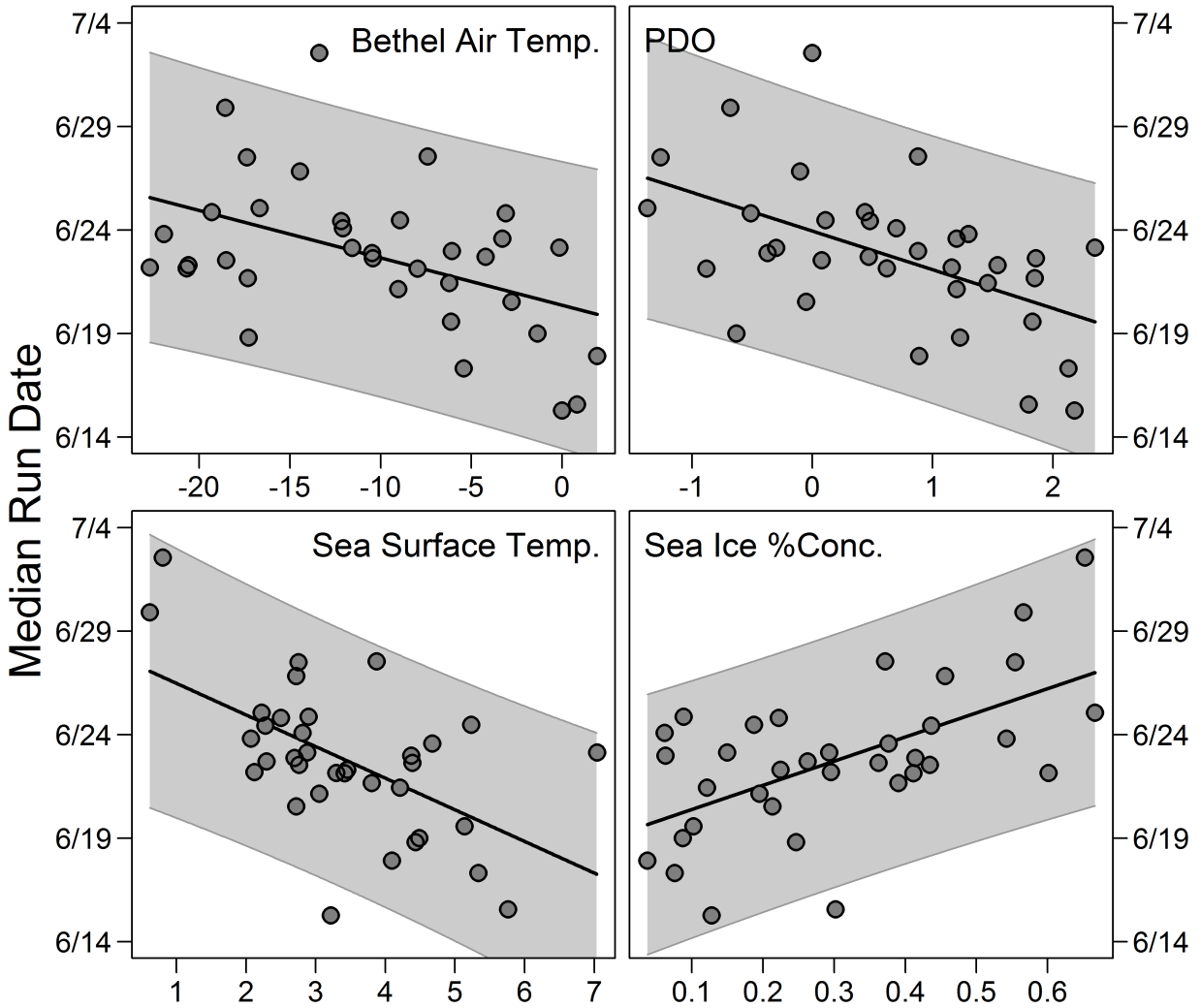
FIGURE 2.3: Relationships between the four single environmental variables and run timing (D_{50}) using data from optimal climate windows when 2018 was included in the training data. For illustration purposes only, gridded variables SST and SIC were combined by weighted averaging, where the weight of each grid cell was assigned the AIC_c weight of that grid cell when grid cell-specific models were fit. Grey bands are 95% prediction intervals.
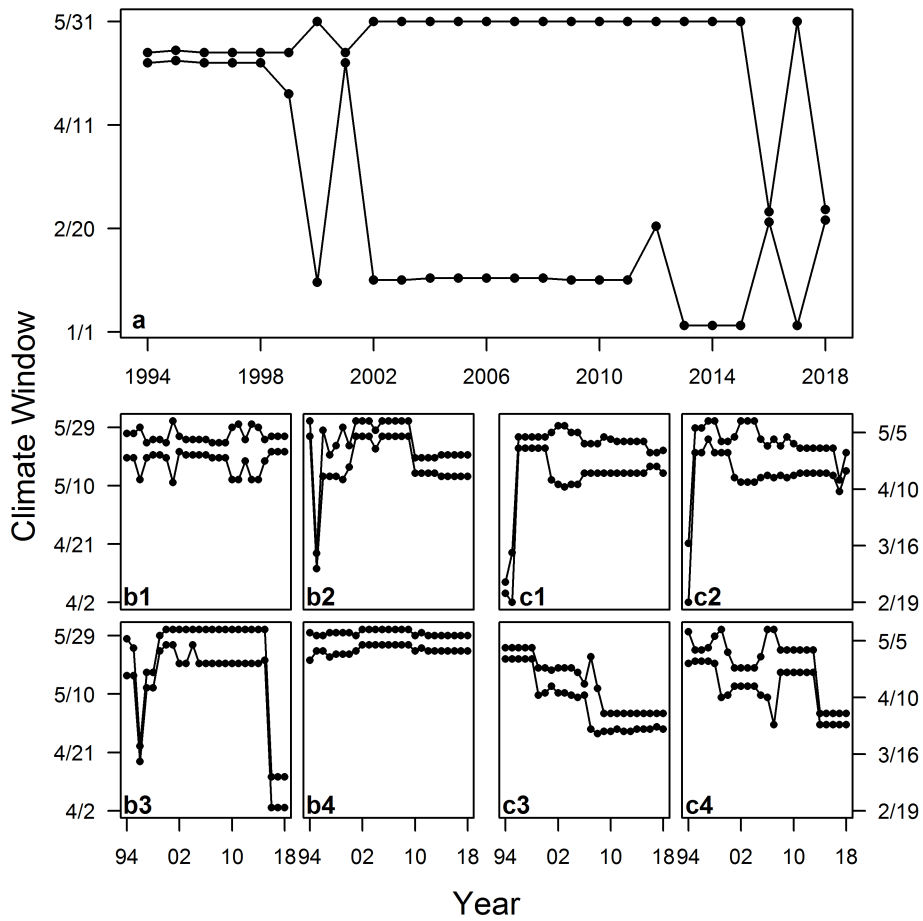
FIGURE 2.4: Changes in selected climate windows as more years of training data were added in the retrospective forecasting analysis. Bottom and top lines show the first and last day of the selected climate window, respectively, as more years were added. The year axis corresponds to the selected window after including environmental and run timing data from that year in the training data. For example, the windows shown for 2017 were used to produce the forecast for 2018. Panel (a) is Bethel air temperature, panels $b1 - b4$ are SST windows for four sample grid cells and panels $c1 - c4$ are SIC windows for the same four sample grid cells. Sample grid cells from Figure 2.1 shown for SST and SIC are as follows: grid cell 8 ($b1$, $c1$), grid cell 44 ($b2$, $c2$), grid cell 12 ($b3$, $c3$), and grid cell 48 ($b4$, $c4$). Selected windows for PDO are not shown because the single month of May was selected in all years.

47
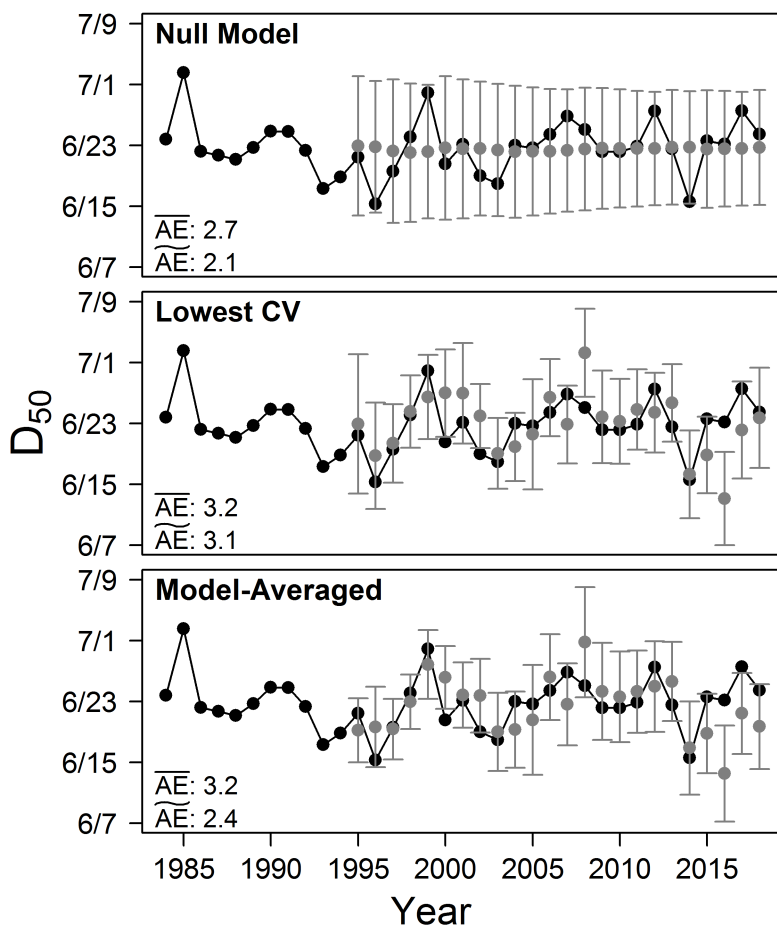
FIGURE 2.5: Observed and forecasted median run date (D_{50}) under the three approaches: the null (intercept only/historical average) model, the single regression model with the lowest forecast cross-validation score up to the forecasting year, and a model-averaged forecast that combined forecasts from 16 nested regression models. Black points/lines are the time series of D_{50} detected by the BTF. Grey points are out-of-sample forecasts with 95% prediction intervals shown as error bars. $\overline{AE}$ and $\widetilde{AE}$ are the mean and median absolute forecast errors from 1995 to 2018, respectively.
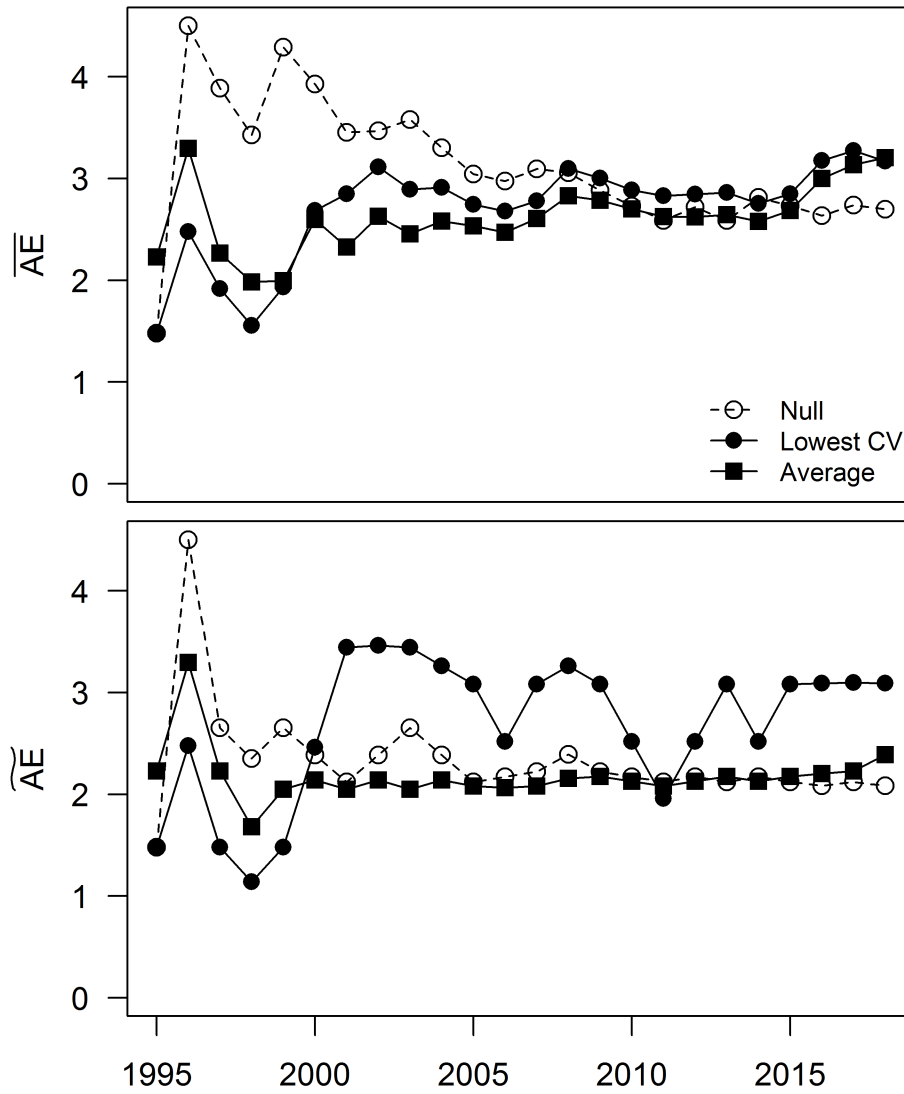
FIGURE 2.6: Evolution of mean ($\overline{\text{AE}}$) and median ($\widetilde{\text{AE}}$) absolute forecast error under the three investigated forecasting approaches: the null (intercept only/historical average) model, the single regression model with the lowest forecast cross-validation score up to the forecasting year, and a model-averaged forecast that combined forecasts from 16 nested regression models. Each point is the average of absolute errors of all years before and including the corresponding year on the x-axis, starting in 1995, which was the first year out-of-sample forecasts were evaluated.
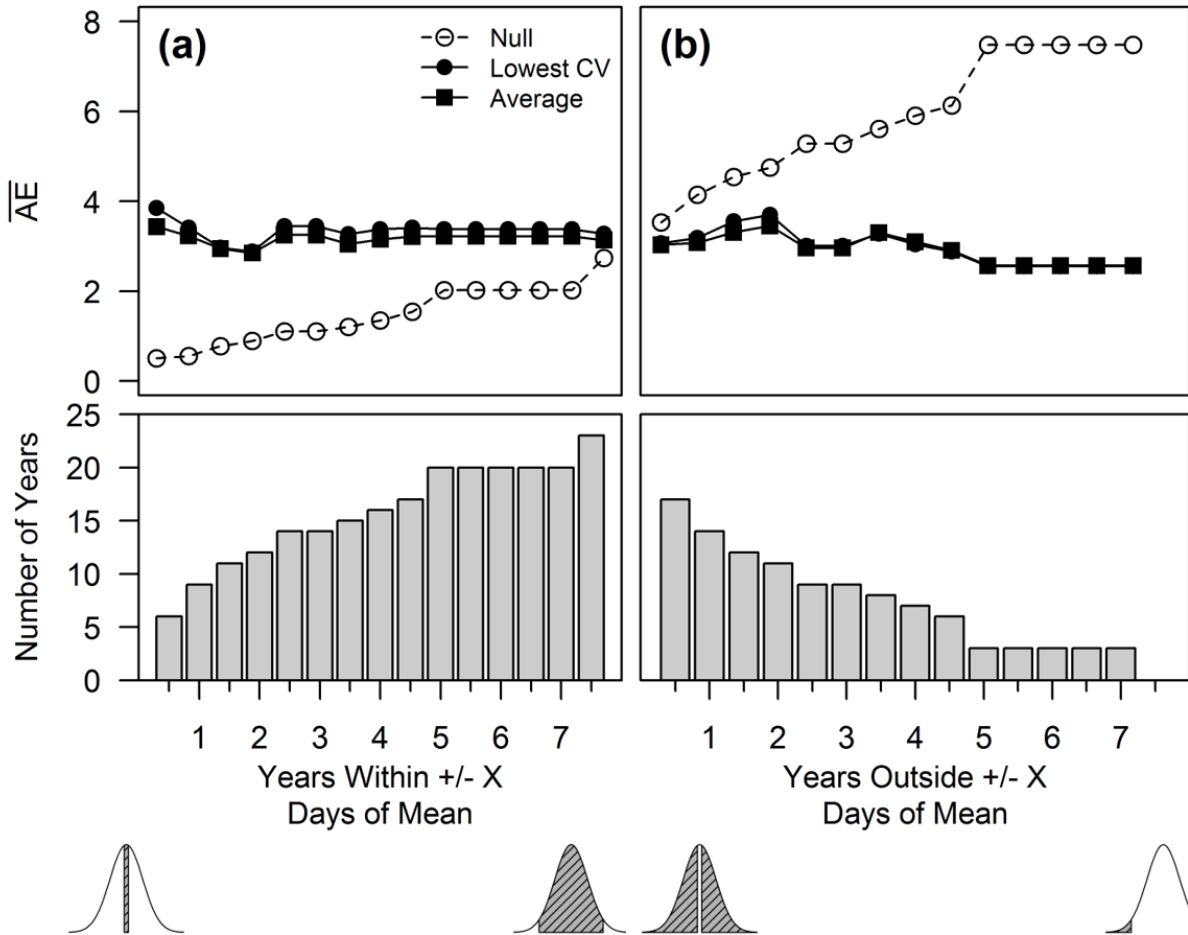
FIGURE 2.7: $\overline{AE}$ under three forecast approaches calculated by either (*a*) including years with a D_{50} value within $\pm x$ days of the all-year average or (*b*) including years with a D_{50} value outside $\pm x$ days of average, where x is the number of days indicated on the x-axis. Bottom panels show the number of observed years in which the appropriate $\pm x$ days criterion was met. Shaded regions in the hypothetical distributions show the types of D_{50} values that were included in the calculation of $\overline{AE}$. One point that may enrich inference from this figure (and is shown in the shaded normal distributions) is that panel (*a*) becomes more inclusive from left to right by adding years that are more dissimilar to the average in the calculation of $\overline{AE}$ whereas panel (*b*) becomes more exclusive from left to right by removing years that are similar to the average.
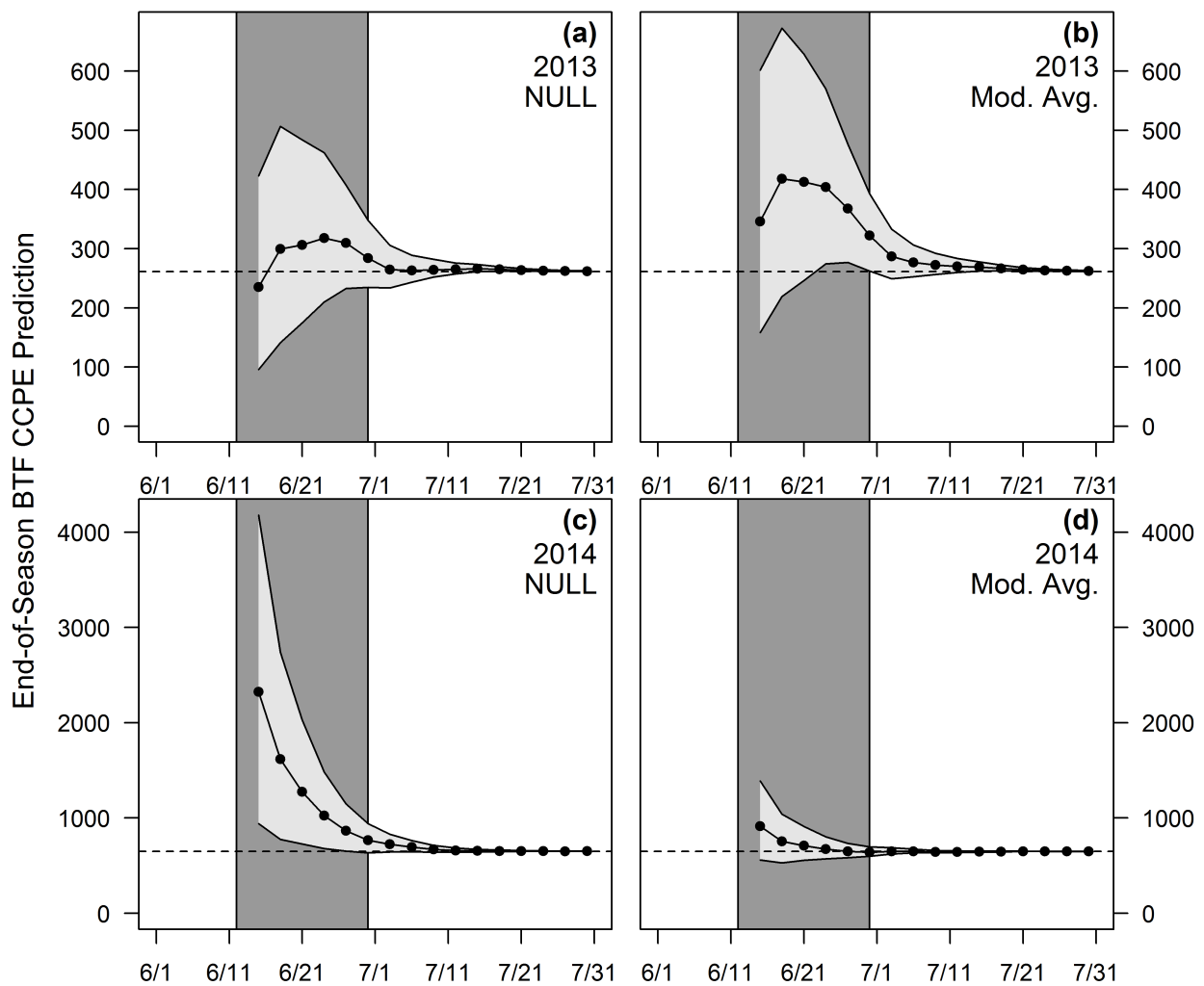
FIGURE 2.8: In-season predictions of end-of-season cumulative BTF CPE under the model-averaged forecast using environmental variables and the forecast under the null model in 2013 (panels *a* and *b*) and 2014 (panels *c* and *d*). Intended to illustrate cases in which a manager would benefit from having access to the model-averaged run timing forecast model using environmental variables (2014) and when the null model would have performed better (2013). Horizontal lines are the true end-of-season cumulative BTF CPE and light grey regions are 80% confidence intervals obtained *via* Monte Carlo simulation. Dark grey vertical regions indicate the period when key harvest decisions are made with respect to Chinook salmon in the Kuskokwim River.
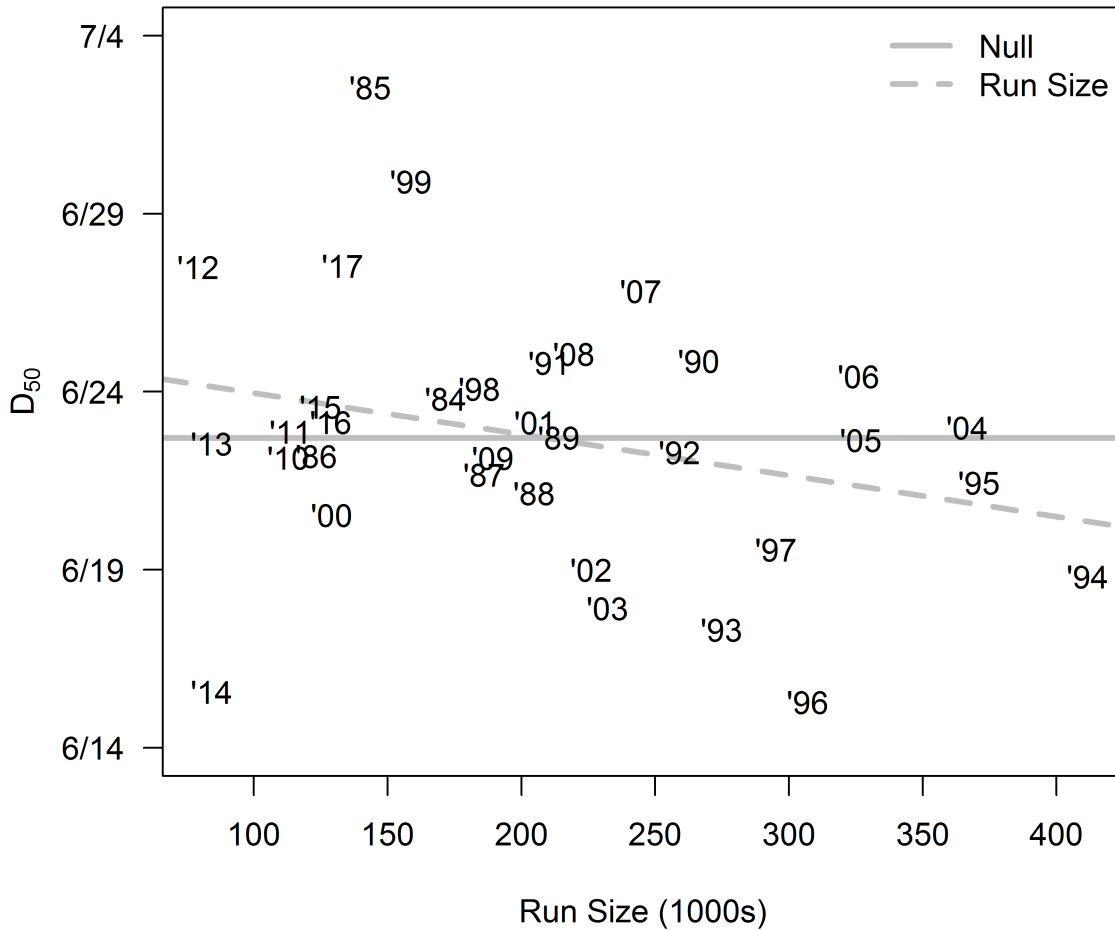
FIGURE 2.9: Relationship between D_{50} and run size for Kuskokwim River Chinook salmon with two fitted models shown: the null model (which assumed constant mean D_{50}) and the run size model (which assumed the mean D_{50} changes as a function of run size). As described in the text (Section 2.3.6), the effect of run size on run timing was weak and not significantly different than no effect. Additionally, knowledge of run size did not result in smaller average prediction errors of D_{50} than not having this knowledge.

Chapter 3

Evaluation of In-Season Harvest Management Strategies For Kuskokwim River Chinook Salmon using a Stochastic Simulation Model

Abstract

In-season management of Chinook salmon subsistence fisheries in large river basins is conducted in the presence of much uncertainty, primarily with respect to run size and timing. Managers must manipulate the amount of time in which fishing is allowed to ensure adequate escapement that will sustain future harvests while simultaneously providing as much opportunity in the current year as possible. In doing so, they may use a set of decision rules to open or close the fishery based on either intuition or assessment information. Inferences about which strategies may perform better in certain circumstances can be informed using management strategy evaluation, an analytical method in which decision rules and information sources are tested against simulated conditions to measure likely management performance. This chapter presents a management strategy evaluation for in-season harvest management for the Kuskokwim River Chinook salmon subsistence fishery in western Alaska to test four primary management strategies that ranged in their complexity and information needs. Findings showed that all assessed strategies can perform well, but that the more complex strategies tended to perform better when the incoming run was small. Additionally, the optimal settings (*i.e.*, aggressive or conservative with respect to fishing opportunity) of each strategy depended on run size, with conservative settings favored in smaller runs. The findings of this chapter extend the knowledge about in-season salmon harvest management strategies, which is mostly regarding commercial fisheries, to include subsistence fisheries as well and should be informative to fishery managers in the region.

3.1 Introduction

In-season harvest management of Pacific salmon *Oncorhynchus* spp. fisheries in large river systems is undertaken in the presence of a large amount of uncertainty about how to schedule fishing opportunities. In order to manage in a fully informed way, a manager would require continuous and accurate information on arrival timing, run size, fleet dynamics, and harvest. With knowledge on these components, it would be theoretically possible to perfectly harvest the available surplus each year (Adkison and Cunningham 2015). In reality, these quantities (when available) are often highly uncertain (Adkison and Peterman 2000; Flynn and Hilborn 2004; Hyun *et al.* 2012) which results in difficulties in decision-making about how to best implement the fishery in order to meet a set of pre-defined objectives dealing with both conservation and exploitation.

In addition to the substantial uncertainty in decision-making, there are often sharp trade-offs among competing objectives, such as the desire to provide adequate and equitable harvest opportunity *versus* the desire to ensure adequate escapement (Catalano and Jones 2014). Oftentimes, managers are also concerned with spreading exploitation evenly among stock subcomponents (Schindler *et al.* 2010), but this may conflict with aspects dealing with the ideal time to harvest salmon as a result of weather or fish quality conditions (Adkison and Cunningham 2015; Carney and Adkison 2014b). When given the task of balancing trade-offs such as these, the manager has the ability to manipulate the fishing gear used as well as the spatiotemporal distribution of fishing effort by opening or closing the fishery for various amounts of time, though it is rarely clear as to how to manipulate these management "levers" to achieve the desired outcomes. Presumably, different approaches for performing these manipulations (termed "management strategies") will exhibit differential performance at meeting the objectives and balancing trade-offs.

Management strategy evaluation (MSE) has been proposed as a powerful tool for determining how to manage exploited natural resource systems with competing management objectives (Butterworth 2007; Cooke 1999). MSE is a stochastic simulation-based analytical technique whereby management strategies are evaluated by comparing their relative performance at meeting pre-defined objectives under simulated (though realistic) conditions. A management strategy can be thought of as all of the steps that encompass the collection of data, subsequent analyses, and resulting decision-making surrounding the exploitation of a resource. The MSE approach tests a range of such strategies to find the one(s) that are likely to be most robust to uncertainty and balance trade-offs. This approach is powerful as it can provide general insights without having to test strategies on the real system, which would be incredibly time-intensive (each year is one sample) and costly given that some candidate strategies can be risky (Walters and Martell 2004). Punt *et al.* (2014) outlined a set of seven steps to an MSE that must be conducted in order for the analysis to be meaningful:

(1) identification of management objectives and performance measures for each; preferably under the direction of stakeholders and managers,

(2) identification of the key uncertainties present in the system (biological, assessment, implementation, *etc.*),

(3) identification of candidate management strategies for evaluation,

(4) development of one or more models that serve as the representation of the real system including reasonably realistic representations of biological and fishery components (termed the "operating model"),

(5) selection of parameters to drive the operating model in accordance with the real system,

(6) simulation of executing each strategy using the operating model(s), and

(7) summary of performance measures, and presentation to managers and stakeholders.

Two broad classes of strategies could be conceived for in-season salmon management: effort control using either (1) a fixed schedule set at the start of the season or (2) a feedback

strategy where the fishery is opened or closed in response to in-season data (*i.e.*, management by emergency order, Adkison and Cunningham 2015). There exist many substrategies that fall into these two broad categories based on (1) the level of risk aversion on the part of the manager (*i.e.*, aggressive *versus* conservative) and (2) the timeliness and reliability of information available to the manager. In general, more complex strategies will require more data to inform their implementation (Carney and Adkison 2014b). Given the wide range of strategy complexity, it is worthwhile investigating if more complex (and data-intensive) strategies provide better management performance than simpler strategies that use less information. Carney and Adkison (2014a) and Carney and Adkison (2014b) evaluated feedback *versus* fixed schedule strategies for sockeye salmon *O. nerka* stocks in Bristol Bay, Alaska, and found trade-offs between maximizing harvest and reducing inter-annual variability in harvest magnitude as well as spreading harvest pressure among substock components. Su and Adkison (2002) evaluated a set of schedule-based strategies that ranged in their aggressiveness and found differences in strategy performance based on which objective carried most weight in utility functions, which implies that trade-offs exist.

An MSE analysis for subsistence salmon fisheries in large drainages (such as the Yukon and Kuskokwim systems in western Alaska) necessitates different considerations than these other examples which focused on commercial fisheries. While the types of strategies considered and conservation-based objectives (adequate escapement and temporally distributed harvest) are broadly consistent, the fleet dynamics and harvest-based objectives may be different. Subsistence fishers are less concerned with maximizing harvest as they are with maintaining consistent harvests that meet their needs and that harvest opportunities allow exploitation consistent with cultural practices (*e.g.*, time of season and frequency of opportunities). The fleet dynamics of subsistence fisheries are quite different than commercial fisheries in that they are limited by processing capacity and have a fixed targeted harvest for the season. Due to this processing capacity, harvest of targeted species (such as Chinook salmon *O. tshawytscha*)

in subsistence fisheries is limited by the species composition, sometimes expressed as a ratio of chum *O. keta* + sockeye:Chinook salmon. Subsistence fishers must stop fishing when they reach their processing capacity, and when this ratio is high (*e.g.*, > 20), the catch will be dominated by chum/sockeye salmon. In-season harvest management strategies that acknowledge these characteristics have not been evaluated for subsistence salmon fisheries, highlighting a clear need for work that focuses on this topic.

In this chapter, the performance of a variety of in-season harvest control rules for subsistence salmon fisheries in large drainage systems is investigated using a MSE approach. Though the analysis will be tailored to the Kuskokwim River Chinook salmon subsistence fishery, the framework developed will be general enough for application to other in-river salmon fisheries in large drainages in which the primary users are subsistence fishers, and relative slight modifications would be necessary to accommodate simulation of commercial fisheries. The objectives of the analysis were to:

(1) develop a stochastic simulation model tailored to the Kuskokwim River fishery system that would allow simulation of a wide range of biological conditions,

(2) assess the performance of several realistic in-season harvest management strategies that capture a range of complexity in their management dexterity and need for information, and

(3) highlight the strength of trade-offs among competing objectives, and find management strategies that might balance them better than others.

3.2 Methods

The analysis was carried out by developing a stochastic simulation model of a subsistence salmon fishery system and imposing several management strategies separately. The operating model, which simulated the system dynamics, was tailored to the Kuskokwim River subsistence salmon fishery and had a reasonably complex spatiotemporal structure (see Section 3.2.3).

Four primary strategies were identified (see Section 3.2.2) based on input from managers, biologists, and stakeholders from the Kuskokwim River drainage, as well as from academic experts in the field of Pacific salmon management. These strategies were explicitly selected to explore a range of complexity, with more complex strategies requiring more information for their implementation. Each primary strategy had several substrategies varied in the degree of aggressiveness in allowing fishing opportunities according to the rules of the primary strategy. Each management strategy was tested by simulating many hypothetical and independent salmon seasons in a Monte Carlo framework such that performance was tested at many different run scenarios including run size, run timing, and species composition. Performance of each strategy and substrategy was assessed relative to the attainment of four objectives (Section 3.2.1) using a set of utility functions (Section 3.2.5).

3.2.1 Identification of management objectives

As indicated by Punt *et al.* (2014), the objectives selected for evaluation in an MSE analysis should be informed by communications with stakeholders and managers to determine what outcomes are deemed desirable. As part of a complementary project intended to build capacity in the engaged representatives from the local stakeholder group, four multi-day workshops were held in Alaska over the period spanning autumn 2015 – 2017. The workshops were led by by experts in meeting facilitation and salmon biology and management and were highly interactive. Presentations were given about the difficulties in salmon management, the basics of their biology, the ways information can be used in decision-making, and ways that simulation models can be used to evaluate management strategies. In the first of these workshops, stakeholders and managers were solicited for input regarding which outcomes are important from their perspective. Based on the themes that emerged, four main objectives for Chinook salmon management at the in-season level were identified. This is a critical component of this analysis, because the objectives define the necessary complexity of the

operating model and they provide the context for measuring which strategies might perform better than others. They can be grouped as follows:

Sustainability-based

(1) Ensure adequate drainage-wide Chinook salmon escapement to the spawning grounds to support sustained subsistence yields into the future,

(2) Ensure that the Chinook salmon substocks have even exploitation rates within a given year,

Exploitation-based

(3) Ensure that Chinook salmon subsistence harvest needs are met at the basin-scale,

(4) Ensure that when Chinook salmon harvest restrictions are necessary, the burdens are spread evenly among the various villages.

This list is provided here to set the context for the rest of the methods, see Section 3.2.5 for a description of the utility functions used to measure the attainment of each objective. In this analysis, it was assumed that the abundance of chum/sockeye salmon was high enough to meet both harvest and escapement needs, so no objectives were developed regarding their management.

3.2.2 Assessed management strategies

A set of four primary in-season harvest management strategies were evaluated for this analysis. Managers in large salmon-producing river basins have the tools of time, area, and gear restrictions at their disposal for managing harvest. Strategies assessed here focused primarily on the time (*i.e.*, when in the season fishing is allowed) aspect of these tools. Each of the four strategies represented a different way of determining if the fishery should be open on a given day of the season. Given the historical season for Chinook salmon (the species of interest in this analysis) management in the Kuskokwim River, each strategy focused on a five week period between June 1 and early July. Based on Chinook salmon run timing through

the lower Kuskokwim River (50% complete on June 22 in an average year, see Chapter 2, this dissertation) and the timing that chum and sockeye salmon become vastly dominant in the species composition of the run (Figure A.2), it is only during this time that management actions affecting subsistence harvest can have any meaningful impact on the attainment of Chinook salmon objectives (both those based in conservation and exploitation).

3.2.2.1 Strategy #1: "Closed until open"

Under this first and most naïve management strategy, the simulated manager selected a single day on which to open the entire fishery, before which it remained completely restricted (closed) and after which it remained unrestricted (open) for the rest of the season. The decision of which day to open was not explicitly informed by any "previous data" on the part of the manager, or changed based on in-season information. Three reasonable dates to start the fishery were evaluated: June 1, June 12, and June 23. These dates represent the historical average 1%, 12%, and 55% percentage points of the Chinook salmon run as indexed by the Bethel Test Fishery (Bue and Lipka 2016).

3.2.2.2 Strategy #2: "Forecast-based fixed schedule"

Under this strategy, the manager used a pre-season run size forecast (described in Section 3.2.4.1) with which to inform the decision about how often fishing opportunities should be provided. This was conducted by developing categories (hereafter "bins") of run sizes that triggered a decision regarding how many days to allow fishing in each week: *e.g.*, if the run was forecast to be less than 80,000 Chinook salmon, the number of days of fishing allowed per week would be less than if the the run was forecast to be between 130,000 and 180,000. Substrategies were represented by three different sets of schedules conditional on the pre-season forecast, ranging from conservative (fewer fishing days per week) to aggressive (more days per week).

In developing these schedules that dictated how many days (D) the fishery would be open during week w conditional on a forecast falling in bin b, three main qualities were desired. First, for any week $w \geq 0$ (first week) and forecast bin $b \geq 0$ (smallest bin), $D_{w,b}$ for conservative schedules should be less than the neutral and aggressive schedules, and aggressive schedules should have the highest $D_{w,b}$ in the same w and b. Second, $D_{w,b}$ should generally increase as the forecast bin increases – *i.e.*, years with larger anticipated runs can allow fewer restrictions to the fishery. Finally, $D_{w,b}$ should generally increase as the season progresses (increasing w), because the species composition shifts towards chum/sockeye salmon later in the season lessening the concern for high catches of Chinook salmon that may endanger the ability to meet escapement needs.

I developed a linear model that would return $D_{w,b}$ depending on the week w, forecast bin b, and schedule type (*i.e.*, aggressive *versus* conservative; Figure 3.1). The model took the form:

$$D_{w,b} = \delta_0 + \delta_1 C + \delta_2 A + \delta_3 w + \delta_4 Cw + \delta_5 Aw + \delta_6 b^2 + \delta_7 bw, \quad (3.1)$$

where C and A are dummy variables indicating either conservative or aggressive schedules, respectively, w is the week index (five weeks: $0 \leq w \leq 4$), b is the forecast bin index (five bins: $0 \leq b \leq 4$). A and C are mutually exclusive and $A = C = 0$ for the neutral schedule. The vector δ contains coefficients for how $D_{w,b}$ depends on the values of the covariates (C, A, w, and b):

$$\delta = \begin{bmatrix} 0.25 & -0.25 & 0.25 & 0.25 & -0.50 & 0.50 & 0.50 & 0.50 \end{bmatrix}$$

For example, in the first week ($w = 0$), first bin ($b = 0$), and the neutral schedule ($A = C = 0$), $D_{w,b} = \delta_0 = 0.25$. For the same b and w, $D_{w,b} = \delta_0 + \delta_1 = 0$ for the conservative schedule ($C = 1$) and $D_{w,b} = \delta_0 + \delta_2 = 0.5$ for the aggressive schedule. The slope of conservative and aggressive schedules differ from the neutral schedule by -0.5 and 0.5 days/week in all bins,

respectively, and all slopes increase by 0.5 days/week for each increase in bin. The intercept of all schedules increases by $0.5b^2$ days for each increase in the bin. Cases in which $D_{w,b}$ would exceed 7 days were rescaled such that $D_{w,b} = 7$, the same was done to prevent $D_{w,b} < 0$.

3.2.2.3   Strategy #3: "Forecast/ratio-based variable schedule"

This strategy was similar to Strategy #2 in that it used a pre-season forecast to set a schedule for each week, though rather than treating the different possible schedules as conservative or aggressive substrategies, the manager treated them as alternatives to be employed selectively based on additional information. The manager made this selection based on in-season species composition information collected at a simulated test fishery site (described in Section 3.2.4.2). The species composition (expressed as a ratio in terms of chum+sockeye:Chinook salmon) is an important aspect of the fishery, because subsistence fishers are self-limited in the number of fish they can successfully process per fishing trip, and Chinook salmon harvest can be limited during times when the species ratio is high. Based on the historical percentile of the ratios in the previous week ($\phi_{p,w-1}$), the manager selected either the conservative, neutral, or aggressive schedule for the appropriate forecast bin b for use in week w as indicated in Figure 3.1.

Three substrategies were assessed, dealing with how the trigger percentiles were selected, as shown in Table 3.1. The "neutral" set of ratio trigger points specified that the manager would employ conservative schedules in accordance with the forecast bin until $\phi_{p,w-1}$ exceeded the 33% percentile of all historical ratios, at which point they would use the appropriate neutral schedule (from Figure 3.1). If at any w, $\phi_{p,w-1}$ exceeded the 66% percentile, the manager would switch to the aggressive schedule. The rationale here is that the more chum and sockeye there are relative to each Chinook salmon, the fewer Chinook will be caught per fishing trip and the more opportunity can be allowed for species of non-conservation concern. The "conservative" substrategy used cut-offs of 66% and 85% to make these transitions, and

the "aggressive" substrategy used cut-offs at 15% and 33% (Table 3.1). The resulting ratio trigger points are shown in Table 3.2.

3.2.2.4   Strategy #4: "Explicit harvest target"

Under this strategy, the manager took on a much more active decision-making process wherein they decided how many days to allow fishing in each week of the season based on an explicit harvest target (H_T) selected probabilistically to ensure some escapement limit point (S_L) would be exceeded that season. This was the most complex management strategy, as the manager needed information on how much harvest had been taken to date and how long they should allow fishing each week based on how many fish they wish to allow to be caught. H_T was apportioned among weeks ($H_{T,w}$) according to historical Chinook salmon run timing and represented the number of Chinook salmon the manager wishes to see harvested in week w. $H_{T,w}$ could be updated in response to (1) whether in-season abundance index data suggest the Chinook salmon run is either smaller or larger than forecast or (2) whether harvest data suggest the fishery is either ahead of or behind schedule in meeting H_T.

This strategy had two main phases as shown in Figure 3.2. In the pre-season phase, managers used a forecast, management target, and risk tolerance to set a value for H_T and $H_{T,w}$ to start the season. Then, the in-season phase proceeded as a weekly cycle of Bayesian run abundance estimation (described in Section 3.2.4.3), re-evaluation of H_T in accordance with updated knowledge and S_L, determination of remaining harvest, a decision of the number of days to fish based on an updated $H_{T,w}$, and estimation of harvest outcomes.

Three substrategies were formulated by building three different "harvest tables" which dictated how many days the fishery should be open in week w based on the value of $H_{T,w}$ and differed in how aggressive or conservative they were (Figure 3.3). The neutral table started with 0.5 days for the case of $0 < H_{T,w} \leq 5,000$ and increased by 1 day for each additional 5,000 Chinook salmon in $H_{T,w}$. The aggressive harvest table resulted in fishing 1.5 times as

63

many day as the neutral table for all $H_{T,w} > 0$. If this rule would result in greater than 7 days it was capped at 7 days. The conservative table was constructed the same way except with 0.5 times as many days as the neutral table.

The probabilistic approach to selecting and updating the season-wide harvest target (H_T) in this fourth and most complex assessed management strategy is a relatively novel approach to the management of Pacific salmon fisheries (but see Catalano and Jones 2014, for another application using simulation techniques). The problem is to select some value for H_T that will ensure the drainage-wide total escapement (S) will exceed some critical escapement limit threshold (S_L) with probability greater than $1 - P^*$. The quantity P^* represents a manager's tolerance for risk of seeing the undesirable outcome of $S \leq S_L$ occur. $\Pr(S \leq S_L \mid H_T)$ can be calculated from a cumulative probability density function expressing beliefs about total run size. If F_N is this expression of beliefs (and H_T is assumed to be a constant), then $F_N(S_L + H_T) = \Pr(N \leq S_L + H_T) = \Pr(S \leq S_L \mid H_T)$. The value H_T can be manipulated to ensure the condition $\Pr(S \leq S_L \mid H_T) \leq P^*$ is satisfied. When new information accumulates in F_N (*i.e.*, as a posterior probability density function from daily or weekly Bayesian updates of the pre-season run forecast; Section 3.2.4.3), H_T can be updated as well to ensure the condition is still satisfied. For this analysis, $S_L = 65,000$ (the lower bound of the current drainage-wide escapement goal for Chinook salmon; Hamazaki *et al.* 2012) and $P^* = 0.1$. This probabilistic harvest control rule is similar to those used in marine fisheries when setting sustainable fishing mortality targets, and explicitly accounts for uncertainty and risk when determining allowable fishing activity based on limit management reference points (Prager *et al.* 2003; Shertzer *et al.* 2010).

3.2.3 Description of the operating model

The role of the operating model was to simulate the true dynamics of the fishery system, which included the important dynamics of the biological (*i.e.*, the salmon) and social (*i.e.*, the

fishers) components of the fishery. The operating model was structured such that important spatial and temporal dynamics of fish and fishers in the Kuskokwim River subsistence salmon fishery could be captured. The biological and fishery components of the operating model were informed using as much empirical information as possible (see Appendix A for a description of data sources and preparation for use in these contexts). Furthermore, simulated outcomes of the fishery (*i.e.*, magnitude and spatiotemporal distributions of Chinook salmon harvests) under a "no management" scenario were compared to those observed in historical data in years the subsistence fishery was unrestricted (Appendix B). This was an important validation of the behavior of the operating model to ensure it adequately reproduced the patterns and variability of inter-annual observations from the real system according to the best available scientific information.

The operating model tracked in-river salmon abundance, fishing effort, harvest, and escapement in each of in each of 26 discrete river reaches (hereafter indexed by r) along the main-stem Kuskokwim River over the span of approximately 130 days (late-May to the start of October; hereafter indexed by d). Although the month of June and early July are the primary salmon harvest periods in the Kuskokwim River subsistence fishery, this long temporal scale was needed to allow all simulated fish to migrate completely through the entire Kuskokwim River model. The operating model was written in Program R (R Core Team 2018).

3.2.3.1   Biological components

The biological submodel was made up of two aggregate salmon populations: one Chinook salmon population and one of chum and sockeye salmon together. Chinook salmon are the species of primary management interest in this analysis; the other species were included because harvest dynamics for Chinook salmon are influenced by the relative abundance of all three species in the harvesting gear. The Chinook salmon population was subdivided into

three spatially explicit substocks representing spawning aggregations in the lower, middle, and upper reaches of the drainage, which was necessary to assess the equal exploitation rate objective and enforce the realities of in-river sequential (*i.e.*, "gauntlet") fisheries. River entry timing and relative abundance of each Chinook substock was informed by Kuskokwim River telemetry studies (Stuby 2007; Smith and Liller 2017a,b). These studies indicate that the middle river substock is the largest (~60% of the total abundance) and enters the river mixed with the tail-end of the upper river substock (~20% of the total abundance). The lower river substock enters mixed with the middle river substock and is approximately the same size as the upper river substock.

To initialize the model, the size of the total abundance of Chinook salmon (N_{tot}) that would return to the system in the simulated year was obtained as a random sample from a distribution with density equal to that of the historical distribution of run sizes over the period (1976 – 2017; as presented in Liller *et al.* 2018, and further described in Appendix A.1.1). The total annual abundance of each Chinook salmon substock (N_s) was then obtained:

$$N_s = N_{tot}\pi_s, \tag{3.2}$$

where π_s is a Dirichlet random vector representing the proportion of the total run made up of fish returning to each of the three Chinook salmon substocks with hyperparameters informed by the distribution of radio telemetry tagged fish (see Appendix A.1.2 for details). The number of fish from each Chinook salmon substock that entered the first reach each day of the season was then populated:

$$A_{d,1,s} = N_s p_{d,s}, \tag{3.3}$$

where $A_{d,1,s}$ is in-river abundance on day d in reach $r = 1$ for substock s and $p_{d,s}$ is a run timing variable representing the fraction of the run from that substock entering on that day

of the season. $p_{d,s}$ was modeled using a logistic density function, standardized to sum to one within each substock over the season:

$$p'_{d,s} = \frac{e^{\frac{d-D_{50,s}}{h_s}}}{h_s \left(1 + e^{\frac{d-D_{50,s}}{h_s}}\right)^2}, \tag{3.4}$$

$$p_{d,s} = \frac{p'_{d,s}}{\sum_d p'_{d,s}}, \tag{3.5}$$

where $p'_{d,s}$ are elements of the unstandardized timing curve as given by the substock-specific location ($D_{50,s}$) and scale (h_s) parameters, also informed using the telemetry data (Appendix A.1.3.2).

Detailed information regarding total abundance or spatial differences in run timing of various substocks of Kuskokwim River chum and sockeye salmon is not available. Accordingly, the aggregate population representing these species was modeled using historical estimates of daily relative abundance from a long time series of a standardized catch per effort (CPE) index (the Bethel Test Fishery – BTF; Bue and Lipka 2016). Daily relative abundance was represented by ϕ_d, calculated as the observed ratio of the CPE of chum + sockeye salmon to Chinook salmon (Appendix A.1.5). Simulated entry timing and abundance of the chum/sockeye aggregate stock was obtained from the total daily entering abundance of Chinook salmon and a randomly drawn annual vector of ϕ from the historical data set:

$$A_{d,1,4} = \phi_d \sum_{s=1}^{3} A_{d,1,s} \tag{3.6}$$

The movement of fish through the main-stem of the river was modeled using a "boxcar" approach (Walters and Martell 2004), in which each reach had associated rates of in-river "mortality" (*i.e.*, the removal of fish from the main-stem due to fishery harvest and escapement). The main-stem mortality rate resulting from fish escaping to spawning tributaries for reach r

for substock s ($\psi_{r,s}$) was obtained using the historical telemetry studies (Appendix A.1.4) and represented the fraction of all fish from substock s that survived all harvesters prior to and including reach r that would spawn in a tributary with a confluence with the main-stem in reach r. As telemetry information was only available for Chinook salmon ($s = 1$, 2, or 3), $\psi_{r,s}$ for the chum/sockeye stock ($s = 4$) was assumed to be the same as for Chinook salmon, though with the removal of the spatial substock structure (Table A.2). Many factors contributed to the simulated fishing mortality rate in reach r on day d, as described in Section 3.2.3.2, though it was assumed that fishing mortality occurred before escapement mortality:

$$S_{d,r,s} = \psi_{r,s} \left(A_{d,r,s} - H_{d,r,s} \right), \tag{3.7}$$

where $S_{d,r,s}$ is escapement and $H_{d,r,s}$ is harvest. Any fish that survived these sources of main-stem mortality remained in the main-stem, but would transition to the next reach on the next day with probability equal to one:

$$A_{d+1,r+1,s} = A_{d,r,s} - H_{d,r,s} - S_{d,r,s} \tag{3.8}$$

All reaches were assigned a length of 35 km, which is the approximate mean estimated travel distance per day for Chinook salmon in the main-stem Kuskokwim River (Smith and Liller 2017a,b).

3.2.3.2 Fishery components

There were five primary factors used to model the subsistence fishery dynamics in each reach: (1) maximum daily effort ($E_{\text{MAX},r}$; effort expressed in boat trips per day), total maximal salmon need by species, maximum daily salmon catch per boat per day (abbreviated by CPB; maximum is denoted CPB_{MAX}), (4) effort responses to fishery conditions, and (5) a measure of fishery selection for different species. Since 1990, the Alaska Department of Fish

and Game (ADF&G) has conducted rigorous post-season sampling from the 26 villages in the Kuskokwim River documenting the number of fishing households and salmon harvest by species (these estimates are presented in Hamazaki 2011; Carroll and Hamazaki 2012; Shelden *et al.* 2014; Shelden *et al.* 2015; Shelden *et al.* 2016a; and Shelden *et al.* 2016b). This wealth of information was used to inform maximal salmon need and effort (described in Appendices A.2.1 and A.2.2, respectively) for villages in each reach r. CPB_{MAX} and effort responses were informed by recent studies of the in-season subsistence fishery dynamics in the lower Kuskokwim River (Staton and Coggins 2016, 2017; Staton 2018d) and fishery selection was obtained by comparing these data with the catches at the BTF on the same day in the same years.

An effort response model was needed to replicate observed patterns in effort dynamics in recent years, namely that effort declines as the season progresses (Staton and Coggins 2016, 2017; Staton 2018d). This decline is thought to be a result of two primary factors: attainment of harvest needs and in-river species composition, but finer-scale factors (*e.g.*, weather) are certainly at play as well. A logit-linear model was constructed to specify the fraction of maximum fishing effort that would fish in each reach each day if the fishery were open ($p_{E,d,r}$):

$$\text{logit}(p_{E,d,r}) = \beta_0 + \beta_1 full_{d,r} + \beta_2 stop_{d,r} + \beta_3 \delta_{d-1,r,CH} + \beta_4 \delta_{d-1,r,CS} + \beta_5 \phi_{d,r} \qquad (3.9)$$

The effort response model operated on a reach-specific basis, and had five terms in addition to the intercept (β_0):

Time of season effects: β_1 and β_2

β_1 was an effect used to increase effort to near-full capacity after a critical date. The indicator $full_{d,r}$ took on a 0 value prior to this date and a 1 after it; the critical date was in early June

for the first reach and increased by one-quarter day for each upstream reach. This effect was intended to capture the behavior that few fishers will participate early in the season before many fish have arrived in their area. Additionally, it is reasonable to expect that essentially all lower-river fishers will be done fishing for Chinook, chum, and sockeye salmon by mid-July (Hamazaki 2008), so the B_2 coefficient was included to force effort to drop to near 0 around this time for lower-river villages, where $stop_{d,r}$ had the same one-quarter day lag for upstream villages as done for $full_{d,r}$.

Attainment of subsistence needs effects: β_3 and β_4

The covariates $\delta_{d-1,r,CH}$ and $\delta_{d-1,r,CS}$ represented the cumulative fraction of met needs by villages in reach r as of the previous day for Chinook and chum/sockeye salmon, respectively. β_3 and β_4 had negative values, which reflected the nature of a subsistence fishery that more fishers will exit the fishery as the season progresses and more harvest needs are met.

Species composition effect: β_5

β_5 was a response to the local in-river species ratio of chum+sockeye:Chinook salmon. It has been observed in recent years that effort declines as the season progresses (and chum/sockeye become more abundant in-river) even when Chinook, chum, and sockeye salmon needs are far from being met (as defined by the Amounts Reasonably Necessary for Subsistence Needs as determined by the Alaska Board of Fisheries; ANS; Table A.4). The important mechanism captured here is that the species composition and abundance of chum and sockeye salmon becomes so high in late June (Figure A.2) that it is not uncommon to catch several dozen fish of these species in a single gillnet drift, which may be undesirable to some fishers given limited processing and storage capacity.

The general pattern that arises from this model is low effort early in the season due to low in-river abundance and catch rates, a peak when most harvesting activity occurs due to favorable catch rates, and a rapid decline as salmon needs are met. The coefficients were selected to generally reproduce recent observations of effort dynamics (Staton and Coggins

2016, 2017; Staton 2018d) and historical harvest timing data (Hamazaki 2008, 2011, and see Appendix B for a validation). Coefficient values were $\beta_0 = 0$; $\beta_1 = 3$; $\beta_2 = -100$; $\beta_3 = -4$; $\beta_4 = -5.5$, and $\beta_5 = -0.05$ – note that the effect for attainment of Chinook salmon needs was weaker than that of chum/sockeye. This indicates that effort should decline more quickly with the attainment of chum/sockeye needs rather than for Chinook salmon, which was intended to reflect the desirability of the latter species to subsistence fishers in the Kuskokwim drainage.

Subsistence fishers are limited by processing time and space, and thus have a self-imposed catch limit. CPB_{MAX} was needed to prevent CPB from being proportional to in-river abundance at high salmon densities. A value of 60 total salmon per day was used, and came from a mixture of recent observations (Staton and Coggins 2016, 2017; Staton 2018d) and from speaking with stakeholders about their harvest and processing behavior.

It has been observed that fishers in the Kuskokwim River do not target all salmon species in proportion to their relative abundance as indexed by the BTF (Staton and Coggins 2016, 2017; Staton 2018d). Whether due to a size-selective bias of the gear or due to fisher preference, the observed species ratio in the fishery is typically skewed more towards Chinook salmon than is the BTF on the same days, by a factor of approximately 0.6. That is, if the BTF (which is assumed to sample the vulnerable relative abundance representatively) exhibits a species ratio of 15:1 (chum+sockeye:Chinook), the fishery would be expected to exhibit a species ratio of 9:1. This selectivity correction was included into the fishery model when apportioning harvest to species.

Realized effort on day d in reach r $(E_{d,r})$ was calculated by combining $E_{\mathrm{MAX},r}$, $p_{E,d,r}$, and the fraction of a 24-hour day the fishery was open $(F_{d,r})$:

$$E_{d,r} = p_{E,d,r} E_{\mathrm{MAX},d,r} F_{d,r} \tag{3.10}$$

$F_{d,r}$ was manipulated by the management strategies presented in Section 3.2.2. Total salmon harvest ($H_{d,r,tot}$) was obtained as:

$$H_{d,r,tot} = \min\left(1 - e^{-E_{d,r}q}\sum_{s=1}^{4} A_{d,r,s}, E_{d,r}CPB_{\text{MAX}}\right) \tag{3.11}$$

The term $1 - e^{-E_{d,r}q}$ is equivalent to a daily exploitation rate in the absence of processing capacity, and includes effort and catch efficiency (*i.e.*, catchability; q). The minimum statement in (3.11) enforces the maximum daily harvest per boat trip. This total salmon harvest was apportioned to each Chinook salmon substock based on (1) the known level of selectivity towards Chinook salmon and (2) the relative abundance of each substock s. That is, the species ratio of $H_{d,r,tot}$ was reduced from the true species ratio $\phi_{d,r}$ by a factor of 0.6 to obtain Chinook and chum/sockeye salmon harvest, then the Chinook salmon harvest was apportioned by the substock relative abundance. The maximum daily exploitation rate of any $A_{d,r,s}$ was capped at 0.9.

3.2.4 Simulated assessment data collection

The simulated assessment structure differed based on the management strategy used based on the richness of information required for each management strategy: *e.g.*, Strategy #1 (closed until open) required no information whatsoever whereas Strategy #4 (explicit harvest target) required a pre-season forecast, in-season abundance data, a method to update beliefs about run abundance, and weekly in-season harvest estimates. Only data sources that could be useful for in-season management were simulated, *e.g.*, because weir projects assessing escapement to specific tributaries are located so far from the bulk of the fishery they are not useful to determining in-season harvest opportunities.

3.2.4.1 Pre-season run size forecast

Pre-season forecasts of Chinook salmon total abundance were obtained as a bias-corrected lognormal random deviate from the true run:

$$\log\left(N_{tot,fcst}\right) \sim \mathrm{N}(\log(N_{tot}) - \frac{\sigma_F^2}{2}, \sigma_F^2) \tag{3.12}$$

where $\sigma_F^2 = 0.07$ which is the estimated variance ($\sigma_F = 0.27$) of historical forecast errors using the current forecast method (presented in Staton and Catalano 2019). When used in Strategies #2 and #3, only the point estimate of $N_{tot,fcst}$ was used to categorize the run as being a member of one of five discrete run size "bins", as displayed in Figure 3.1. When used in Strategy #4, the uncertainty in the forecast method was incorporated by treating the forecast as a bias-corrected lognormal probability density function (PDF), with variance equal to σ_F^2.

3.2.4.2 Test fishery index

A test fishery that produced daily catch per effort ($CPE_{TF,d,j}$) values for each salmon stock j ($n_j = 2$; one aggregate Chinook salmon stock and one aggregate chum/sockeye stock) was simulated in the first river reach and was assumed to index the run prior to any fishery harvest. The test fishery had an expected daily catchability (q_{TF}) and two sources of sampling variability: a catchability deviation representing annual fluctuations in river conditions and age/size composition of the incoming run (Flynn and Hilborn 2004) and daily fluctuations in fish vulnerability:

$$CPE_{TF,d,j} = A_{d,1,j}\frac{e^{q_{TF}+\varepsilon_{TF,y}+\gamma_{TF,d}}}{1+e^{q_{TF}+\varepsilon_{TF,y}+\gamma_{TF,d}}} \tag{3.13}$$

where $A_{d,1,j}$ is the total abundance of fish from species j each day in the first reach, and $\varepsilon_{TF,y}$ and $\gamma_{TF,d}$ are logit-scale sampling errors operating on the annual and daily time scales,

respectively. These sampling errors were normally distributed with standard deviations equal to $\sigma_\varepsilon = 0.15$ and $\sigma_\gamma = 0.2$ and q_{TF} was set to 0.004 – these settings resulted in simulated test fishery with similar properties as the Bethel Test Fishery. Daily species compositions were expressed as the ratio of chum+sockeye:Chinook salmon:

$$\phi_{TF,d} = \frac{CPE_{TF,d,CS}}{CPE_{TF,d,CH}} \tag{3.14}$$

where $j = CH$ for Chinook salmon and $j = CS$ for chum/sockeye salmon.

3.2.4.3 Bayesian updates of perceived run abundance

In assessed Strategy #4, the manager used in-season information regarding run abundance contained in the sampled values of $CPE_{TF,d,CH}$ to update the PDF provided by the pre-season forecast in a Bayesian framework. The analytical methods to perform this Bayesian update were identical to those presented in Staton and Catalano (2019), however, a brief description will be provided here. Based on a regression relationship between run size and cumulative test fishery CPE fitted to historical data of the form:

$$\log(N_{tot,y}) = \hat\beta_{0,d} + \hat\beta_{1,d} \sum_{k=1}^{d} CPE_{TF,k,CH,y} + \hat\varepsilon_{N,y,d}, \tag{3.15}$$

it is possible to predict total annual abundance on any day d of the season from the sum of all observed $CPE_{TF,CH}$ data through day d. Thirty historical years were simulated for fitting this historical relationship, which is highly variable for low values of d as a result of run timing and sampling variability, though becomes more informative as d increases and the run approaches completion. Uncertainty was propagated to predictions of abundance *via* Monte Carlo simulation of the regression parameters and residuals from their respective estimated sampling distributions as described in Staton and Catalano (2019). This process yields a daily distribution of likely run size outcomes according to the in-season data alone, and can

be viewed as new evidence with which to update prior information. The prior distribution each day was the PDF of the pre-season run forecast, and the PDF of abundance predictions from (3.15) was used as the likelihood to obtain the posterior PDF.

3.2.4.4 Weekly harvest estimates

In assessed Strategy #4, the manager had the ability to track in-season Chinook salmon harvest, such that progress toward attainment of the season-wide harvest target (H_T) could be monitored. Weekly harvest estimates were produced as random deviates from a symmetric truncated normal distribution with mean equal to the true weekly harvest, coefficient of variation (CV) equal to 15%, and lower and upper boundaries at 0 and 2 $\times$ true weekly harvest, respectively – these boundaries ensured unbiased and all positive harvest estimates. A CV of 15% was used because this is the approximate CV obtained using the in-season harvest estimation method developed and employed by Staton and Coggins (2016), Staton and Coggins (2017), and Staton (2018d). Cumulative estimated harvest was obtained by summing weekly estimates; uncertainty in harvest estimation was not considered. Estimates were created only for the villages within the Yukon Delta National Wildlife Refuge (YDNWR; reaches 1 – 9; Table A.4); this is a small enough area to be surveyed feasibly and it accounts for approximately 95% of all historical subsistence Chinook salmon harvest in the Kuskokwim River drainage (Hamazaki 2011).

3.2.5 Utility functions

Due to the lack of a common scale to the various objectives (Section 3.2.1), it was important to devise metrics than can be compared among objectives. These metrics are termed "utility functions", and here they are on the scale of [0,1], where zero indicates complete failure to meet an objective and one indicates complete success. Objectives can then be weighted based on their importance to different managers and an aggregate score can be obtained as a

weighted sum across the different utilities. Each objective received a unique utility function, as described below.

3.2.5.1 Attainment of aggregate escapement needs

Adequate escapement is the primary conservation objective in many salmon fisheries, and is necessary to ensure the stock can continue to produce adequate subsistence yields in the future. Thus, a rational metric to use is one based on the likely ability of the escaping spawning abundance to produce enough adult recruits to allow for attainment of subsistence harvest needs. The best scientific understanding of this ability is based in population dynamics of the stock, specifically the spawner-recruit dynamics. If the Ricker (1954) spawner-recruit model is to believed (as is often done in salmon population analyses, Fleischman *et al.* 2013, see Chapter 4, this dissertation as well), then there is a theoretical spawner abundance, termed S_{MAX}, that is most likely to produce maximum recruitment, termed R_{MAX}. R_{MAX} may be a more important metric for subsistence salmon fisheries than maximum sustained yield, given subsistence fishers tend to value consistently high abundance and catch rates over simply maximizing their long-term catch (Hamazaki *et al.* 2012). The utility function was a curve that represented the probability that a given escapement will produce 90% of R_{MAX} under equilibrium conditions, which would ensure high future catch rates and enough surplus of Chinook salmon to meet subsistence needs in the long-term.

To obtain this curve, termed a probability profile (Fleischman *et al.* 2013), the Bayesian state-space model presented in Hamazaki *et al.* (2012) was fitted to the Kuskokwim River aggregate population data over the period 1976 – 2017 using JAGS (Plummer 2017). This utility function assigned high utility (> 0.9) to escapements between approximately 70,000 – 125,000, with lower utility on either end outside of this range (Figure 3.4). One important consideration, however, is that if the Chinook salmon run is larger than approximately 230,000 fish, the subsistence fishery alone, which has historically harvested a maximum of

approximately 110,000 fish (Hamazaki 2011), cannot harvest enough fish to place escapement within that range. This fact is important when considering the value of this metric in very large runs.

3.2.5.2 Even substock exploitation rates

In the absence of any information regarding the productivity of the different Chinook salmon substocks within the Kuskokwim River drainage, the default preference would be that all substocks should receive the same exploitation rate ($U_s = \frac{H_s}{N_s}$; but see Chapter 4 for a study regarding methods used to obtain such estimates). Thus, a metric was needed that would have a high value (near one) if all Chinook salmon substocks had relatively equal U_s and that would provide a low value (near zero) if the U_s were vastly uneven. One such metric is the Schutz coefficient (Habib 2012; Schutz 1951), which is often used in econometrics to measure income inequality (*e.g.*, Kennedy *et al.* 1996). The Schutz coefficient takes the form:

$$z = \frac{\sum_i^n |x_i - \bar{x}|}{2 \sum_i^n x_i}, \tag{3.16}$$

where x_i is the income of earner i, $\bar{x}$ is the average income among all n earners, and z is the Schutz coefficient. In a technical sense, this index represents the fraction of the total income that would need to be redistributed reach perfect equity ($z = 0$), which has earned it an alternate name: the "Robin Hood Index." Here it is viewed simply as an index of evenness among substock-specific exploitation rates within a given year.

Several modifications were made to the Schutz coefficient in (3.16) for use in this utility metric. First, U_s was substituted for x_i and $n = 3$ to represent the three simulated Chinook salmon substocks. Second, given perfect equity (or evenness) of exploitation rates would be deemed a success, the complement of the Schutz coefficient was obtained for the utility function: $z' = 1 - z$. Third, the smallest value attainable for z' is n^{-1} but a complete failure

needed to be represented by zero to be consistent with the other utility functions. Thus, z' was normalized to be on the [0,1] scale:

$$z'' = \frac{z' - n^{-1}}{1 - n^{-1}} \qquad (3.17)$$

Finally, if all x_i elements are zero, z'' is undefined. In these cases, z'' was assigned the utility of one, given the U_s are even. Several examples of this utility function are presented in Table 3.3.

3.2.5.3 Attainment of aggregate subsistence needs

The Alaska Board of Fisheries has produced ANS ranges, which represent the range of salmon harvests by species that would reasonably be expected to meet subsistence salmon needs of fishers in the Kuskokwim River drainage (Appendix A.2.1). This range is 67,200 – 109,800, with a midpoint of 88,500. A "hockey-stick" utility function for drainage-wide Chinook salmon harvest was used that reached its maximum at one if harvest was above the midpoint of the ANS range, and a fraction of it ($H_{CH}/88{,}500$) otherwise.

3.2.5.4 Evenness of subsistence harvests

In addition to meeting the needs of the aggregate population of subsistence fishers, it is also generally desirable that Chinook salmon harvest be distributed evenly among the villages in each region (relative to their salmon needs). Thus, the same modified Schutz coefficient (z'') shown in Section 3.2.5.2 was used to quantify evenness of need-adjusted harvests (harvest/need) for villages located in the lower, middle, and upper regions of the Kuskokwim Drainage (Table A.4). In this case, high utility would be placed on outcomes in which a relatively equal fraction of Chinook salmon needs were harvested by villages in these regions.

3.2.5.5 Total Utility

The four objectives and utility metrics described above were collapsed into one measure that allowed quantification of overall performance and simple comparisons among strategies. This metric, termed total utility (V_T) was calculated as the weighted sum across each of the four objective-specific metrics:

$$V_T = V_S\omega_S + V_U\omega_U + V_H\omega_H + V_E\omega_E \tag{3.18}$$

where V_x and ω_x represent the utility measure and weighting factor for objective x, respectively (S = aggregate escapement, U = even U_s, H = aggregate harvest, E = equitable harvest). The default case assigned equal weight to each objective, but three alternate weighting schemes were assessed as well to determine the sensitivity of conclusions to this choice (Section 3.2.7.3).

3.2.6 Monte Carlo simulation

For each assessed strategy, $M = 5,000$ hypothetical runs were simulated with different total Chinook salmon run size, aggregate and substock-specific entry timings, substock compositions, and species compositions. Assessment errors were introduced randomly as well and each substrategy was tested on the Monte Carlo sample. The utility for each objective was calculated for each simulated year and strategy and was saved for summarization.

3.2.7 Summarization of management performance

Two levels of post-stratification of run types was conducted to facilitate inference. First, runs were stratified into five categories based on total Chinook salmon abundance (N_{tot}): [50K,80K], (80K,130K], (130K,180K], (180K,230K], and (230K,450K], and were the same as the bins used to categorize pre-season run size forecasts for Strategies #2 and #3 (shown

in Figure 3.1). These bins were selected roughly based on the level of needed management restrictions to ensure the subsistence fishery would not harvest an amount of fish that might damage escapement utility. Runs in the first two strata may require substantial restrictions, those in the third and fourth may require light or no restrictions, and the majority of runs in the fifth strata should require no management whatsoever to ensure near full attainment of the escapement and harvest objectives. Second, run timing was stratified into three categories: >3 days early, >3 days late, and all runs. The average utility value across Monte Carlo samples for each strategy/substrategy was calculated for each objective in each run size/timing stratum.

3.2.7.1 Within-strategy comparisons

Substrategies within each of the four primary strategies were compared at each run size stratum for each utility metric. The effect of run timing variability was assessed by qualitatively evaluating which substrategies had largely different outcomes for either the "early" or "late" strata than the "all" stratum.

3.2.7.2 Among-strategy comparisons

The best-performing substrategy in each run size stratum across all run timing strata according to the total utility measure (V_T) was extracted and its performance was compared to that of other strategies. In selecting the best substrategy, it often occurred that negligible differences were found among substrategies according to V_T: in these cases of a "tie" (defined as one where the second best substrategy was within 5% of the best) the substrategy that performed best with respect to escapement utility was selected for comparison; if that was again a tie, then utility measures from all substrategies included in the tie were averaged and noted as a "hybrid" substrategy.

80

3.2.7.3 Evaluation of sensitivity to weighting schemes

The default case was to weight all four metrics equally when obtaining total value ($\omega_S = \omega_H$ = $\omega_E = \omega_U = 1$), but three other weighting schemes were used for sensitivity analyses:

- *Simple-view*: $\omega_S = 1$; $\omega_H = 1$; $\omega_E = 0$; $\omega_U = 0$

- *Escapement-oriented*: $\omega_S = 1$; $\omega_H = 0.5$; $\omega_E = 0.25$; $\omega_U = 0.75$

- *Harvest-oriented*: $\omega_S = 0.5$; $\omega_H = 1$; $\omega_E = 0.75$; $\omega_U = 0.25$

The "simple-view" is intended to focus only on the two primary objectives of salmon management, and the escapement- *versus* harvest-oriented scenarios are opposites of each other, with the aggregate objective (V_S or V_H) in each case carrying the most weight followed by the spatial distribution objectives (V_U or V_E).

In calculating the total utility (V_T; the only basis for comparison here), it was important to restandardize it for comparisons among weighting schemes (ω_x). This is because these different combinations have differing maximally attainable V_T. For example, the maximum attainable V_T for the "simple-view" case is two, but it is four for the default case. For this comparison, the different V_T values for each weighting scheme were rescaled to be a fraction of the maximally attainable V_T for that weighting scheme, which is equal to $\sum_x \omega_x$.

3.3 Results

3.3.1 Operating model realism

The operating model was found to adequately capture the important dynamics of the fishery when left unrestricted with respect to total harvest magnitude as well as spatiotemporal patterns in the distribution of Chinook salmon harvest at a range of all simulated run sizes, timings, and species/stock compositions (Appendix B). Some amount of fine-tuning was required of the catchability parameter (q) and the effort response model coefficients (β_n) to reproduce these patterns, however no behaviors arose that seemed highly questionable. In

general, the level of simulated inter-annual variability was similar to that observed in the historical data (Appendix B). Based on these findings, inference regarding policy performance proceeded under the assumption that the operating model reasonably represented the system dynamics.

3.3.2 Within-strategy comparisons

3.3.2.1 Strategy #1: "Closed until open"

Strong patterns were found in the relative performance of the different substrategies of assessed Strategy #1 (Figure 3.5), particularly with respect to the expected utility for the aggregate harvest (V_H) and escapement (V_S) objectives. In the "smallest" simulated runs (< 80,000), only the June 23 substrategy resulted in any measurable amount of escapement utility ($V_S \approx 0.2$), the other two assessed earlier dates resulted in V_S near 0. This finding in the smallest runs was not at all sensitive to the timing with which simulated Chinook salmon entered the river (as indicated by the overlap in the three lines, Figure 3.5). In "small" simulated runs (80,000 – 130,000), more escapement utility was attained for each substrategy, but the declining pattern remained where earlier fishing would result in lower V_S. In these runs, however, run timing variability did greatly impact the ability to meet escapement needs: late runs had the tendency to result in higher V_S even when the river was opened completely beginning on June 1. V_S was generally highest in the "medium-sized" runs (130,000 – 180,000), with all three substrategies resulting in $V_S \geq 0.8$ and little sensitivity to run timing. The June 23 substrategy resulted in the lowest V_S in "large" runs between 180,000 – 230,000, as a result of allowing many Chinook salmon to escape; the case was the same for all substrategies of the "largest" runs (> 230,000), but in these runs there is no management action that could allow the subsistence fishery to harvest enough fish to obtain high escapement utility according to the function used (Figure 3.4). Greater V_H was obtained with earlier opening dates, as would be expected given Chinook salmon abundance becomes

overwhelmed by chum and sockeye salmon in the later part of June. These results highlight a trade-off between harvest and escapement in runs smaller than 130,000: earlier fishing resulted in more harvest, but less escapement utility.

Harvest equity (V_E) was maximized at the intermediate substrategy (June 12) for small runs, but reached its maximum with the earliest date with larger runs. The utility resulting from equal exploitation rates (V_U) was flat over the continuum of assessed start dates, however there was a slight trend for later fishing dates to have higher values of V_U. Run timing influenced the value of V_U as well, with earlier runs generally having greater utility. The default total utility metric (V_T; obtained with all weights $\omega_x = 1$) was roughly equal among substrategies in small runs (indicating that each balanced the trade-offs differently), whereas V_T was greatest for the intermediate and earliest start dates in larger runs (*i.e.*, more aggressive start dates).

Given the relatively small changes in the V_E and V_U metrics among substrategies, other patterns with the raw output of exploitation rate by substock (U_s; Figure 3.6) and the fraction of salmon needs that were met (Figure 3.7) were investigated. With respect to U_s, the most noticeable difference among substrategies was that the exploitation rate for all substocks was lower for the late opening dates than for early opening dates (Figure 3.6). Due to the limiting nature of the subsistence fishery, the exploitation rates declined with increasing run sizes. In all substrategies, the exploitation rate of the upper river substock was greater than for the lower and middle river substocks, however this difference declined as the opening date was delayed. Regarding the evenness of attainment of harvest needs among villages in different regions of the drainage, the greatest unevenness was found for large runs combined with the June 23 substrategy, in which upper river fishers often exceeded their minimal needs but lower river fishers obtained less than half of theirs. Overall, the changes in these raw output values seemed more substantial than what was indicated by the use of the modified Schutz coefficient, as shown in Figure 3.5.

3.3.2.2   Strategy #2: "Forecast-based fixed schedule"

Just as in assessed Strategy #1, the expected utilities for the aggregate harvest (V_H) and escapement (V_S) objectives were those most influenced by the choice of substrategy of assessed Strategy #2. The conservative substrategy resulted in higher V_S in small runs, but at the cost of lower V_H (Figure 3.8). In large runs, however, there was much less contrast in substrategy performance. Run timing variability again played a key role in determining management success, but more so in large runs than in small runs. It was only in small runs that harvest equity (V_E) or harvest rate evenness (V_U) were sensitive to the selection of substrategy – in medium and the large run scenarios these metrics were essentially equal along the continuum of conservative to aggressive fishing schedules. In general, Strategy #2 was also influenced by run timing variability (though less so than Strategy #1), and particularly in larger run sizes. V_T was largely the same among substrategies, with a slight tendency to favor more aggressive schedules in nearly all run size categories.

3.3.2.3   Strategy #3: "Forecast/ratio-based variable schedule"

Substrategies of assessed Strategy #3 (Figure 3.9) showed high similarity to the patterns in Strategy #2. The only difference of note between these two strategies was the difference in utility between substrategies was smaller for Strategy #3 than for Strategy #2 (*i.e.*, overall shallower slopes in Figure 3.9 than in Figure 3.8).

3.3.2.4   Strategy #4: "Explicit harvest target"

The choice of the particular substrategy used for Strategy #4 had less of an impact on escapement or harvest utility than substrategies of Strategies #1-3 (Figure 3.10) in small runs. This indicates that the performance of this strategy in small runs was insensitive to the particular harvest table used (*i.e.*, the linkage between the weekly harvest target and number of fishing days, Figure 3.3). This is likely a result of the probabilistic choice of a

harvest target – because this method accounted for uncertainty in run abundance and risk in failing to meet the escapement limit, the harvest target was probably low enough in these small runs to where it did not matter which substrategy was used, they would all suggest very few fishing days per week. In larger runs, where the harvest target was larger, more contrast was found in substrategy performance with respect to harvest and escapement. Run timing variability affected the performance of this strategy, and as in other strategies the effect was strongest at large run sizes. V_E and V_U were generally unaffected by the choice of substrategy.

3.3.3   Among-strategy comparisons

After extracting the best substrategy from each of the four primary strategies at each run size (across all run timing scenarios), it was clear that conservative/neutral substrategies were favored in small runs and aggressive substrategies in large runs (Figure 3.11). Ties among substrategies were more common in larger runs, indicating the details of strategy implementation had less influence in these cases. The largest differences in management performance among strategies were with respect to aggregate escapement and harvest in small and intermediate sized runs – the harvest equity and evenness of exploitation rate metrics were largely insensitive to the selection of strategy at all run sizes. In the smallest runs, Strategy #4 was strongly favored over other strategies with respect to escapement, likely as a result of its inherent risk aversion built into the probabilistic selection of the harvest target. However, Strategy #4 tended to result in less harvest utility in nearly all run sizes than the other strategies, indicating that more complexity in the decision rules still leaves room for management mistakes, but that they err on the side caution. Within a run size category, there was a high degree of similarity in total utility among strategies, though a weak pattern emerged that favored more complex strategies (#4) in small runs and simpler strategies (#1-3) in larger run scenarios.

3.3.4   Sensitivity to weighting schemes

It is important to consider how these findings regarding total utility might depend on how the various utility functions were weighted. The major pattern that arose was that when the weights were adjusted to the simple-view or escapement-oriented scenarios, the tendency to favor conservative substrategies in small runs was more apparent – the "harvest-oriented" weighting scenario favored either neutral or aggressive substrategies in even the smallest run sizes (Figure 3.12). The pattern of high similarity in overall performance among strategies remained, but there was a tendency to favor Strategy #4 (the most complex) more in the escapement-oriented weighting scenario than in the harvest-oriented weighting scenario (Figure 3.12). According to the "simple-view" weighting scenario, relative performance in the smallest runs was much lower in comparison to other run sizes than using other weighting schemes (Figure 3.12). This is because only the aggregate harvest and escapement objectives were considered in the simple view case (and both objectives score low in these smallest runs), whereas other weighting scenarios included the spatial distribution of these quantities in measuring management performance.

3.4   Discussion

The dominant trade-off I found, not surprisingly, was between harvest and escapement in small runs ($< 130{,}000$). These runs do not have enough fish to allow for both high escapement and harvest utility, and given every fish that is harvested cannot also escape, it is clear as to why this is the case. The trade-off was identified because for most strategies, the more conservative substrategies tended to have higher escapement utility and lower harvest utility, and *vice versa.* In larger runs, this trade-off was not present, given enough fish were available for both objectives.

One of the more surprising findings in my view was the high degree of similarity in total utility (V_T) among management strategies (after filtering the best-performing substrategy). As an example, I expected that the explicit harvest target approach in Strategy #4 would strongly outperform the fixed schedule strategy (Strategy #2) because of its timely response to information. I see two plausible explanations for why the more complex strategies did not perform overwhelmingly better in most cases. First, it is possible that the harvest table approach was too simple in Strategy #4. It is likely that managers may adapt the table based on run abundance or species composition. A more involved approach still would be to select fishing duration each week (D_w) based on an explicit prediction of how many fish would be captured conditional on each value of D_w under consideration. The candidate that results in predicted weekly harvest nearest to the desired weekly harvest ($H_{T,w}$) would then be selected. Understanding of the fishery dynamics at various run sizes would be required to trust these predictions strongly, but recent studies (Staton and Coggins 2016, 2017; Staton 2018d) have gone a long way towards providing this understanding for runs in the small category. These predictions can be made very simply as the product of three anticipated quantities and one policy variable: total boats/day $\times$ salmon catch/boat/day $\times$ % Chinook in catch $\times D_w$. A second explanation is that the additional dexterity gained by a more complex strategy is only as good as the information it relies on, and possibly the simulated data sets were too weak to implement Strategy #4 well. I attempted to mimic the properties of the data sources collected for in-season management in the Kuskokwim River; it is possible that more precise run assessment methods (*e.g.*, sonar) would provide better information to implement this policy.

There are relatively few studies in the literature similar to the one presented here to allow comparison. Carney and Adkison (2014a) and Carney and Adkison (2014b) conducted analyses comparing fixed schedules and feedback strategies (referred to as "daily management" or "management by emergency order" therein). In both cases, they found the more complex

87

feedback strategy did increase average annual catch without putting escapement at risk, but also resulted in more inter-annual variability in catch than the fixed schedule strategy. Inter-annual variability was not investigated in my analysis, but I found that the feedback strategies (those that used information collected in-season; Strategies #3-4) tended to result in the same or less harvest in most run sizes compared to the simpler fixed-schedule strategies (those that used no or only pre-season information; Strategies #1-2; Figure 3.11). Additionally, Carney and Adkison (2014a) stated that a fixed schedule should perform better at spreading exploitation among stock components, but my analysis did not support this claim: I found that all strategies had highly similar performance with respect to evenness of exploitation rates. This was probably a result of the multi-species component: because chum/sockeye enter later than Chinook salmon, their limiting nature on Chinook salmon harvest only takes effect on the later part of the run.

Examination of substrategy performance revealed that often the choice regarding the best depended on run size: more neutral or conservative substrategies were selected in small runs ($< 130{,}000$) and more aggressive substrategies in larger runs ($> 180{,}000$). This finding simply suggests that, regardless of the particular strategy being employed, it should not be implemented the exact same each year. Fishing schedules must be updated to adequately target which outcomes are likely to influence success that year. For example, in the smallest runs, it is possible to obtain moderately high escapement utility ($V_S = 0.7$ with Strategy #4) with very little fishing activity. However, the most harvest utility possible in these runs is low ($V_H = 0.4$ with Strategy #1, June 1) and if this were enacted V_S would be 0. Clearly the more important objective in small runs such as these is escapement, so managers should adapt the strategy to behave in a more conservative way. The situation reverses in intermediate and large sized runs, where it is possible that both V_H and V_S benefit from more aggressive fishing (within reason). This finding makes intuitive sense, and to most salmon managers it was almost certainly known *a priori* to this analysis. However, this analysis (and ones like it)

are useful in helping define these transition points and defining what a set of conservative *versus* aggressive schedules might look like.

In terms of robustness to variability in run timing, I found that all strategies were sensitive in large runs, but that only Strategy #1 was sensitive in small runs. This makes intuitive sense given Strategy #1 used only one "decision", and that was the day to open the fishery completely. As a result, an early opening date coupled with an early run is likely to produce more Chinook salmon harvest than the same opening and a late run because of the timing with which chum and sockeye salmon begin to dominate the species composition. If the Chinook salmon run shows up early, then a larger fraction of their run is vulnerable to lower river fishers before chum and sockeye salmon enter in large numbers and trigger the processing capacity limit, and when coupled with an early opening date would result in high harvest. Conversely, if the Chinook run is early but coupled with a late opening, then Chinook harvest is likely to be low because much of the run has passed the lower river harvest areas before much fishing activity occurs. These dynamics can be easily understood because of the simple nature of Strategy #1. In the more complex strategies, more factors influenced the number of fishing days per week than simply the time of the season. For example, schedules for Strategies #2-3 were explicitly chosen to have fewer days early in the season than later in the season to prevent catching too many Chinook salmon when the abundance is high relative to other species. Furthermore, sampling variability was introduced into the decision-making process that could also serve to swamp the influence of run timing variability. Increased sensitivity at large run sizes was likely a result of the fact that more is at stake in large runs from a harvest perspective: there is more surplus in these years, and a proportional reduction in a large harvest affects the V_H utility function more than the same proportional reduction in a small harvest.

The final primary finding was that the weighting of objectives in the total utility function did influence the inference, but only regarding which substrategies (not strategies) were

best and only in small runs. The escapement-oriented weighting scheme suggested that conservative substrategies performed better in these small runs. Weighting did not, however, change the inference that the different strategies performed similarly within a run size category. This indicates that perhaps managers with different inherent weights on their objectives should not necessarily change their decision rules entirely from other managers, but instead that they may just fine-tune the details of their implementation (*e.g.*, specific trigger points).

The approach I used did have some weaknesses. First, inference regarding strategy performance conditioned on the particular objectives selected for this purpose, and more specifically, on the utility functions used to measure the degree of their attainment for each hypothetical salmon run. Each manager/stakeholder may have different objectives (or different ways to weight them), but the performance metrics were built around the four dominant themes discussed in management and stakeholder meetings regarding management objectives and I assessed the sensitivity to different weighting schemes (which might represent different managers or stakeholders). Second, this simulation analysis required that the management strategy be expressed as a rigorous control rule, where the decision would be made the same way each time the same information was available. In reality, managers do not operate this way – one control rule cannot simultaneously consider all sources of information in such a programmatic way. This fact limits the realism of the analysis, but is not unique to these kinds of MSE analyses.

A population dynamics submodel (*i.e.*, one including reproduction of spawners resulting in future run sizes; outputs organized as a time series) was not incorporated into the operating model primarily because I deemed it unnecessary to evaluate performance of in-season management strategies. Incorporation of a population dynamics submodel would move the analysis away from its focus on in-season strategies to long-term harvest control rules and the optimal level and spatial distribution of escapement. While these issues are important to address, I wanted this analysis to focus on the different ways a manager could behave

in-season to meet a set of objectives they care about meeting in any given year (irrespective of population dynamics). This focus assumes that attainment of objectives specified here is indicative of good long-term performance. In contrast, if the focus of the analysis was on the long term performance of the objectives themselves (*e.g.*, finding the best escapement goal), then a multi-year simulation approach with embedded population dynamics would be necessary. This type of approach would allow for management mistakes or successes in any given year to propagate to future years which, while potentially interesting, was beyond the scope of my study.

The primary characteristic that sets my analysis apart from other salmon MSE analyses is that it focused on in-season decision rules for subsistence fisheries, which behave quite differently than commercial fisheries such as the ones modeled by Carney and Adkison (2014a) and Carney and Adkison (2014b). As a result, I was able to assess strategies that acknowledge the characteristics of subsistence fisheries when considering fishing schedules: namely declining effort with attainment of harvest needs and the self-limiting nature resulting from processing capacity. To my knowledge, strategies that acknowledge these characteristics have not been assessed using stochastic methods such as the one I used, making this work a novel contribution to the body of knowledge regarding in-season salmon management. Managers and stakeholders in the region may find the results informative as an objective evaluation of management strategy performance – a key point of interest here will be that there are several suitable strategies that could be implemented. At the very least it should serve to illustrate the concepts of how the in-season component could be modeled should a more-engaged participatory strategy evaluation process aimed at long-term performance is desired in the future.

TABLE 3.1: Species ratio (chum+sockeye:Chinook salmon) trigger point cut-offs used in Strategy #3. $\phi_{p,w-1}$ is the percentile of the average daily species ratio detected in the previous week at the simulated test fishery site when taken in context of all historical species ratios for week $w-1$. Different substrategies are shown in the three columns: different thresholds that indicate when the manager should switch from using different schedules (as shown in Figure 3.1). For example, the neutral version of Strategy #3 would employ the conservative schedules following the pre-season forecast as shown in Figure 3.1 until the species ratio exceeds the 33% percentile of all historically observed ratios.

| | Species Ratio Thresholds for Substrategies of #3 | | |
Schedule	Conservative	Neutral	Aggressive
Conservative	$\phi_{p,w-1} \leq 66\%$	$\phi_{p,w-1} \leq 33\%$	$\phi_{p,w-1} \leq 15\%$
Neutral	$66\% < \phi_{p,w-1} \leq 85\%$	$33\% < \phi_{p,w-1} \leq 66\%$	$15\% < \phi_{p,w-1} \leq 33\%$
Aggressive	$\phi_{p,w-1} > 85\%$	$\phi_{p,w-1} > 66\%$	$\phi_{p,w-1} > 33\%$

TABLE 3.2: Specific species ratio trigger points used in assessed Strategy #3 when selecting which schedule type (Figure 3.1) to employ. For example, a manager in week 3 using the conservative substrategy would use the conservative schedule unless the average species ratio in the previous week was above 1.4, at which point they would switch to the neutral schedule. Date ranges belonging to each week are shown in Figure 3.1. These trigger points were obtained from the cut-off rules shown in Table 3.1.

	Week				
Ratio Trigger Substrategy	**1**	**2**	**3**	**4**	**5**
Trigger Switch from Conservative to Neutral Schedules					
Conservative	0.7	2.4	1.4	4.6	17.1
Neutral	0.2	0.7	0.4	2.6	11.0
Aggressive	0.1	0.1	0.0	1.3	9.1
Trigger Switch from Neutral to Aggressive Schedules					
Conservative	1.2	3.1	1.8	6.5	26.2
Neutral	0.7	2.4	1.4	4.6	17.1
Aggressive	0.2	0.7	0.4	2.6	11.0

TABLE 3.3: Example values of the modified Schutz coefficient (z''; Section 3.2.5.2) used as the utility function for the objectives dealing with evenness of exploitation rates and harvest equity. Examples are in decreasing order of equity, with the top rows representing more equitable/even cases than those at the bottom of the table. When used for evenness of exploitation rates, the x_{Region} represent substock-specific exploitation rates (U_s). When used to measure equity, the x_{Region} represent the fraction of needed Chinook salmon harvested by villages within reaches located in each region.

x_{Lower}	x_{Middle}	x_{Upper}	z''
100%	100%	100%	**1**
10%	10%	10%	**1**
40%	40%	50%	**0.92**
40%	40%	70%	**0.8**
30%	30%	90%	**0.6**
15%	30%	90%	**0.5**
0%	10%	90%	**0.15**
0%	0%	10%	**0**
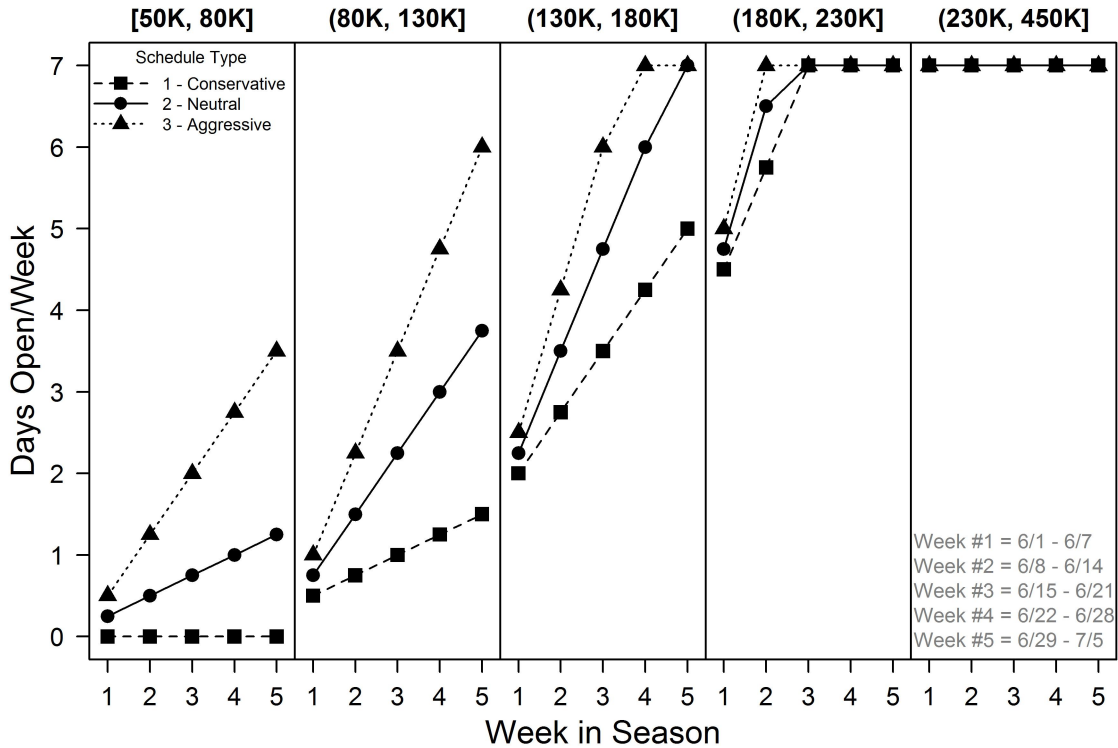
FIGURE 3.1: Decision rules for setting fishing schedules in assessed Strategies #2 and #3. The number of days the fishery is to be opened per week is a function of the pre-season forecast, as shown by each of the five panels. The three lines in each panel represent the different substrategies of Strategy #2 or schedule types for Strategy #3. In Strategy #3, the manager would select to be conservative, neutral, or aggressive based on the percentile of recently observed species ratios, as indicated in Table 3.1. In other words, the manager using Strategy #3 could adapt fishing schedules to in-season conditions, where as the #2 manager could not.
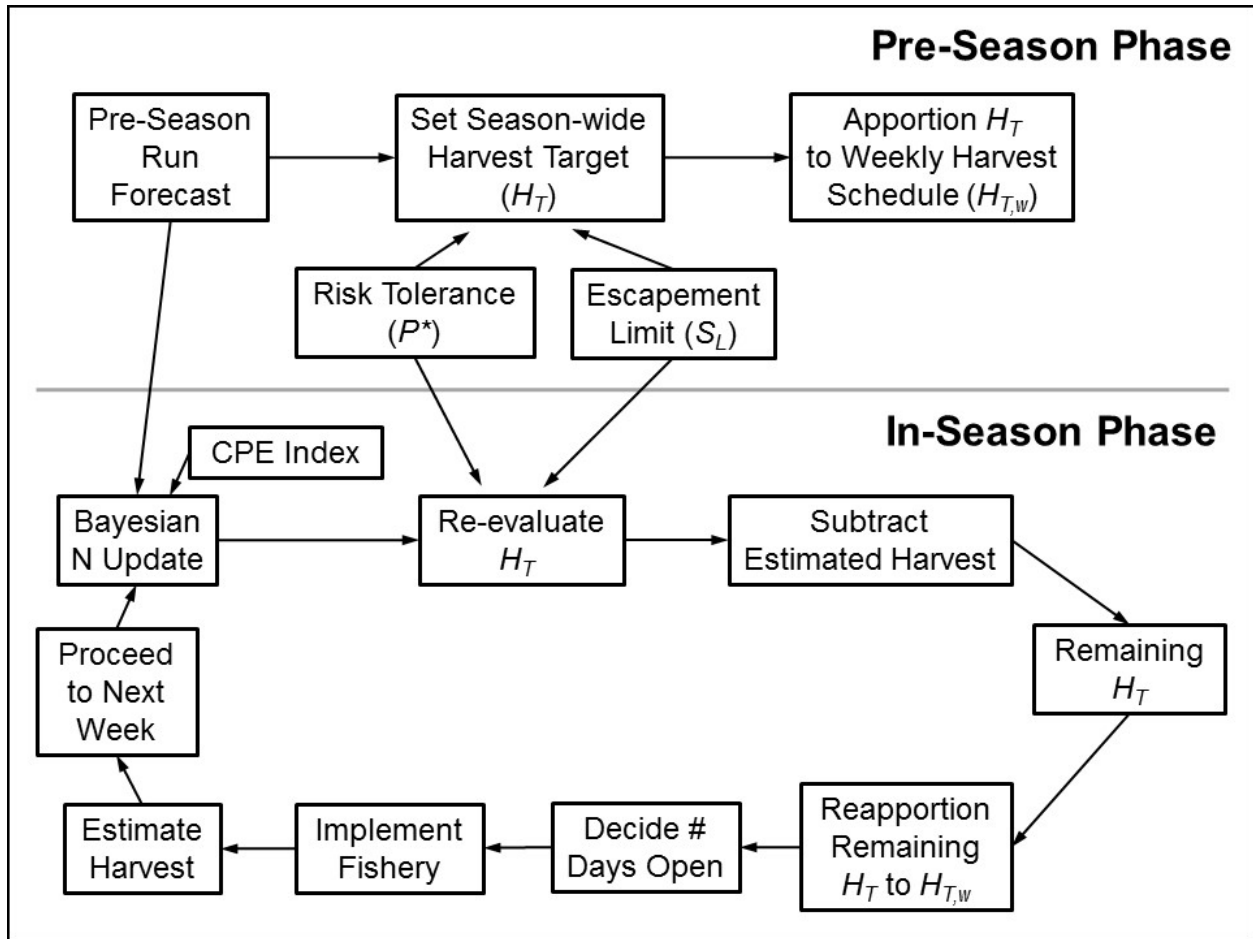
FIGURE 3.2: Depiction of the use of information to guide decision-making in assessed Strategy #4, partitioned into pre-season and in-season phases. All actions are taken with regards to Chinook salmon. **Pre-season actions** occur only once per season, and involve producing a pre-season forecast (with error) and using it to set a season-wide harvest target (H_T) based on (*a*) the probability distribution representing uncertainty in the pre-season forecast, (*b*) a limit point that escapement should not fall below (S_L), and (*c*) the maximal acceptable probability for seeing the outcome $S < S_L$ (P^*). Targeted harvest by week ($H_{T,w}$) is initially set by apportioning the total among weeks according to a fixed schedule based on historical run timing data. **In-season actions** are represented by a weekly cycle that involves updating perceptions of abundance and adapting the season-wide harvest target H_T as appropriate to ensure the current posterior probability of attaining at least S_L given H_T still conforms with P^*, and the remaining allowable harvest for the season is obtained *via* subtracting cumulative estimated harvest already taken. Remaining harvest is then apportioned to the remaining weeks, and based on the value of $H_{T,w}$, the fishery will be opened for between zero and seven days for the week according to the harvest tables displayed in Figure 3.3. Harvest outcomes are monitored such that a weekly harvest estimate is available for use in the next week, which begins with obtaining a new posterior understanding of total run abundance.
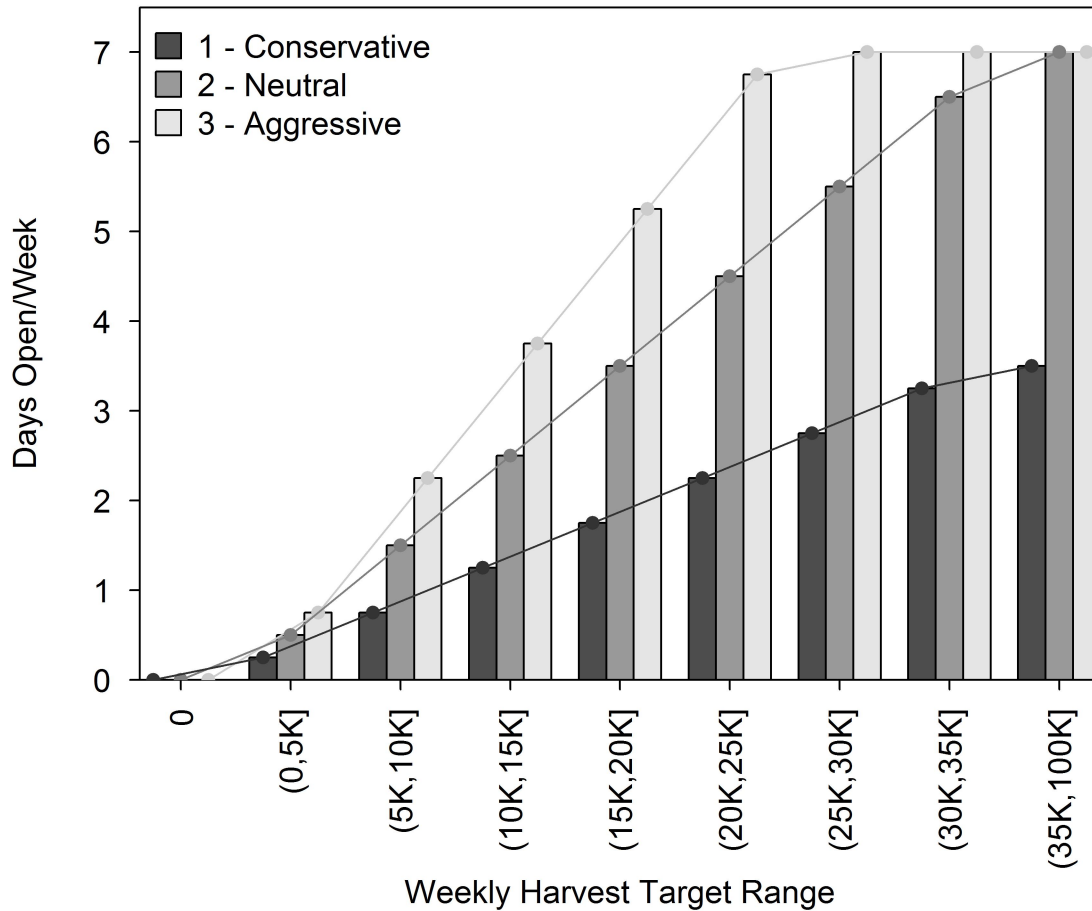
FIGURE 3.3: The "harvest tables" used in assessed Strategy #4. Based on how many fish are targeted each particular week ($H_{T,w}$), the manager would select the number of days to open the fishery. The process to obtain $H_{T,w}$ was rather involved, requiring pre-season forecasts, in-season abundance index data, and in-season harvest data to inform its value, as shown in Figure 3.2.
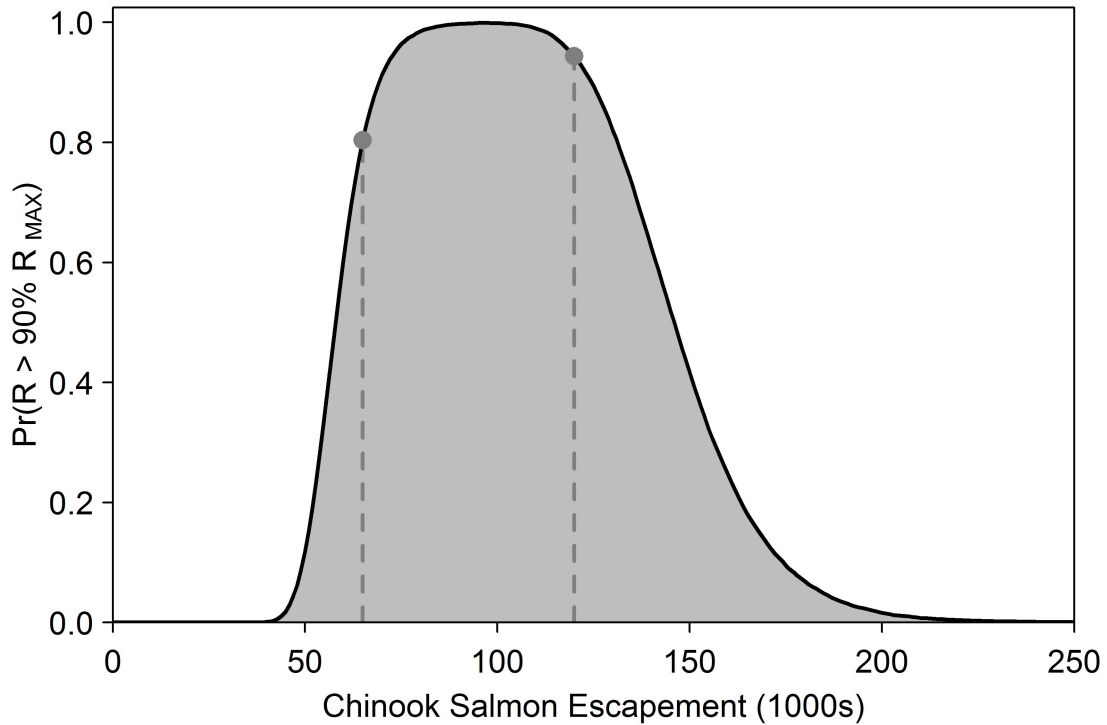
FIGURE 3.4: Estimated probability profile for Kuskokwim River Chinook salmon used as the utility function for drainage-wide escapement (V_S) in this analysis. The height of the curve represents the currently understood probability that expected recruitment produced by a given escapement level will exceed 90% of R_{MAX}, and was obtained for the aggregate Chinook salmon stock using the Bayesian state-space estimation model presented in Hamazaki *et al.* (2012) updated with abundance, harvest, and age composition data through 2017. The vertical dashed lines are the endpoints of the current escapement goal range: 65,000 – 120,000.
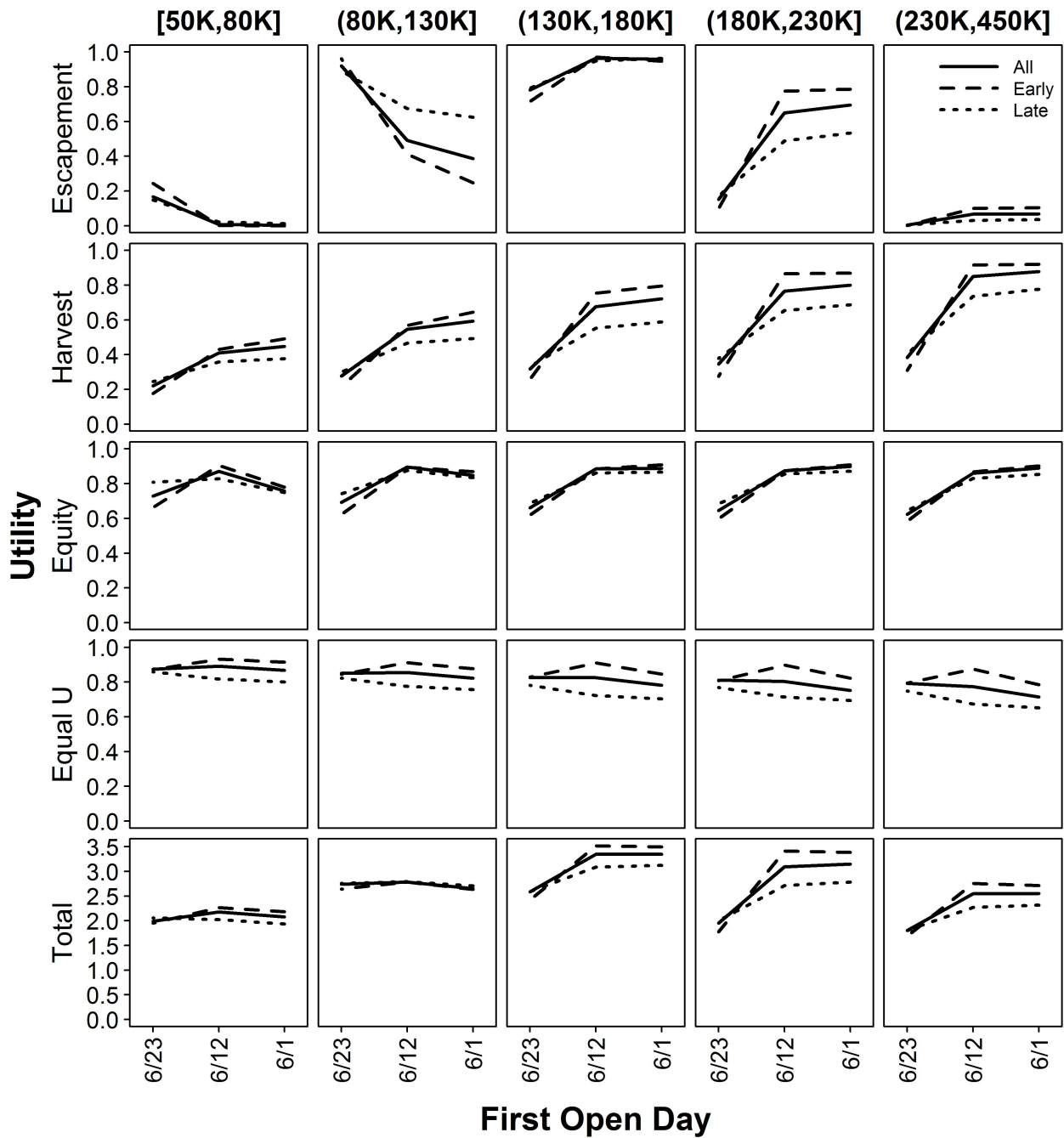
FIGURE 3.5: Detailed performance of assessed Strategy #1. Values of the utility functions (rows) separated by run size category (horizontal panels), run timing category (line type), and substrategy (x-axis, ordered from most conservative to aggressive). Substrategies of this policy differ in the date at which the fishery is opened completely. The form of each utility function is described in Section 3.2.5, and the total metric shown uses the default weighting scheme (all objective weights equal to 1).
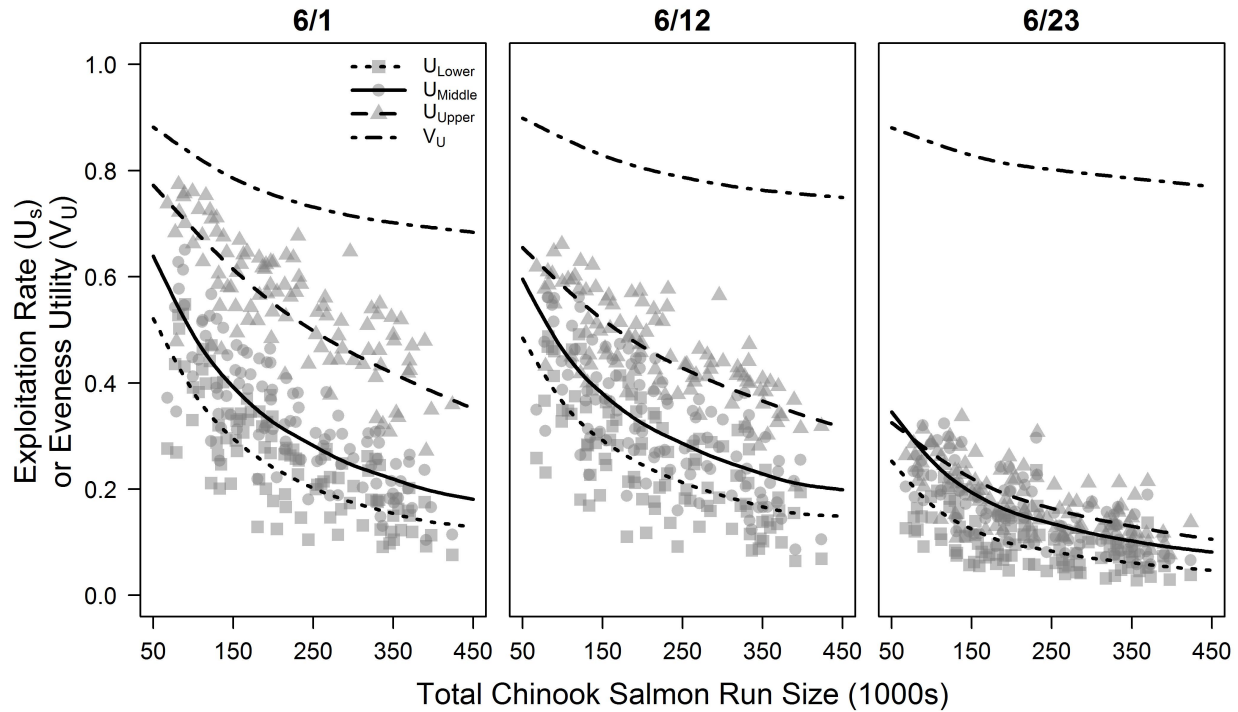
FIGURE 3.6: Chinook salmon substock-specific exploitation rates as a function of run size from 500 Monte Carlo trials, separated by different substrategies (*i.e.*, opening dates; panels) of assessed Strategy #1. Lines are fitted generalized additive models. The line denoted by V_U represents the model fitted to the utility metric as defined by the modified Schutz coefficient used (points not shown).
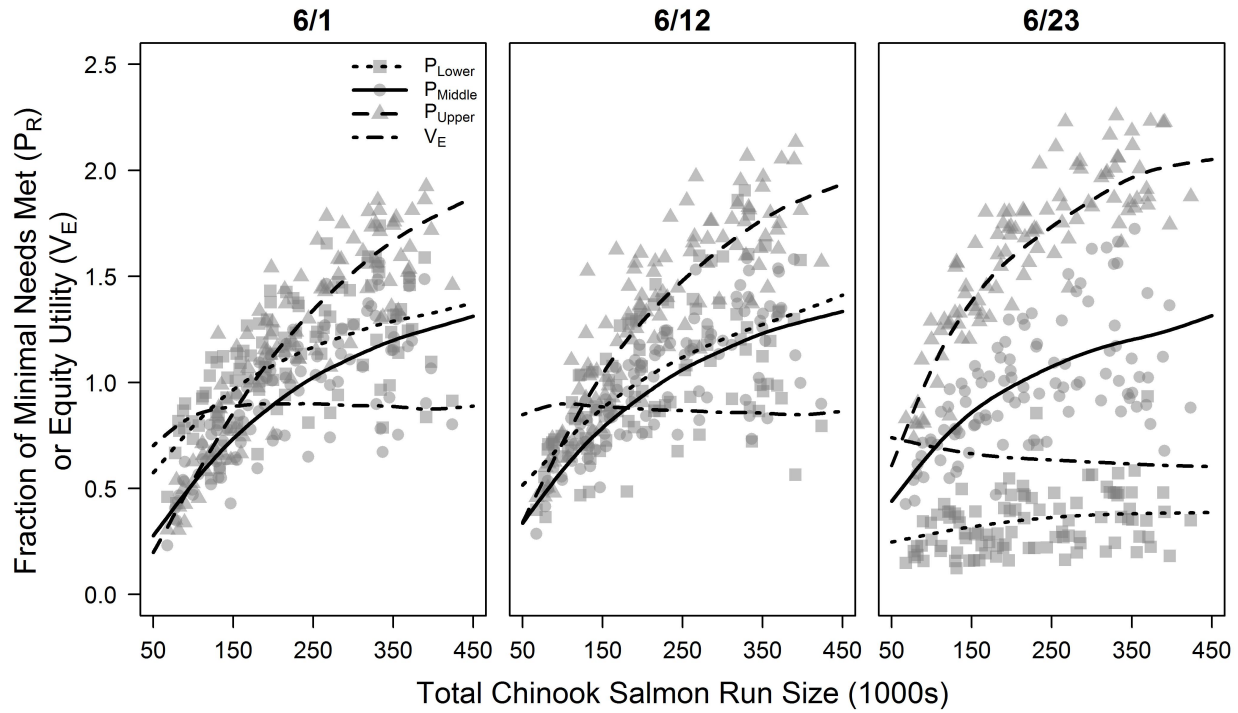
FIGURE 3.7: The fraction of minimal Chinook salmon harvest attained by villages in the lower, middle, and upper regions of the simulated Kuskokwim River as a function of run size from 500 Monte Carlo trials, separated by different substrategies (*i.e.*, opening dates; panels) of assessed Strategy #1. Lines are fitted generalized additive models. The line denoted by V_E represents the model fitted to the utility metric as defined by the modified Schutz coefficient used (points not shown).
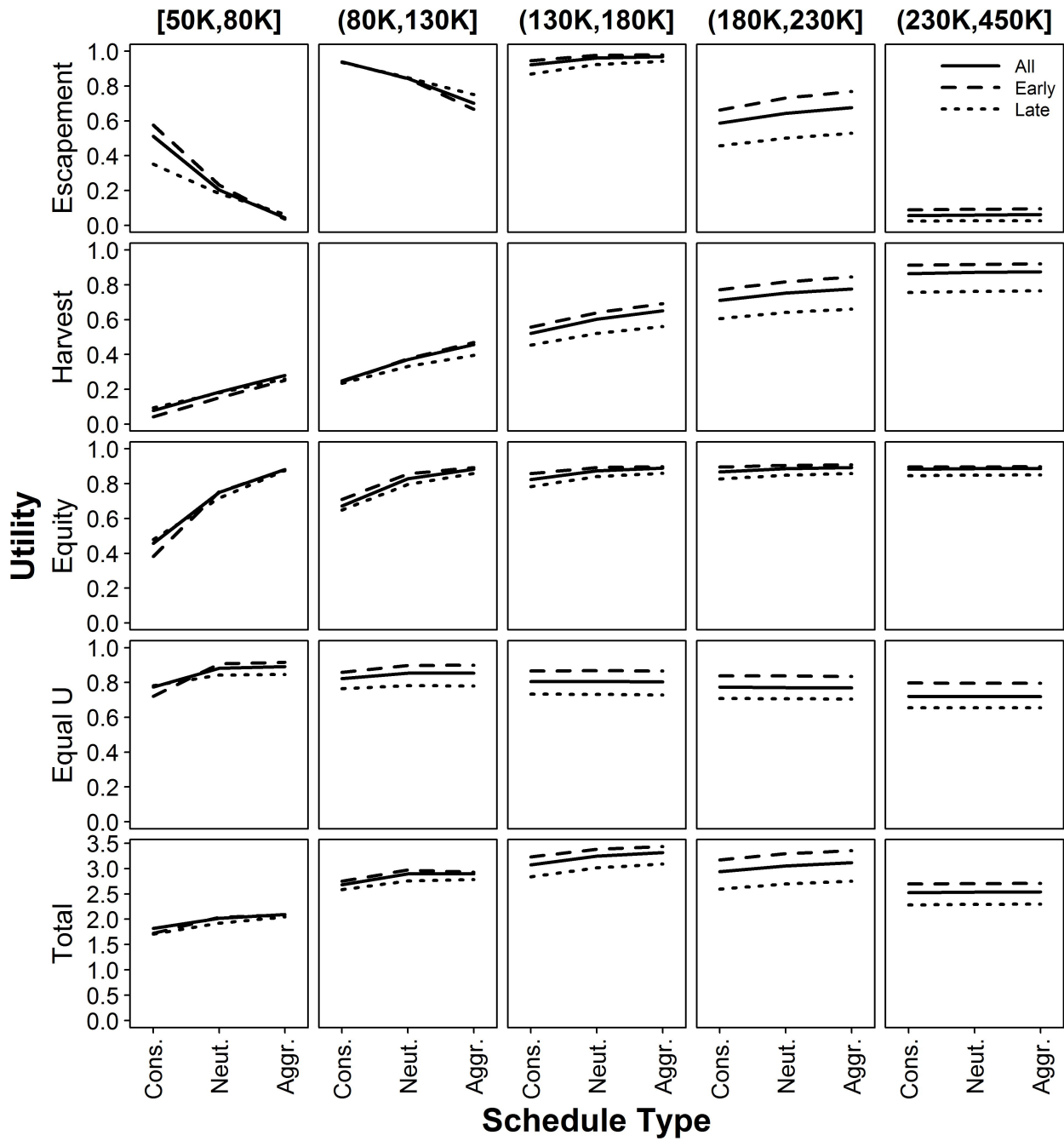
FIGURE 3.8: Detailed performance of assessed Strategy #2. The layout of panels in this figure is the same as in Figure 3.5, only substrategies represent different schedules conditional on a pre-season run size forecast (schedules shown in Figure 3.1).
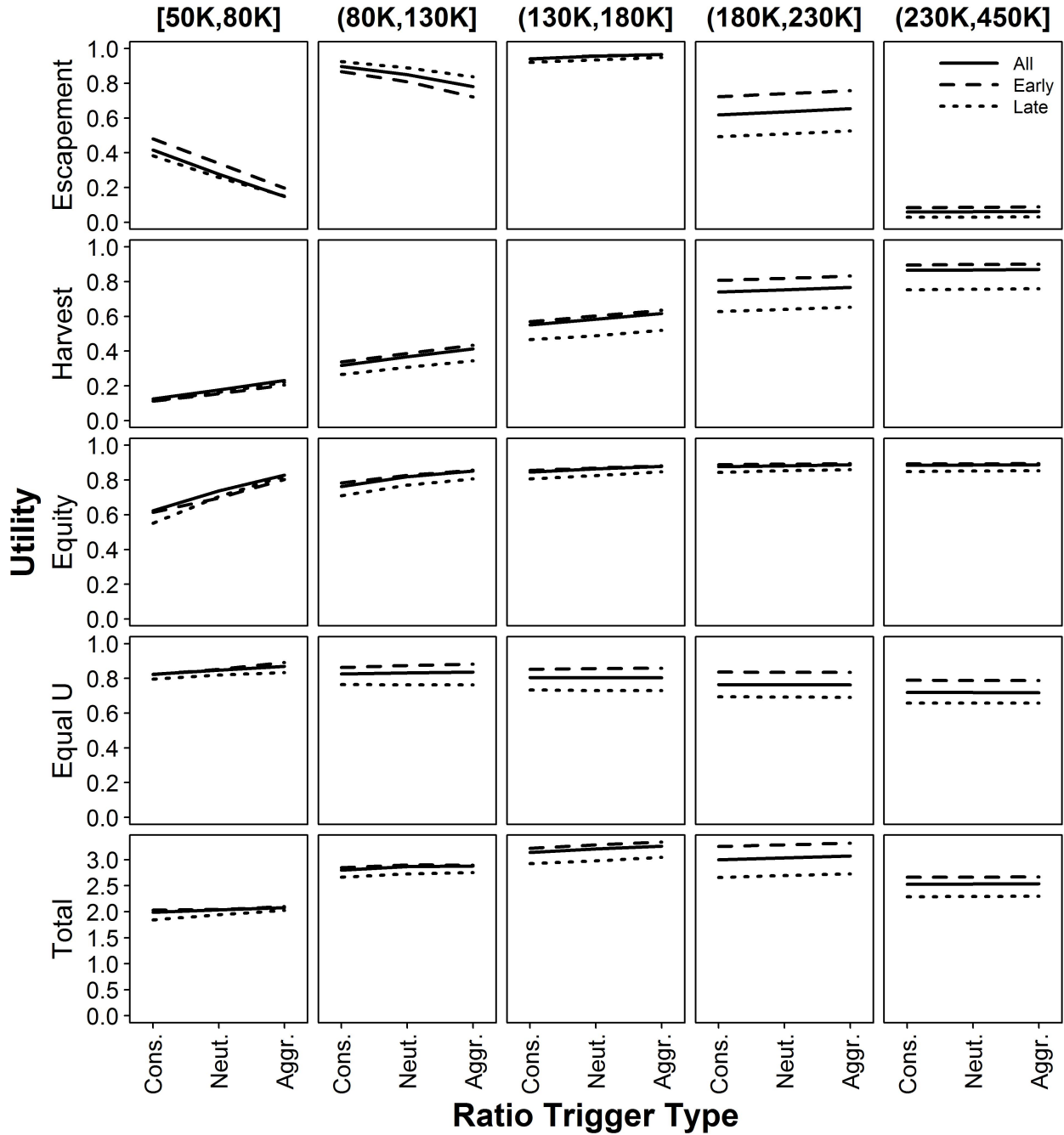
FIGURE 3.9: Detailed performance of assessed Strategy #3. The layout of panels in this figure is the same as in Figure 3.5, only substrategies represent different species ratios cut-offs used to pick fishing schedules conditional on a pre-season run size forecast (schedules shown in Figure 3.1, ratio thresholds shown in Table 3.2).
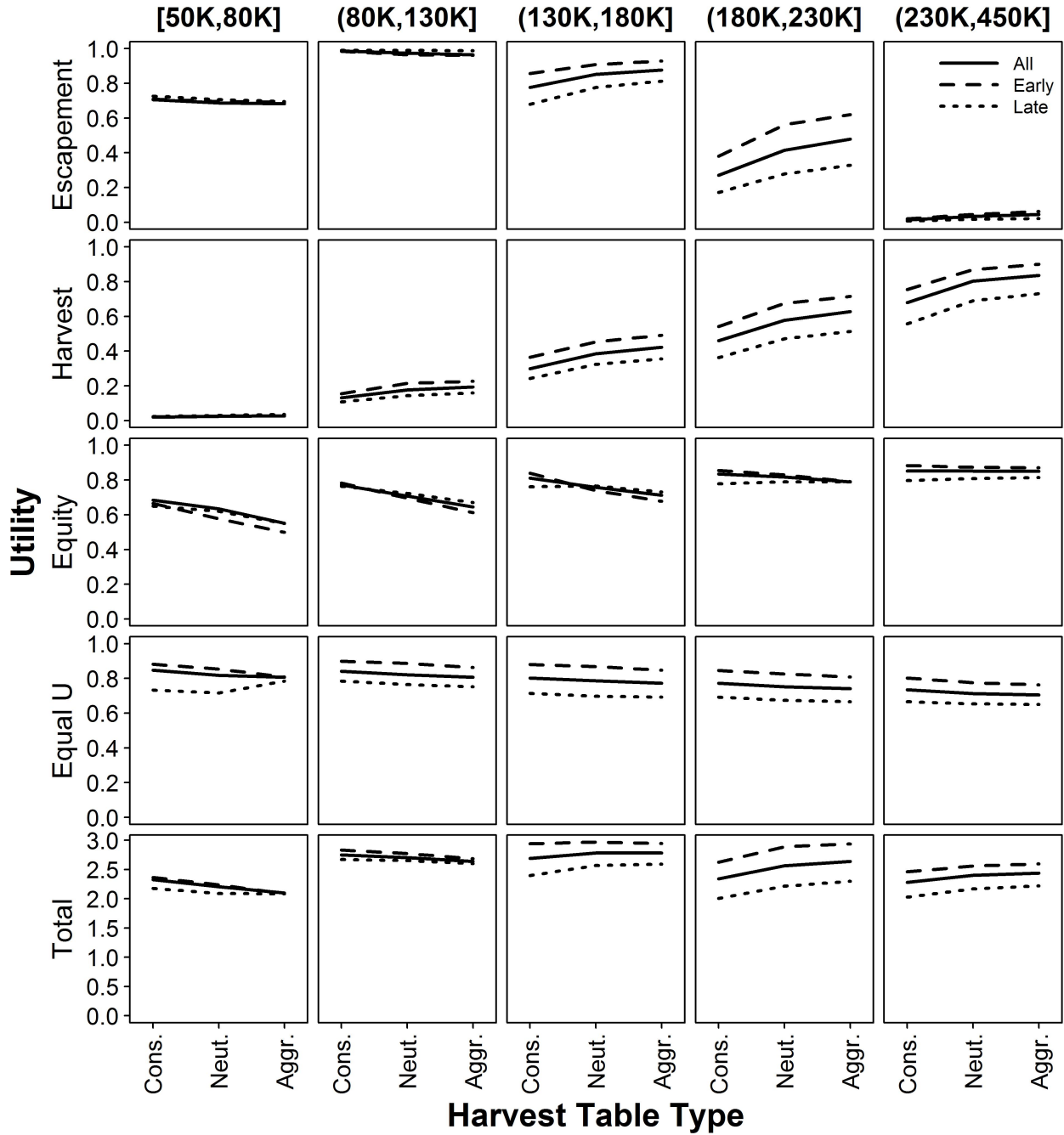
FIGURE 3.10: Detailed performance of assessed Strategy #4. The layout of panels in this figure is the same as in Figure 3.5, only substrategies represent different harvest tables used to set the number of days of open fishing per week based on how many fish are targeted to be harvested that week (harvest tables shown in Figure 3.3).
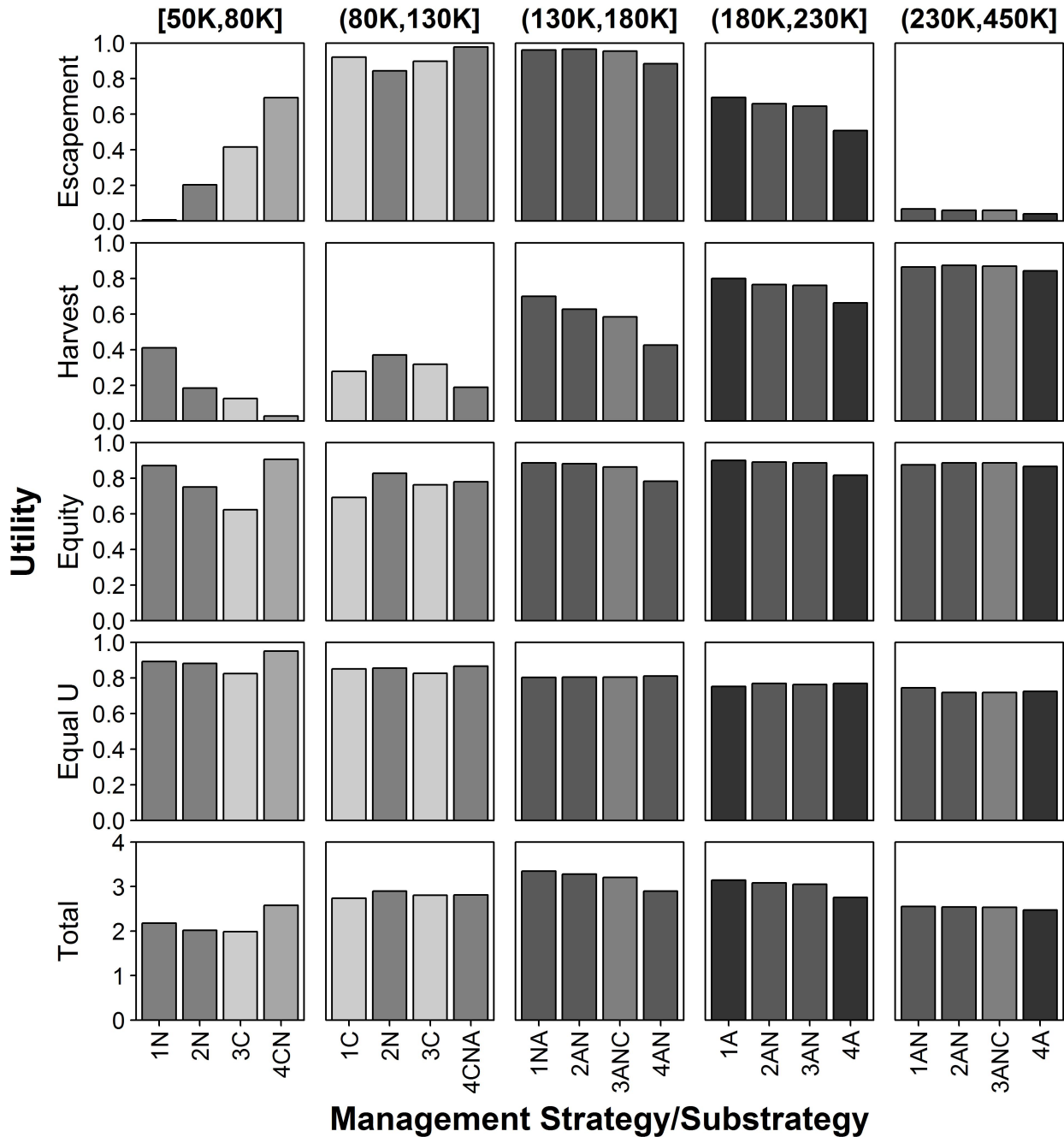
FIGURE 3.11: Comparison of utility according to the different metrics among strategies (with the best substrategy selected for comparison) and run sizes. Numbers represent the strategy, letters and colors indicate the selected substrategy (darker colors represent more aggressive substrategies; C = conservative, N = neutral, A = aggressive; multiple letters indicate a hybrid strategy resulting from a tie). Total utility was calculated according to the default weighting scheme, where all objectives received equal weight. Note the predominance of light grey bars (more conservative) in smaller runs and darker grey bars (more aggressive) in larger runs.
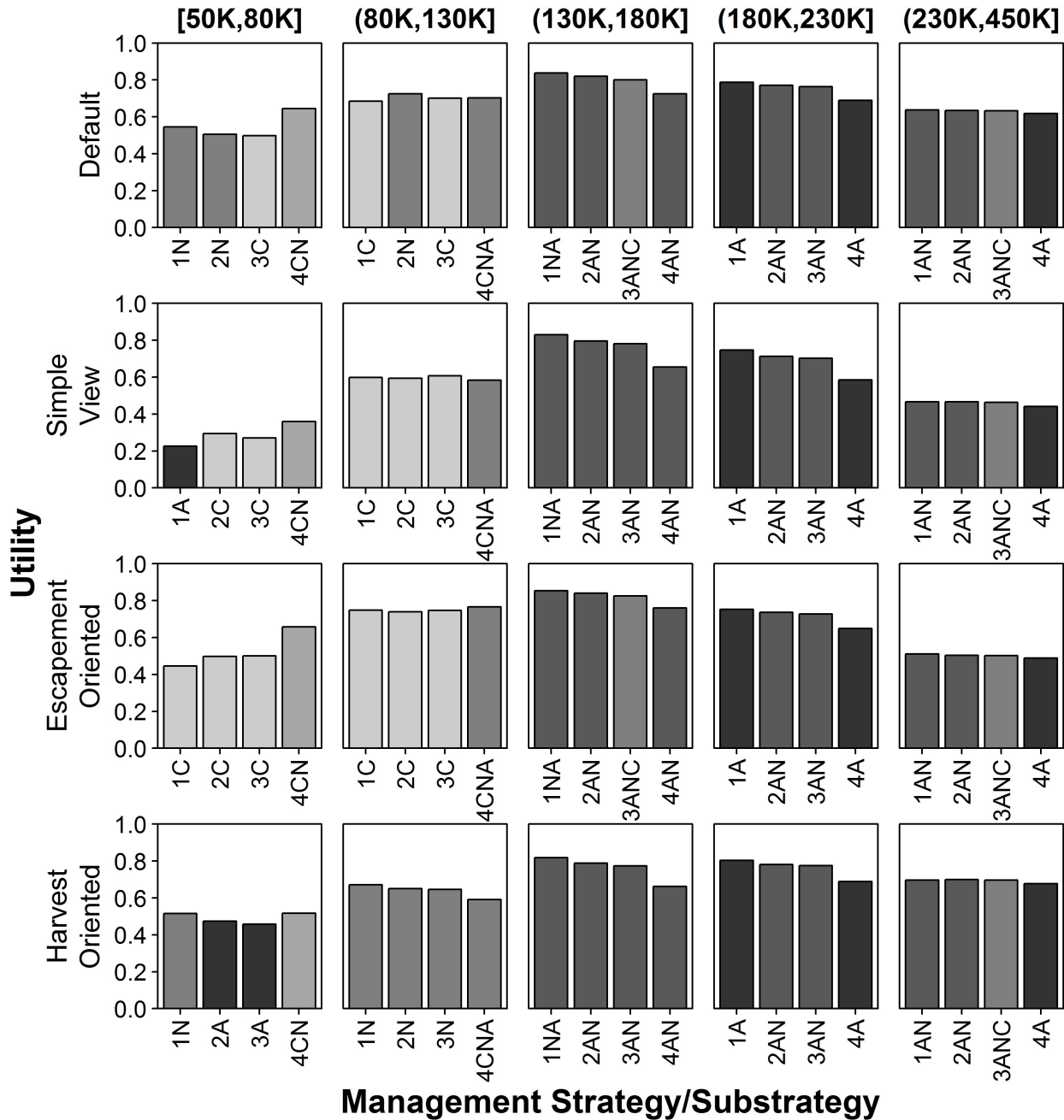
FIGURE 3.12: The total utility metric of the best-performing substrategy (letters/colors) for each strategy (numbers), when considering different weighting schemes (schemes described in Section 3.2.7.3). Bars are shaded based on the best substrategy, with darker greys representing more aggressive substrategies (C = conservative, N = neutral, A = aggressive; multiple letters indicate a hybrid strategy resulting from a tie). Total utility was scaled to the maximum attainable total utility for each weighting scheme, which is equal to the sum of the weights. Note the predominance of light grey bars (more conservative) in smaller runs and darker grey bars (more aggressive) in larger runs.

Chapter 4

Assessment Approaches for Mixed-stock Pacific Salmon Fisheries: Empirical and
Simulation-Estimation Applications

Abstract

Although many salmon populations in large drainages are composed of many smaller substocks,
they are often managed and harvested as a mixed-stock. There is likely to be heterogeneity
in productivity among the substocks, implying that some subset may need to be overfished if
the mixed-stock harvest is to be maximized. Methods to quantify substock productivity in
the context mixed-stock salmon fisheries are not well established, and fitting substock-specific
models independently may result in a loss or less adequate use of information. In this chapter,
an age-structured state-space framework was developed for simultaneously fitting multiple
spawner-recruit relationships. This approach may have advantages over simpler methods
such as allowing for more complete use of available data and sharing of information regarding
recruitment anomalies experienced by individual substocks. Four alternative state-space
models were developed and differed in their complexity in the treatment of recruitment
and maturity process error, and were compared to simpler regression-based approaches.
The models were fitted to reconstructed data from 13 Chinook salmon substocks in the
Kuskokwim River and estimation performance was assessed using simulation-estimation
trials. Biological and policy conclusions were found to be largely consistent among different
state-space model structures, but differed strongly from the regression-based approaches
which suggested substantially more aggressive harvest policies. The state-space framework
was shown to be largely unbiased when tested across a range of true population parameters,
though strong biases were found for regression approaches. The reliability and richness of

biological inferences from the state-space model suggest it shows promise for populating policy evaluations directed at investigating the strength of harvest-biodiversity trade-offs in mixed-stock salmon fisheries.

4.1   Introduction

Many salmon populations in large drainage systems are commonly harvested in a relatively small spatial area and are managed as a single stock. However, these "stocks" are instead stock-complexes, in which the aggregate stock is composed of several (and sometimes, many) substocks. These substocks are known to show differences in genotypic (Templin *et al.* 2014), phenotypic (*e.g.*, morphology; Hendry and Quinn 1997), behavioral (*e.g.*, run timing; Clark *et al.* 2015; Smith and Liller 2017a,b), and life history (*e.g.*, age-at-maturation, Blair *et al.* 1993) characteristics that are the result of adaptations to local environments after many generations of high spawning-site fidelity and reproductive isolation from conspecifics in other tributaries located within the same basin. It has been widely proposed that maintaining this diversity of local adaptation (hereafter, "biodiversity") is favorable both from ecosystem and exploitation perspectives. One argument is that in a system where many parts contribute to the whole, the variability in the aggregate characteristics can be dampened due to asynchrony in the subcomponent dynamics, a phenomenon known as the "portfolio effect" (Schindler *et al.* 2010, 2015).

Variability in these characteristics of portfolio components can ultimately lead to heterogeneity in productivity among them (Walters and Martell 2004). Productivity in this context is the ability of a population to replace itself after harvesting, often represented for salmon populations as the maximum number of future migrating adults (recruits) produced by one spawner (hereafter, α), which is attained at low spawner abundances due to density-dependent survival. Substocks j with higher α_j values can sustain greater exploitation rates (U) than

those with smaller α_j values, in fact, α_j can be expressed in terms of the exploitation rate that maximizes sustained yield from substock j ($U_{\text{MSY},j}$; Schnute and Kronlund 2002):

$$\alpha_j = \frac{e^{U_{\text{MSY},j}}}{1 - U_{\text{MSY},j}}. \tag{4.1}$$

Given that there is likely some level of heterogeneity in α_j and $U_{\text{MSY},j}$ among individual substocks, the logical conclusion is that in a mixed-stock fishery where the exploitation rate in year t (U_t) is common among all substocks, some weaker substocks must be exploited at $U_t > U_{\text{MSY},j}$ in order to fish the more productive substocks at $U_{\text{MSY},j}$[1]. This of course implies a trade-off, such that it may be necessary to over-exploit some substocks in order to maximize harvest benefits from others (Figure 4.1, Walters and Martell 2004).

Before these trade-offs can be considered by managers in a well-informed way, the shape and magnitude of the trade-off must first be quantified similarly to the examples as shown in Figure 4.1. Curves like this are generated using the estimated productivity and carrying capacity of all (or a representative sample) of the substocks contributing harvest to a mixed-stock fishery. These quantities are obtained using a spawner-recruit analysis, which involves tracking the number of recruits produced in each brood year (*i.e.*, parent year) by the number of fish that spawned in that same year and fitting a curve to the resulting pattern. The spawner-recruit literature is extensive, but focuses primarily on assessing single populations as opposed to substock components (but see the work on Skeena River, British Columbia sockeye salmon *O. nerka* substocks; Korman and English 2013; Walters *et al.* 2008). Substock-specific analyses are uncommon because of two factors: (1) the data to conduct well-informed substock-specific spawner-recruit analyses are often unavailable (20 – 30 years of continuous spawner and harvest counts/estimates and age composition for each substock) and (2) management actions in large mixed-stock fisheries may not be dexterous enough to deliberately exert higher exploitation rates on more productive substocks, so

---

[1]MSY is used here only as an example, this is true for other harvest levels, *e.g.*, 90% of MSY.

deriving substock-specific estimates could be of little utility. Regarding the former reason, there are some cases where the data do exist to perform these kinds of analyses, however methods to conduct multi-stock spawner-recruit analyses are not well-developed. Regarding the latter, even in cases where management cannot target particular substocks over others, understanding the nature of the trade-offs can be informative for evaluating candidate harvest policies for the mixed-stock in the context of substock biodiversity (Walters *et al.* 2018).

The methods to fit spawner-recruit models can be grouped into two broad categories: regression-based approaches (*e.g.*, Clark *et al.* 2009) and state-space (*i.e.*, time series) models (*e.g.*, Fleischman *et al.* 2013; Su and Peterman 2012). The regression-based approaches treat spawner-recruit pairs as independent observations, and are thus subject to substantial pitfalls when dealing with the inherent time-dependent properties and oftentimes large amounts of observation error found in spawner-recruit data sets (Walters and Martell 2004, Ch. 7). The consequence of ignoring the first issue is the "time-series bias", which chronically causes positive biases in α and negative biases in β, resulting in the same directional biases in U_{MSY} and S_{MSY}, respectively (*i.e.*, spuriously providing too aggressive harvest policy recommendations; Walters 1985). The second is known as the "errors-in-variables bias" and is known to cause an apparent scatter which inserts additional variability that commonly used regression estimators do not account for (Ludwig and Walters 1981). Though these methods have been known for their problems for over 30 years, they are still somewhat widely used (*e.g.*, Korman and English 2013). Unlike the regression-based approaches, the state-space class of models captures the process of recruitment events leading to future spawners while simultaneously accounting for variability in the biological and measurement processes that gave rise to the observed data (de Valpine and Hastings 2002; Fleischman *et al.* 2013). For these reasons, state-space spawner-recruit analyses have been rapidly gaining popularity, particularly in Alaska (Fleischman *et al.* 2013; Staton *et al.* 2017b; Su and Peterman 2012). Including this level of additional model complexity comes at computational costs, as these

models are well-suited for Bayesian inference with Markov Chain Monte Carlo (MCMC) methods (Newman *et al.* 2014, Ch. 4), but has been shown to reduce bias in estimates in some circumstances (Su and Peterman 2012; Walters and Martell 2004).

In the Kuskokwim River Chinook salmon fishery, there has been recent interest in considering biodiversity to explicitly inform the drainage-wide escapement goal. In obtaining substock-specific parameters to inform such policy analyses, it will be difficult to determine which method is appropriate, given many possible model structures, frequently missed sampling, and unknown biases in the data. Before strong inferences can be made from the ultimate trade-off analyses of interest, the performance of the estimation models used to parameterize them needs to be evaluated, as well as the appropriate level of model complexity needed to address the problem with sufficient accuracy. In this chapter, I evaluate the performance of a range of assessment models for mixed-stock salmon fisheries *via* simulation-estimation and apply them to Chinook salmon substocks in the Kuskokwim River as a case study. The objectives were to:

(1) develop a set of varyingly complex multi-stock versions of state-space spawner-recruit models,

(2) determine the sensitivity of biological and trade-off conclusions to assessment model complexity (including those obtained using regression-based approaches) using empirical data from Kuskokwim River Chinook salmon substocks, and

(3) test the performance of the assessment models *via* simulation-estimation trials.

4.2   Methods

This analysis was conducted in both an empirical and a simulation-estimation framework to evaluate the sensitivity and performance of assessment strategies for the mixed-stock assessment problem in Pacific salmon fisheries. First, all assessment methods (two regression-based and four state-space models) were fitted to observed data from the Kuskokwim River

substocks ($n_j = 13$) to determine the extent to which the choice of assessment model structure might influence biological and management conclusions. Next, a hypothetical exploited system with known properties was simulated and was composed of several age-structured substocks. Then, these hypothetical populations were sampled per a realistic sampling scheme and each of the assessment models were fitted to the resulting data sets. Estimation performance regarding quantities used in management (*e.g.*, U_{MSY} and S_{MSY}) was calculated from the resulting estimates. Inference from the simulation can then be used to justify an appropriate level of model complexity.

4.2.1 Multi-stock spawner-recruit models

4.2.1.1 Regression-based models

Two regression-based approaches to estimating Ricker (1954) spawner-recruit parameters in the multi-stock case were assessed: (1) a single mixed-effect regression model with random intercepts and (2) independent regression models. The Ricker (1954) spawner-recruit model can be written as:

$$R_y = \alpha S_y e^{-\beta S_y + \varepsilon_y} \tag{4.2}$$

where R_y is the total recruitment produced by the escapement S_y in brood year y, α is the maximum expected recruits per spawner (RPS), β is the inverse of the escapement that is expected to produce maximum recruitment (S_{MAX}), and ε_y are independent mean zero random variables often assumed to be attributed solely to environmental fluctuations in juvenile survival. Primary interest lies in estimating the population dynamics parameters α and β as they can be used to obtain biological reference points off of which sustainable harvest policies can be developed. The Ricker (1954) function in (4.2) is increasing at small escapements and declining at large ones, though can be linearized:

$$\log(\text{RPS}_y) = \log(\alpha) - \beta S_y + \varepsilon_y, \tag{4.3}$$

allowing for estimation of the parameters $\log(\alpha)^2$ and β in a linear regression framework using the least squares method or likelihood methods under the assumption that $\varepsilon_y \sim \text{N}(0, \sigma_R^2)$ (Clark *et al.* 2009; Hilborn and Walters 1992). This relationship is nearly always declining, implying a compensatory effect on juvenile survival (*i.e.*, RPS) with reductions in spawner abundance (Rose *et al.* 2001). Obtaining the R_y component of RPS_y can be problematic for salmon populations that spawn at multiple ages (like Chinook salmon) given the prevalence of non-consecutive sampling years of either age composition or escapement data (see Appendix D.5 for more details, also Figure D.1).

A multi-stock formulation of this model can be expressed by including substock-specific random effects on the intercept $[\log(\alpha)]$:

$$\log(\text{RPS}_{y,j}) = \log(\alpha_j) - \beta_j S_{y,j} + \varepsilon_y,$$
$$\log(\alpha_j) = \log(\alpha) + \varepsilon_{\alpha,j}, \tag{4.4}$$
$$\varepsilon_{\alpha,j} \sim \text{N}(0, \sigma_\alpha^2).$$

It would be nonsensical to include substock-level random effects on the slope, given that β is a capacity parameter related to the compensatory effect of resource limitation experienced by juveniles, likely in the freshwater environment (*i.e.*, amount of habitat as opposed to quality of habitat). Fitting the individual substock models in this hierarchical fashion allows for the sharing of information such that the more intensively assessed substocks can help inform those that are more data-poor.

---

[2]Throughout, the use of $\log(x)$ denotes the natural logarithm of x.

The mixed-effect model may have the benefit of sharing information to make some substocks more estimable, but it should also have the tendency to pull the extreme α_j (those in the tails of the hyperdistribution) toward α. This behavior may not be preferable for policy recommendations, as it should tend to dampen the extent of heterogeneity estimated in α_j. For this reason, independent regression estimates for each substock were also obtained for evaluation. In estimating the parameter $\log(\alpha)$, a lower bound constraint of zero was used in all regression models. This was necessary to prevent the models from estimating biologically implausible parameters: if $\log(\alpha) < 0$, then no amount of spawners would be expected to replace themselves (let alone provide surplus) in their most productive state, in which case the population would not likely be in existence.

4.2.1.2 State-space models

Four versions of the state-space formulation were developed. As three versions were simplifications of the full model, the full model is presented completely here and the changes resulting in the other three model structures are described following the description of the full model. The JAGS code for the full model is presented in Appendix C.

The state-space formulation of a multi-stock spawner-recruit analysis developed and evaluated here is an extension of various single-stock versions (*e.g.*, Fleischman *et al.* 2013). Walters *et al.* (2008) used a similar model using maximum likelihood methods to provide estimates of >50 substocks in the Skeena River drainage, British Columbia. The model presented here was fitted in the Bayesian mode of inference using program JAGS (Plummer 2017), and allows for relaxation of certain assumptions made by Walters *et al.* (2008) such as the important notion of perfectly shared recruitment residuals (*i.e.*, anomalies – deviations from the expected population response).

The state-space model is partitioned into two submodels: (1) the process submodel which generates the latent (*i.e.*, true but unobserved) states of $R_{y,j}$ and the resulting calendar

year states (*e.g.*, $S_{t,j}$) and (2) the observation submodel which fits the latent states to the observed data (notation summarized in Tables 4.1 and 4.2). Note that this method allows for missing calendar year observations and does not require excluding brood year recruitment events that were not fully observed as was necessary for the regression-based models (see Appendix D.5).

The recruitment process operated by producing a mean prediction from the deterministic portion of the Ricker (1954) relationship in (4.2) for n_y brood years for each of the n_j substocks. From these deterministic predictions, auto-correlated process variability was added to generate the realized latent recruitment states. To populate the first n_a calendar year abundance states with recruits of each age a, the first a_{max} brood year expected recruitment states were not linked to a spawner abundance through (4.2) (because the S_y component was not observed), but instead were assumed to have a constant mean equal to the unfished equilibrium recruitment (where non-zero S_j produces $R_j = S_j$ when unexploited and in the absence of process variability):

$$\dot{R}_{y,j} = \frac{\log(\alpha_j)}{\beta_j}, \tag{4.5}$$

where $\dot{R}_{y,j}$ is the expected (*i.e.*, deterministic) recruitment in brood year y from substock j with Ricker parameters α_j and $\beta_j{}^3$. The remaining $n_y - a_{max}$ brood years had an explicit time linkage:

$$\dot{R}_{y,j} = \alpha_j S_{t,j} e^{-\beta_j S_{t,j}}, \tag{4.6}$$

where $t = y - a_{max}$ is the tth calendar year index in which the escapement produced the recruits in the yth brood year index.

---

[3](4.5) can be derived by substituting R in for S in (4.2) and solving for R. Similarly, equilibrium recruitment in the case where the stock is fished at U can be obtained by substituting $R(1 - U)$ in for S.

From these deterministic predictions of the biological recruitment process, lag-1 auto-correlated process errors were added to produce the latent states:

$$\log(R_{y,1:n_j}) \sim \text{MVN}\left(\log(\dot{R}_{y,1:n_j}) + \omega_{y,1:n_j}, \Sigma_R\right), \tag{4.7}$$

where

$$\omega_{y,1:n_j} = \phi\left(\log(R_{y-1,1:n_j}) - \log(\dot{R}_{y-1,1:n_j})\right), \tag{4.8}$$

and $R_{y,1:n_j}$ is the vector[4] of latent recruitment states across the n_j stocks in brood year y, $\omega_{y,1:n_j}$ is the portion of the total process error attributable to serial auto-correlation, ϕ is the lag-1 auto-correlation coefficient (constant across substocks), and Σ_R is a covariance matrix representing the white noise portion of the total recruitment process variance. In the full model, Σ_R was estimated such that each substock was assigned a unique variance and covariance with each other substock. This was achieved by using an inverse Wishart prior distribution, with degrees of freedom equal to $n_j + 1$ and the scale matrix populated with zero-value elements along the off-diagonals and each diagonal element equal to one, which inserts little information about the covariance matrix Σ_R (Plummer 2017). The multivariate normal errors were on the logarithmic scale so the variability on $R_{y,j}$ was lognormal, which is the most commonly used distribution for describing recruitment variability (Walters and Martell 2004). Further, the multivariate normal was used as opposed to n_j separate normal distributions so the degree of synchrony in brood year recruitment deviations (*i.e.*, recruitment process errors) among substocks could be captured and freely estimated in Σ_R.

The maturity schedule is an important component of age-structured spawner-recruit models, as it determines which calendar years the brood year recruits $R_{y,j}$ return to spawn (and be observed). Recent state-space spawner-recruit analyses have accounted for brood

---

[4]Throughout, the syntax *a:b* denotes a sequence $a, a+1, \cdots, b$ with $b - a + 1$ elements for $b > a$. For example, $R_{y,1:n_j}$ denotes the vector of recruitments across substocks within a brood year.

year variability in maturity schedules as Dirichlet random vectors drawn from a common hyperdistribution characterized by a mean maturation-at-age probability vector ($\pi_{1:n_a}$) and an inverse dispersion parameter (D) (see Fleischman *et al.* 2013; Staton *et al.* 2017b, for implementation in JAGS), and the same approach was used for the full model with maturity schedules shared perfectly among substocks within a brood year. Brood year-specific maturity schedules were treated as random variables such that:

$$p_{y,a} \overset{\text{iid}}{\sim} \text{Dirichlet}(\pi_{1:n_a} \cdot D). \tag{4.9}$$

where $p_{y,a}$ is the probability a fish spawned in brood year y will mature at age a.

In order to link $R_{y,j}$ with calendar year observations of escapement from each substock, $R_{y,j}$ was allocated to calendar year runs-at-age ($N_{t,a,j}$) based on the maturity schedule:

$$N_{t,a,j} = R_{t+n_a-a,j} p_{t+n_a-a,a}, \tag{4.10}$$

and the total run returning to substock j in year t was the sum of mature fish within a calendar year across ages:

$$N_{t,j} = \sum_{a=1}^{n_a} N_{t,a,j}. \tag{4.11}$$

The harvest process was modeled using a freely estimated annual exploitation rate (U_t) time series, which was assumed to apply equally to all substocks (by default, but see Section 4.2.2.3 for a relaxation of this assumption):

$$H_{t,j} = N_{t,j} U_t, \tag{4.12}$$

and escapement was obtained as:

$$S_{t,j} = N_{t,j}(1 - U_t). \tag{4.13}$$

The quantity H_t aggregated among all substocks was obtained by summing $H_{t,j}$ within a t index across the j indices. The true age composition returning in year t to substock j ($q_{t,a,j}$) was obtained as:

$$q_{t,a,j} = \frac{N_{t,a,j}}{N_{t,j}}. \tag{4.14}$$

Three data sources were used to fit the state-space model: (1) observed escapement from each substock ($S_{obs,t,j}$) with assumed known coefficients of variation (CV), (2) total harvest arising from the aggregate stock ($H_{obs,t}$) with assumed known CV, and (3) the age composition of substocks with these data each calendar year ($q_{obs,t,a,j}$; which had associated effective sample size $ESS_{t,j}$ equal to the number of fish successfully aged for substock j in year t). The CVs were converted to lognormal variances:

$$\sigma^2_{\log} = \log(\mathrm{CV}^2 + 1), \tag{4.15}$$

and used in lognormal likelihoods to fit the time series $S_{t,j}$ to $S_{obs,t,j}$ and H_t to $H_{obs,t}$. Calendar year age composition was fitted using independent multinomial likelihoods with parameter vectors $q_{t,1:n_a,j}$ and observed vectors of ($q_{obs,t,1:n_a,j} \cdot ESS_{t,j}$).

Three alternative formulations of the state-space model were evaluated, and all were simplifications of the full model described above regarding the structure of (1) the covariance matrix on recruitment residuals (Σ_R) and (2) the maturity process (see Table 4.3 for a summary of the models fitted in this analysis). The simplest state-space model did not include brood year variability in maturity schedules and Σ_R was constructed by estimating

a single σ_R^2 and ρ (common across all substocks and substock pairs[5]), and populating the diagonal elements with σ_R^2 and off-diagonal elements with $\rho\sigma_R^2$. This simplest model is denoted as SSM-vm throughout the rest of this chapter. In one intermediate model (SSM-vM), brood year maturation variability was included but Σ_R was constructed as in the simplest model. In the other intermediate model (SSM-Vm), brood year variability in maturation was not included but Σ_R was fully estimated as in the full model (SSM-VM). As for the choice of notation, lowercase letters indicate the simple version of a structure and uppercase letters indicate the complex structure; "v" refers to recruitment covariance structure and "m" refers to maturity. These two structural uncertainties (complexity in recruitment covariance and maturity variability) were chosen for evaluation here because they are two key areas where an analyst might question if the available data are adequate for model fitting and inference. In other words, these are two key model components where may be important to know if the complex versions are reliably estimable with a reasonable amount of data.

4.2.2 Kuskokwim empirical analysis

4.2.2.1 Study system

All six assessment models (two regression-based and four state-space models) evaluated were fitted to empirical data from Chinook salmon substocks of the Kuskokwim River located in western Alaska (Figure 4.2). The Kuskokwim River salmon fishery can very well be described as a mixed-stock fishery, both for multiple salmon species (predominately Chinook, chum *O. keta*, and sockeye salmon) and for multiple substocks of the same species. Fish originating from and returning to the various tributaries enter through the bulk of the fishery as a mixed-stock, though Chinook salmon stocks traveling to the headwaters have been

---

[5]One drawback of constructing Σ_R this manner is that MCMC samples of $\rho < -0.05$ for a 13×13 covariance matrix results in JAGS crashing with an error related to positive-indefiniteness, which is prohibited by JAGS (Plummer 2017). Thus, a constraint was required to maintain $-0.05 \leq \rho < 1$ to prevent the sampler from crashing.

illustrated to enter the main-stem earliest in the summer migration (Smith and Liller 2017a,b) so a limited ability to direct harvest toward or away from these substocks is possible by manipulating the front portion of the fishery (Chapter 3, this dissertation). It is acknowledged that the assessment program does not sample all tributaries within the Kuskokwim River where Chinook salmon spawn (Figure 4.2), but total run size between 1976 and 2017 has been estimated *via* run reconstruction (Liller *et al.* 2018) and large-scale mark-recapture studies (Schaberg *et al.* 2012; Smith and Liller 2017a,b; Stuby 2007).

4.2.2.2 Data sources

The data set used included counts of Chinook salmon at many locations throughout the Kuskokwim River system (Figure 4.2). Nearly all data were collected by projects managed by the Alaska Department of Fish and Game (ADF&G) and a complete description of data needs and preparation procedures is provided in Appendix D. The raw escapement data set available spanned 20 different escapement monitoring projects (six weirs and 14 aerial surveys) and 42 calendar years from 1976 – 2017; see Head and Smith (2018) for details on tributary escapement monitoring. Some pre-processing was required to convert the aerial survey index counts to estimates of total spawners (Appendix D.2). Annual estimates of Chinook salmon harvest originating from both subsistence and commercial fisheries in each year was also available, as was the estimated exploitation rate of the aggregate stock (details in Appendix D.3). Finally, age composition data were available for the six substocks monitored *via* weir programs (details in Appendix D.4).

4.2.2.3 Sensitivity Analyses

Two sensitivity analyses were conducted to test the robustness of inference from the state-space models with the Kuskokwim data. First, the default assumption that all substocks have been fished at the same rate each year is tenuous. A term was included in the model that

allowed for substocks to have differential vulnerability (v_j) by replacing U_t with $U_t v_j$ in (4.12) and (4.13) which makes an adjustment that acknowledges some substocks experienced higher exploitation rates than others. This alteration changes the interpretation of the parameter vector U_t to be the exploitation rate of fully vulnerable substocks. Without additional information on what portion of $H_{obs,t}$ was attributable to each substock going back in time, the v_j elements are not estimable. In the absence of this information for Kuskokwim River Chinook salmon, the value of the vulnerability parameters was assigned by calculating the fraction of the fishing households along the main-stem of the river that each substock must travel past in order to reach their natal spawning grounds. Fishing household data were available from post-season interviews conducted by ADF&G (*e.g.*, Hamazaki 2011; Shelden *et al.* 2016b)[6]. Although this method ignores the temporal overlap of the fishery (Hamazaki 2008) with the arrival timing of particular substock groups (Smith and Liller 2017a,b), it was intended as a first step at determining how much the conclusions might depend on how the internal harvest accounting was specified. No attempt was made to alter how harvest was apportioned for use in the regression-based models.

As a secondary sensitivity analysis, the information content of the age composition data was reduced. In the default case, each annual multinomial age composition vector had sample size equal to the number of fish successfully aged for that substock/year combination. For some substocks/years, this number was quite high from a multinomial sampling perspective when the number of categories is small (*e.g.*, $\approx 1{,}200$ samples across four age categories). To assess whether this strength of information had an impact on the inference, the effective sample size was manipulated such that the maximum number of fish sampled for a substock was assigned $ESS_{t,j} = 100$, and the other years with data were scaled proportionately.

---

[6]These data were the same as those used in Chapter 3 of this dissertation when determining the spatial distribution of fishing effort and salmon needs in the management strategy evaluation operating model.

4.2.2.4 Comparisons of model output

Key population dynamics parameters and biological reference points were compared among six assessment models wherever possible to determine the extent to which the management conclusions for Kuskokwim River Chinook salmon substocks might change based on the model structure used. Where appropriate, quantities were averaged across substocks to facilitate comparisons and included indicators of the average substock's productivity ($\bar{\alpha}_j$ and $\bar{U}_{\mathrm{MSY},j}$; the "bar" denotes average across substocks) and size ($\bar{S}_{\mathrm{eq},j}$, $\bar{S}_{\mathrm{MSY},j}$, and $\bar{S}_{\mathrm{MAX},j}$). Reference points for the aggregate mixed-stock included U_{MSY} and S_{MSY} and a set of metrics that incorporate biodiversity aspects as well: S_p^* H_p^*, and U_p^*, which represent the equilibrium mixed-stock escapement, harvest, or exploitation rate, respectively, that would ensure no more than $p \cdot 100\%$ of substocks would be overfished, defined here as the case where a substock would be fished at $U > U_{\mathrm{MSY},j}$. Three levels of p were extracted: 0.1, 0.3, and 0.5.

Model fits to the data ($S_{obs,t,s}$, $H_{obs,t}$, and $q_{obs,t,a,j}$) were examined for all state-space models and noteworthy differences among model structures were identified. Estimates of synchrony in recruitment anomalies were examined, both between two average substocks ($\bar{\rho}_{i,j}$), and among all substock pairs. For the simple recruitment variance models (SSM-vm and SSM-vM), correlations between each substock pair were conducted by applying Pearson's r coefficient to two substocks' estimated recruitment anomaly time series; in the complex variance models these correlations were captured in the freely estimated covariance matrix (Σ_R), and they were extracted and summarized. Although the state-space models were ignorant of spatial relationships among the substocks, visual comparisons were made to determine if substocks closer in proximity showed higher synchrony than those spaced more distantly, as might be expected. The auto-correlation parameter (ϕ) and characteristics of the maturity schedules (π_a and D) were also compared among state-space models.

Harvest-biodiversity trade-offs were visualized in two ways. First, the equilibrium mixed-stock escapement and harvest were calculated[7] at each level of a mixed-stock exploitation rate that affected all substocks equally. The fraction of substocks overfished and trending towards extirpation (defined as the case where equilibrium $S_j \leq 0$) were also calculated. As a secondary, possibly more direct trade-off visualization, the fraction of substocks expected to be overfished and trending towards extirpation was plotted against the fraction of the mixed-stock MSY obtained at each exploitation rate.

Some substocks could not be fitted using the regression approach because they had fewer than three observed pairs of $\text{RPS}_{y,j}$ (see Appendix D.5 for details). When comparing quantities to the estimates from the state-space models, the substocks that could not be fitted with the regression approaches were removed from the state-space model outputs. When comparisons were made among state-space models only, these substocks were retained.

4.2.3 Simulation-estimation trials

To test the performance of these models, 160 hypothetical salmon data sets were created and were designed to mimic the Kuskokwim River empirical data set. Each of the 160 data sets was passed to each of the six assessment models (Section 4.2.1 and Table 4.3) to evaluate which methods return estimates closest to the true parameters.

Given that the state-space model is a much more natural model of this system (which has intrinsic time series properties) than the regression-based versions, it was used as the foundation of the operating model (*i.e.*, state-generating model). The biological submodel was more complex than the most complex estimation model – namely with regards to the maturity schedule, which had a modest level of substock variability in mean maturity but with highly correlated brood year variability; see Appendix E for details. In order to serve

---

[7]Equilibrium equations for the Ricker (1954) model can be found in Table 2 of Schnute and Kronlund (2002) for the case where $\gamma = 0$. These were applied to individual substocks then summed across substocks to obtain the mixed-stock quantities.

as the state-generating model for the simulation, the state-space model needed only to be populated with true parameters, initial states, and a harvest control rule. A fixed exploitation rate policy was used (chosen to maximize yield without overfishing more than 30% of the substocks) with a modest amount of implementation error to ensure the data time series were generated with enough contrast in spawner abundance. $n_j = 13$ substocks were simulated with different parameters $U_{\mathrm{MSY},j}$ and $S_{\mathrm{MSY},j}$ which took on the values of random posterior draws from the most complex state-space model fitted to the Kuskokwim River Chinook salmon substock data. All other parameters were chosen to mimic the estimated values from the Kuskokwim analysis, with the exception of Σ_R, which was set to have a modest amount of substock recruitment variability ($\bar{\sigma}_{j,j} \approx 0.4$); $\rho_{i,j}$ for each pair of substocks was simulated randomly to be between -1 and 1, but approximately zero when averaged across all substocks.

For a given set of simulated true states, a set of observed states ($S_{obs,t,j}$, $H_{obs,t}$, $q_{obs,t,a,j}$) was generated by adding sampling error to each year[8], which represented the value that would have been observed if the sampling project operated that year. Observation errors in escapement and harvest estimates were generated with lognormal variability and multinomial sampling for the age composition, as assumed in the state-space estimation model. Frequency of sampling on each substock (*i.e.*, simulated data collection) was set to approximately mimic the Kuskokwim River historical monitoring program (Figure D.1). The sampling frequency was designed to continue to generate sampling schedules until one was found that ensured no substock had fewer than three observations of $\mathrm{RPS}_{y,j}$ which allowed the regression-based models to be fitted to all substocks. Aggregate harvest data ($H_{obs,t}$) was assumed to be available every year and it was assumed that the annual mixed-stock exploitation rate could be estimated in an unbiased fashion.

---

[8]Fifty years of population dynamics with exploitation were simulated prior to data collection. This practice "burns-in" the harvest policy into the population dynamics, and intentionally violates the assumption made by the state-space estimation model that sampling initiated under unfished conditions, as shown in (4.5).

Estimation performance in terms of accuracy among assessment models was calculated using the proportional error ($\frac{x_\text{est} - x_\text{true}}{x_\text{true}}$), and the central tendency and variability of errors among random data sets were visualized using boxplots. Key quantities of interest for comparison between the regression approaches and the state-space models included: S_MSY, U_MSY, S_p^*, U_p^*, α_j, $U_{\text{MSY},j}$, and $S_{\text{MSY},j}$. For the state-space models, the ability to accurately estimate the abundance states of $R_{y,j}$, $S_{t,j}$, and H_t was also assessed, summarized for early years and all years separately to investigate the influence of the assumption that the observation time series began at unfished equilibrium, as described by (4.5). Credible interval coverage at the 95% level was assessed by determining the fraction of obtained credible intervals included the true value. Additionally, model run times and convergence diagnostics were summarized for all models that successfully fitted the simulated data sets.

4.2.4 Computation

All parameter estimation was conducted in the Bayesian framework using the MCMC engine JAGS (Plummer 2017) invoked through R (R Core Team 2018) using the package `{jagsUI}` (Kellner 2017). All priors for the regression-based methods (Table 4.4) and state-space models (Table 4.5) were selected to be as minimally informative as possible, while still preventing the MCMC sampler from exploring highly unlikely areas of the parameter space which may cause it to crash. Two packages were written to streamline the workflow of this analysis[9]: one to handle the generation of simulated age-structured salmon populations from a mixed-stock fishery (`{SimSR}`; Staton 2018c) and one to handle the fitting and summarization of model output (`{FitSR}`; Staton 2018b). The former is one that other researchers may find useful for streamlining salmon population dynamics studies, whereas the later is highly specific to this particular analysis (in terms of the models included). Additionally, extensive use was made

---

[9]For reference, any of the packages written for this analysis can be installed on any R user's computer by executing: `devtools::install_github("bstaton1/pkgname")`.

of another package for summarizing the output of JAGS models, which other researchers using JAGS may find useful ({codaTools}; Staton 2018a).

MCMC sampling was conducted using sufficiently long chains (Table 4.6) to ensure adequate sampling of the posterior parameter space, and was assessed using visual inspection of MCMC sampling behavior and the convergence diagnostic proposed by Brooks and Gelman (1998). Adequate sampling was further verified for key estimated quantities using the effective sample size and the Raftery-Lewis diagnostic (Raftery and Lewis 1992). All posterior distributions were summarized using the median and 95% equal-tailed credible intervals. In cases where a quantity was averaged over substocks or years (*e.g.*, $\bar{\alpha}_j$), the average was calculated for each joint posterior sample, then the resulting marginal posterior of the average was summarized.

As a result of the long required model run times and the large number of simulated salmon data sets (*i.e.*, replicates) for the simulation study, this analysis required more computing power than the previous chapters in this dissertation. This analysis was conducted using high performance computing services provided by the Alabama Supercomputer Authority. This allowed many jobs to be executed at simultaneously (*i.e.*, in parallel), where each job fitted one model to one data set using one processor core for each chain, also in parallel. In this way, it was possible to complete just over three years of computing time (if each job was executed back-to-back) in less than one month.

4.3   Results

4.3.1   Kuskokwim River empirical analysis

4.3.1.1   State-space models fit to data

In general, the four state-space models produced similar escapement state time series, especially in years with observed escapement data (Figure 4.3). For several substocks, there

were no escapement data prior to the mid-1990s, and this is one area the various state-space models produced different escapement estimates. In this early portion of the time series, models with brood year variability in maturity included (SSM-vM and SSM-VM) tended to estimate higher escapement abundance than the models with time-constant maturity (SSM-vm and SSM-Vm). For example, substocks spawning in the Kwethluk, George, Holokuk, and Takotna rivers all showed this pattern (Figure 4.3).

There were several cases where extremely high (and seemingly unrealistic) escapement states were estimated by the state-space model, though these only occurred in the versions with simple maturation schedules (Figure 4.3). The period in the late-1980s and the early 1990s had much (*i.e.*, $5 - 10$ times) higher estimated escapement than ever observed for the Holitna, Pitka, and Tatlawiksuk substocks under either models SSM-vm or SSM-Vm. The George River substock had abnormally large escapements in the mid-1990s, and was again most exaggerated for the SSM-vm and SSM-Vm versions. All of these cases occurred when no escapement data were available; in years with data all state-space models fitted the escapement data quite well (Figure 4.3).

In general, the fit to the aggregate harvest data was good (Figure 4.4), though the time-varying maturity schedule versions (SSM-vM and SSM-VM) fitted the data nearly perfectly. Constant maturity schedule versions resulted in a harvest state that was greater than twice as large as the observed state in 1976 (the first observed year). These models did not show discrepancies nearly this large for the rest of of the time series (Figure 4.4). Note the precipitous decline in harvest starting in the late-2000s, which coincides with a nearly decade-long period of low productivity for this stock.

Just as for the escapement and harvest data, the state-space models generally fitted the age composition data well. The only differences among model structures came in the distinction of how maturity was treated. Complex maturity models had visually better fit than simple maturity models (Figure 4.5) for most substocks and ages.

4.3.1.2 Regression models fit to data

Ten of the 13 substocks had sufficient data to fit the regression models (*i.e.*, more than three brood year spawner-recruit pairs): those spawning in the Kisaralik, Oskawalik, and Holitna subdrainages. In general, the two regression approaches provided similar fits to the data. At least five substocks had essentially identical fitted lines between approaches (Figure 4.6), and only two substocks showed major discrepancies between the fitted relationships: the Holokuk and the Pitka substocks. For the Holokuk relationship, the three observed data points suggested a negative intercept and an increasing slope (Figure 4.6), however the intercepts of these models were constrained to be positive to prevent biological impossibilities ($\alpha_j < 1$), which explains the poor fit in this case.

4.3.1.3 Comparisons of estimated spawner-recruit dynamics

There were large discrepancies in suggested population productivity parameters between the regression models and the state-space models (Table 4.7). For these comparisons, the three substocks that could not be fitted by the regression approaches were removed from the summaries of the state-space models. Both regression approaches suggested the maximum productivity of the average substock ($\bar{\alpha}_j$) was far higher than any of the state-space models. The independent regression approach provided an estimate of $\bar{\alpha}_j = 7.74$ $(4.47 - 20.22;$ 95% credible limits) and the mixed-effect regression approach suggested $\bar{\alpha}_j = 4.63$ $(3.16 - 7.45)$. Most state-space models suggested $\bar{\alpha}_j < 3$, with the highest upper 95% credible limit being 5.01, obtained by SSM-vM (Table 4.7).

These patterns are well-illustrated at the substock-level by visualizing the expected recruitment at each spawner abundance for each model (Figure 4.7). It is evident that the four state-space models behaved similarly near the origin (which is governed by α_j), whereas in many cases the regression models suggested steeper slopes near the origin (corresponding to higher values of α_j). State-space models tended to disagree more at larger spawner abundances

(Figure 4.7), suggesting that inferences about stock size and the strength of compensation are dependent on the details how recruitment variance and maturity are modeled for some substocks. As would be expected from Figure 4.6, the mixed-effect regression approach differed from the independent regression estimates the most for substocks with fewer observations (*e.g.*, Pitka *versus* Kogrukluk; Figure 4.7).

These differences in estimated productivity translated directly to the maximum sustainable exploitation rate for the average substock ($\bar{U}_{\text{MSY},j}$): the independent and mixed-effects regression approaches resulted in estimates of 0.58 (0.47 − 0.68) and 0.51 (0.41 − 0.6), respectively, whereas the point estimates ranged from 0.37 − 0.45 across the state-space models (Table 4.7). In terms of the exploitation rate that would maximize yield from the aggregate mixed stock (U_{MSY}), the regression approaches suggested the stock should be fished quite hard to obtain MSY (as high as $U_{\text{MSY}} = 0.78$ for the independent regression approach), whereas all state-space models suggested $U_{\text{MSY}} < 0.5$ for point estimates and no 95% credible intervals exceeded 0.7 (Table 4.7).

In comparing the state-space model estimates from these 10 substocks to the complete set of 13 substocks, the average productivity and maximum sustainable exploitation rates were quite similar, indicating the stocks with insufficient data for regression were missing at random in these regards (compare the rows corresponding to $\bar{\alpha}_j$, $\bar{U}_{\text{MSY},j}$, and U_{MSY} between Tables 4.7 and 4.8).

One metric of substock size is the spawner abundance expected to exactly replace itself under unfished equilibrium conditions ($S_{\text{eq},j}$). When averaged across substocks ($\bar{S}_{\text{eq},j}$), the regression approaches suggested the Kuskokwim River substocks were approximately 2,500 fish (~25%) smaller than the state space models did (Table 4.7). The other metrics used for measuring substock size ($S_{\text{MAX},j}$ and $S_{\text{MSY},j}$) followed the same pattern: smaller values for the regression approaches than the state-space models, though the magnitude varied depending on the metric (larger discrepancies found for $\bar{S}_{\text{MAX},j}$ than $\bar{S}_{\text{MSY},j}$). $\bar{S}_{\text{MAX},j}$ was

60% larger for state-space models with time-constant maturity schedules (approximately 15,500) than models with time-varying maturity (approximately 9,500; Table 4.7). In terms of the aggregate mixed stock escapement expected to produce maximum sustained yield (S_{MSY}), regression approaches suggested much smaller escapements were necessary than the state-space models; by a margin of 12,000 – 15,000 fish (~50%; Table 4.7). In comparing the uncertainty in S_{MSY} among state-space models, those with time-constant maturity schedules had much more uncertainty than the time-varying models (*i.e.*, wider credible limits; Table 4.7).

In comparing the state-space model estimates from these 10 substocks to the complete 13 substocks, the average substock-specific size metrics were fairly similar, indicating the stocks with insufficient data for regression were missing at random in this regard (compare the rows corresponding to $\bar{S}_{\mathrm{eq},j}$, $\bar{S}_{\mathrm{MAX},j}$, and $\bar{S}_{\mathrm{MSY},j}$ between Tables 4.7 and 4.8). In terms of the aggregate reference point S_{MSY}, the estimates were quite different – which should be expected given some substocks' abundance and production were excluded from Table 4.7 and included in Table 4.8. With all substocks included, the estimate was approximately 40,000 compared to approximately 30,000 when they were excluded – indicating escapement from the Holitna, Kisaralik, and Oskawalik should make up approximately 25% of the escapement amongst the substocks included in this analysis if the management objective was to maximize long-term yield.

As illustrated in Figure 4.1, the relationship between substock size and productivity matters for policy conclusions about trade-offs. The relationship between substock size (represented by $S_{\mathrm{eq},j}$) *versus* productivity (represented by $U_{\mathrm{MSY},j}$) was examined for the Kuskokwim River substocks (Figure 4.8). Both regression approaches suggested reasonably strong patterns between substock size and productivity (Pearson's correlation r coefficient of 0.42 and 0.54 for the independent and mixed-effects approaches, respectively, though $p > 0.05$ for both approaches). The state-space models generally suggested weaker or

absent relationships between substock size and productivity, except for SSM-VM, which had an r value of 0.56 that was statistically significant ($p < 0.05$). The state-space models suggested that upper-river substocks were mostly smaller than average, but were made of both productive and unproductive substocks, whereas most of the larger substocks were suggested to be located in the middle river (relatively speaking; Figure 4.8).

In terms of recruitment variability, the regression-based approaches suggested that the average standard deviation of the lognormal distribution that describes randomness in the recruitment process ($\bar{\sigma}_{R,j}$) was 0.78 and 0.52 for the independent and mixed-effects versions, respectively (Table 4.7). The state-space models estimated that the average amount of recruitment process variability by substock was higher: point estimates ranged from 0.85 – 1.11 (Table 4.7). The models with time-varying maturity had estimates on the lower end of this range, which could be explained by the inclusion of additional process variability in maturity that could describe variability in the data.

Unlike the regression models, the state-space models estimated the degree of covariance between substock recruitment residuals. All state-space models estimated a moderate amount of correlation in recruitment variance (*i.e.*, synchrony) between two average substocks ($\bar{\rho}_{i,j}$): point estimates ranged between 0.18 and 0.28, and none of the models suggested 95% credible limits that encompassed zero, indicating evidence of some degree of positive synchrony among the Kuskokwim River substocks. The simple variance models (SSM-vm and SSM-vM) estimated a single correlation parameter, whereas the complex versions (SSM-Vm and SSM-VM) estimated a unique value for each substock combination. One might expect that substock pairs belonging in the same region might show higher degrees of synchrony than substock pairs in different regions, though this analysis suggested this was not necessarily the case (the state-space models were unaware of the spatial relations among substocks). Large correlations (*e.g.*, >0.5) were found between pairs of lower *versus* upper river substocks, lower *versus* lower, and upper *versus* upper substocks (Figure 4.9). Relatively few large

131

correlations were found among middle river substocks with other substocks, though both SSM-Vm and SSM-VM suggested that the Holokuk and Oskawalik substocks have highly synchronous recruitment dynamics, which is interesting given their close proximity (Figure 4.2; substocks #5 and #6, respectively). Conversely, all models suggested the Holitna and Kogrukluk substocks have little synchrony, and they fall within the same subdrainage (Figure 4.2; substocks #7 and #8, respectively). Most correlations were positive, especially those that were large in magnitude (Figure 4.9). A notable exception was the correlation between the Kisaralik substock and the Holokuk and Oskawalik substocks: their dynamics were suggested to be largely opposite.

A visualization of the consequences of this correlation structure for the recruitment residual time series is shown in Figure 4.10. As would be expected, the models with only one correlation parameter (SSM-vm and SSM-vM) showed greater synchrony in time trends among substocks (especially SSM-vM) than the models with complex correlation structure. Synchrony was greatest in the early portion of the time series for all models, likely resulting from fewer substocks with observed data in this period. There was a weak cycling tendency of favorable and poor recruitment conditions (Figure 4.10), which is consistent with the degree of estimated lag-1 auto-correlation in each time series (ϕ; range across state-space models: $0.27 - 0.32$; no credible intervals encompassed 0; Table 4.8).

The only difference in mean maturity schedule estimates came in the comparison between models with and without time-varying maturity (and the differences were slight; Table 4.8). The models that allowed time-varying maturity suggested that the inter-brood year variability was quite high (the value of the Dirichlet inverse dispersion parameter $D \approx 18$; Table 4.8).

4.3.1.4   Trade-off comparisons

With increasing mixed-stock exploitation rates, the aggregate equilibrium escapement declined, but did so more rapidly for the state-space models than the regression models (Figure

4.11). State-space models that included time-varying maturity suggested higher equilibrium escapement and harvest would be available at most exploitation rates than models with time-constant maturity. Regression models suggested that MSY was much larger and occurred at a higher exploitation rate than for the state-space models, as would be expected based on comparing the estimated population dynamics parameters between these methods (Table 4.7). In terms of substock diversity, the time-varying maturity models suggested fewer substocks would be overfished or trending towards extirpation at low exploitation rates than the time-constant maturity models, and the regression approaches suggested even more optimistic conclusions (Figure 4.11).

Another and more direct way of visualizing the trade-off between harvest and biodiversity suggested the trade-off conclusions are nearly identical among state-space models in terms of substock overfishing and slight differences in terms of substocks trending towards extirpation (Figure 4.12). The state-space models suggested that in order to maximize yields, approximately 50% of the substocks would be overfished and approximately a 20% would be trending towards extinction (Figure 4.12). The regression approaches suggested a more pessimistic conclusion: approximately 80% of substocks would be overfished at the mixed-stock MSY. These patterns were the same even if only 60% of MSY was desired, though less exaggerated (Figure 4.12).

4.3.1.5   Sensitivity analyses

The alternative vulnerability schedule (v_j) based on the spatial distribution of fishing households resulted in lower river substocks (those spawning in the Kwethluk, Kisaralik, and Tuluksak rivers) having the lowest v_j ranging between $0.7 - 0.8$, the middle river substocks ranging between 0.9 and 0.95, and those in the upper river between 0.95 and 1 (Figure 4.13). When these v_j terms were incorporated in the state-space models, most changes in substock specific $U_{\text{MSY},j}$ and $S_{\text{eq},j}$ were small ($\pm 10\%$; Figure 4.14), and most (eight of 13) substocks

showed increases in $U_{\text{MSY},j}$. Changes in $U_{\text{MSY},j}$ occurred randomly with respect to changes in vulnerability: some substocks in the three regions showed both increases and decreases (Figure 4.14). Changes with respect to substock size ($S_{\text{eq},j}$) showed more of a pattern: lower river substocks (*i.e.*, those that became less vulnerable in this sensitivity analysis) tended to become smaller by $5 - 15\%$, whereas upper river substocks showed increases of $5 - 25\%$. Middle river substocks showed a mix of increases and decreases in $S_{\text{eq},j}$ (Figure 4.14). Despite these changes in substock-specific estimates, derived biological reference points for the aggregate mixed-stock showed a high degree of similarity: escapement-related quantities were ~3,000 fish ($<5\%$ change) smaller and harvest-related quantities were ~3,000 fish (~10-15%) larger for the alternative vulnerability assumption, however the conclusions about substock diversity at MSY were nearly identical (Table 4.9).

The alternative age composition weighting resulted in more substantial differences from the default case. Twelve of the 13 substocks showed increases in $U_{\text{MSY},j}$, with two substocks showing increases of approximately 100%. Most substock-specific $S_{\text{eq},j}$ were estimated to be smaller when the alternative scheme was used, with eight of 13 substocks showing decreases between 10 and 35%. (Figure 4.15). Derived biological reference points for the aggregate mixed-stock substantially differed as well: escapement-related quantities were ~6,000 – 10,000 fish smaller (~10-15% change) and harvest-related quantities were ~6,000 – 12,000 fish larger (~35% change) for the alternative age composition weighting scheme, however the conclusions about substock biodiversity at MSY still were nearly identical (Table 4.10).

4.3.2 Simulation-estimation trials

State-space models took between 1.3 and 2.9 days to fit on average, with longer run times associated with more complex models (Table 4.11) and the regression-based models took less than an hour in all cases. The state-space models fitted successfully to the majority of simulated data sets (~140 out of 160) and the regression-based models were fitted successfully

in all cases (Table 4.11). No attempts were made to try different initial values for the ~20 data sets that failed in fitting for the state-space models, though because they were generally the same data sets that failed across state-space models, this was likely a characteristic of the random data sets. MCMC diagnostics suggested that the vast majority of models converged and that enough samples were retained for adequate posterior inference (Table 4.12).

Regression models were found to systematically overestimate the substock-specific quantities of α_j and $U_{\mathrm{MSY},j}$ and in some cases they produced wildly erroneous estimates. The mixed-effect model was more accurate than the method that fitted independent regressions to each substock (Figure 4.16). State-space models far more accurately and precisely estimated these productivity-based quantities than the regression-approaches, though there was still a slight positive bias (median proportional error $\approx 5\%$; Figure 4.16). Additionally, regression approaches tended to underestimate $S_{\mathrm{MSY},j}$ more than the state-space models. All state-space models tended to overestimate $\sigma_{R,j}$ by approximately 5% regardless of the assumed covariance structure, though the degree of serial auto-correlation (ϕ) was accurately estimated. The individual state-space models showed essentially no differences for these quantities (Figure 4.16). However, models with simple covariance structure tended to overestimate the correlation among substocks and *vice versa* for the complex covariance structure.

The state-space model was a largely unbiased estimator of abundance-related states ($R_{y,j}$, $S_{t,j}$, and H_t). The early portion of the time series had a slight tendency to be overestimated (Figure 4.17), likely a result of the assumption that data collection began when the substocks were unfished.

Just as for the substock-specific quantities, the regression-based methods provided generally poorer estimates of mixed-stock biological reference points than state-space models and there was no loss in performance with state-space model complexity (Figure 4.18). The mixed-effect regression produced more positively biased estimates of $U_{0.1}^*$ and $U_{0.3}^*$ than the independent regression approach, but this pattern switched for $U_{0.5}^*$ and U_{MSY}, likely as a

135

result of the estimated shape of the distribution of substock productivity. The mixed-effect version was less dispersed, meaning that productivities in the lower tail would have been closer to the mean (*i.e.*, larger) than for the independent regression approach. State-space models tended to produce slight underestimates of $U_{0.1}^*$ and $U_{0.3}^*$ and slight over estimates of $U_{0.5}^*$ and U_{MSY} (Figure 4.18).

Credible interval coverage was better for the state-space models than for regression approaches as well. For substock-specific parameters, the regression approaches had lower coverage than the state-space models and the models that had the complex recruitment variance structure had more parameters close to the optimal level of 95% Table 4.13). All state-space models had low coverage for π which resulted from highly narrow credible intervals, not from inaccurate estimates, though complex maturity models did have slightly better coverage (16% *versus* 11%; Table 4.13). In terms of mixed-stock biological reference points, the state-space model provided much better coverage than the regression approaches, particularly for exploitation rate-based points (Table 4.13). All state-space models exhibited poor coverage for abundance-related states (49 – 66%; Table 4.13).

4.4   Discussion

The state-space models presented here provide a novel extension of the age-structured state-space spawner-recruit analytical framework, which has increasingly been applied to single stocks (*e.g.*, Su and Peterman 2012; Hamazaki *et al.* 2012; Fleischman *et al.* 2013; Staton *et al.* 2017b; DeFilippo *et al.* 2018), to the multi-stock realm. The state-space model was shown to (1) have much less bias and better coverage in key management quantities than regression-based approaches, (2) be robust to structural uncertainty in assumed recruitment covariance and maturity variability, (3) provide good fits to the data, and (4) make more full use of the available data for policy and ecological conclusions. Though it was developed and simulation-tested in the context of Kuskokwim River Chinook salmon, the multi-stock

state-space framework presented here should be general enough to be applied to other systems with similar (and possibly dissimilar) properties and data availability.

The simulation-trials illustrated that the state-space model performed superiorly to regression-based approaches, regardless of the assumptions made regarding covariance structure and variability in maturity. Furthermore, the directionality of the regression biases were consistent with expectations from the time-series bias (Walters 1985): positive biases in $U_{\mathrm{MSY},j}$ and downward biases in $S_{\mathrm{MSY},j}$, as well as for the respective aggregate mixed-stock quantities. These biases all but disappeared in comparison for the state-space models for $U_{\mathrm{MSY},j}$ and to a lesser degree for $S_{\mathrm{MSY},j}$. These results speak strongly in favor of the use of the state-space model over the regression-based approaches assessed here. The superior performance of state-space models was likely a result of its ability to (1) explicitly account for the time-series properties in the data, (2) parse observation from process uncertainty, and (3) make more full use of the available data.

With respect to the fuller use of the available data, there were 35 brood years for each substock in which it was possible to jointly observe spawners and recruits across all four ages (if all calendar years between 1976 and 2017 were monitored). The average Kuskokwim River substock fitted using the regression approaches had 17% of the possible observations because partially observed recruitment events were not considered (see Appendix D.5). Conversely, the average substock fitted using the state-space model had 31% of possibly observable pairs where the recruitment was observed for three out of four ages, 39% with two ages out of four, and 42% for one age out of the possible four observed. Ignoring whether recruitment and escapement were observed jointly, the average substock fitted using the state-space model had 48%, 65%, and 76% of possibly observable recruitments observed for the same ages combinations, respectively. Clearly, the use of the regression approaches resulted in a severe loss of information. It could be argued that the rule I employed to only use completely observed recruitment pairs for fitting the regression approaches was too strict and that

137

reasonable methods to impute missing observations could be devised. While this may be true, the state-space model provides a comprehensive, rational, and rigorous method to completely reconstruct the brood tables with latent states informed with partial information by fitting to solely observed data.

Although the simulation results suggested the state-space model is an unbiased estimator, it was assessed under relatively limited conditions. Su and Peterman (2012) illustrated that state-space models can still show bias under some combinations of measurement error, intrinsic productivity, and fishing intensity. I attempted to evaluate performance across a range of true parameters by randomly sampling leading parameters from the joint posterior from one of the empirical model fits. Though it is possible the model could perform more poorly if (1) substock size and productivity were more or less heterogeneous than assumed, (2) fewer substocks were used, (3) available data time series were much shorter or sparser, or (4) the magnitude of observation error was incorrectly assumed. All of these scenarios remain exciting avenues for future research, but were beyond the scope of this study. As it currently stands, this analysis suggests that it is reasonable to conclude that the state-space model developed here can be an appropriate estimator for mixed-stock fishery data, though future applications to specific cases should be simulation-tested if the properties of the system and available data differ significantly from the Kuskokwim and the operating model and simulated sampling scheme designed roughly off of it.

The policy and trade-off conclusions from the state-space model were generally robust to an alternative assumption regarding relative vulnerability of the substocks to harvest, although substock-specific estimates did change moderately (primarily $U_{\mathrm{MSY},j}$). The alternative vulnerability vector used was a first attempt at assessing sensitivity to the assumption that all substocks are equally vulnerable; other more complex approaches to determining this vector could also be assessed in the future. For example, the fishery has historically been focused in on the early portion of the Chinook salmon run in the Kuskokwim River

(Hamazaki 2008), and upper river substocks have been illustrated to arrive earliest in the summer migration (Smith and Liller 2017a,b). This indicates that upper river substocks may be even more vulnerable than what was captured by the vector I used, which was based solely on the spatial distribution of harvesters. Although no data sources were available to directly estimate the vulnerability vector for the Kuskokwim River data, other systems may have these data that could be incorporated, particularly those with precise genetic stock identification programs. One would expect larger discrepancies in policy conclusions to arise when vulnerability covaries more strongly with either substock size or substock productivity; coupled with methods to incorporate information on substock-specific harvest data, this provides an provides an interesting avenue for future research.

The conclusions from the state-space models were less robust to alterations in the assumed weighting of age composition data. Specifically, policy recommendations became more aggressive when age composition data received less weight: higher estimates of U_{MSY} and lower S_{MSY}. It is unsurprising that the estimates changed, as this often happens in stock assessment models when the assumed data weighting structure is altered (Hulson *et al.* 2011), but it is unclear as to why weakening the confidence in the age data had the effect of increasing perceived substock productivity. Regardless of the cause, this finding suggests that policy conclusions may be conditional on the weighting of the data, and that careful thought should be given to the appropriate weighting scheme. It is likely that the optimal weighting scheme falls somewhere between the two schemes assessed here, given that effective sample size is nearly always less than the true sample size due to violations to the multinomial sampling distribution (*e.g.*, sampled individuals show similarities that result in clustering, non-independence, and overdispersion; Maunder 2011).

Regarding state-space model complexity, the important findings were (1) there was no loss in estimation performance with increasing model complexity, (2) credible interval coverage of the complex recruitment covariance models was better than that of the models

with simple structures for substock-level parameters of interest, and (3) the time-varying maturity models never resulted in wildly large estimates of escapement or harvest as the time-constant maturity models sometimes did for the Kuskokwim data. Based on these findings, it seems that the most complex model is most appropriate, however, as previously stated, the simulation-trials were limited in the scope of biological and sampling scenarios considered, and as such the appropriate model may change if applied to other systems with differing characteristics.

This conclusion is contrary to the traditional dogma of assessment model complexity and management performance. Walters and Martell (2004) advise that more complex models may provide more accurate estimates of management quantities, but that their uncertainty will be much greater rendering them less useful for setting harvest policies (see Figure 5.2 and the corresponding discussion therein). These claims have been supported through closed-loop evaluations that have shown simple models known to be wrong but that give conservative advice can provide better management outcomes than complex models that better approximate the true model (*e.g.*, Hilborn 1979; Ludwig and Walters 1985). The simplest model (fewest freely estimated parameters) evaluated here was the mixed-effect regression approach. Although it did provide more conservative and more confident advice than the independent regression approach, the simulation-trials showed that it was biased with respect to U_{MSY} and S_{MSY} in the direction that would lead to more aggressive than optimal harvest policies relative to the state-space models.

In some cases, harvest on strong and highly profitable fisheries has been severely curtailed in the name of conserving a few small and unproductive substocks. Walters *et al.* (2018) discuss an example of the British Columbia commercial salmon fishery, and illustrate that a drastic decline in harvest beginning in the early 1990s and continuing to the present was (in part) a result of intentional reductions in exploitation rates intended to minimize the risk of extinction of a few small and unproductive stocks. The authors argue that fisheries managers

140

have not adequately addressed the harvest-biodiversity trade-off in their decision-making processes, and have instead focused on managing for the weakest stocks in the portfolio. Among the authors' four (mostly controversial) recommendations to address this situation, first on their list is for managers to conduct trade-off analyses so that costs and benefits to both fishery and conservation interests can be more fully addressed in decision-making. The state-space approach developed here shows promise for informing these policy analyses. Particularly for those that involve closed-loop stochastic simulation, the state-space model provides much richer biological information to populate the operating models and understand the strength of the portfolio effect (which is inversely related to the magnitude of shared recruitment trends; Schindler *et al.* 2010, 2015). The model not only provides estimates of leading parameters (α_j and β_j), but also estimates of the extent to which recruitment anomalies are shared among substocks and the strength of serial auto-correlation in these time series, all of which would be valuable in populating operating models for policy evaluation. Furthermore, it is possible to use the estimated states of recruitment and spawner abundance and recruitment anomalies at the end of the time series to populate forward simulations from the present to determine which policies might be most likely to achieve short-term objectives in addition to those more focused on the long-term.

TABLE 4.1: Description of the various indices used in the description of the state-space model. n_t is the number of years observed for the most data-rich stock.

Index	Meaning	Dimensions
y	Brood year index; year in which fish were spawned	$n_y = n_t + n_a - 1$
t	Calendar year index; year in which observations are made	n_t
j	Substock index	n_j
a	Age index; $a = 1$ is the first age; $a = n_a$ is the last age	n_a

TABLE 4.2: Symbology used in presenting the state-space models.

Symbol	Description
Dimensional Constants	
n_y	Number of brood years
n_t	Number of calendar years for the substock with the longest data time series
n_j	Number of substocks
n_a	Number of possible ages of maturation
a_{min}	The first age recruits can mature
a_{max}	The last age recruits can mature
Parameters	
α_j[a]	Maximum recruits per spawner for substock j
β_j[a]	Capacity parameter for substock j; inverse of $S_{\mathrm{MAX},j}$
$\sigma^2_{R,j}$	Recruitment white noise process variance for substock j
$\rho_{i,j}$	Correlation in process variance between subtocks i and j
Σ_R	Recruitment white noise process covariance matrix
ϕ	Lag-1 serial autocorrelation coefficient
$\omega_{y,j}$	Serially autocorrelated portion of recruitment process anomalies (residuals)
π_a	Mean probability a juvenile matures at age a
D[b]	Dirichlet dispersion parameter for brood year-specific maturity schedules
$p_{y,a}$[b]	Probability a juvenile belonging to brood year y matures at age a
U_t[c]	Exploitation rate experienced by fully vulnerable substocks in calendar year t
v_j[c]	Relative vulnerability term for substock j
Biological Reference Points	
$U_{\mathrm{MSY},j}$	Exploitation rate expected to produce MSY for substock j
U_{MSY}	Exploitation rate expected to produce MSY for the mixed-stock aggregate
$S_{\mathrm{MSY},j}$	Spawner abundance expected to produce MSY for substock j
S_{MSY}	Spawner abundance expected to produce MSY for the mixed-stock aggregate
$S_{\mathrm{MAX},j}$	Spawner abundance expected to produce maximum recruitment for substock j
$S_{\mathrm{eq},j}$	Spawner abundance expected to produce the same number of recruits for substock j
S_p^*[d]	Mixed-stock escapement expected to result in no greater than $p \cdot 100\%$ of substocks overfished
H_p^*[d]	Same as S_p^*, but for mixed-stock harvest
U_p^*[d]	Same as S_p^*, but for mixed-stock exploitation rate
$p_{EX,\mathrm{MSY}}$[d]	Fraction of substocks expected to be overfished at MSY
$p_{OF,\mathrm{MSY}}$[e]	Fraction of substocks expected to be trending towards extinction
States	
$\dot{R}_{y,j}$	Deterministic (expected) recruitment in brood year y for substock j
$R_{y,j}$	Realized latent (true) recruitment in brood year y for substock j

Symbol	Description
$N_{t,j}$	Run abundance returning to spawn in calendar year t for substock j
$S_{t,j}$	Spawner abundance in calendar year t for substock j
H_t	Mixed-stock aggregate harvest in calendar year t
$q_{t,a,j}$	Fraction of the run mature at age a in year t for substock j

[a] In the state-space models, $U_{\mathrm{MSY},j}$ and $S_{\mathrm{MSY},j}$ were estimated as leading parameters, α_j and β_j were derived from them using equations found in Schnute and Kronlund (2002).

[b] Used only in complex maturity models: SSM-vM and SSM-VM. For simple maturity models, $p_{y,a}$ took the value π_a.

[c] In the default case, all substocks were assumed to be fully vulnerable, v_j was used in a sensitivity analysis to this assumption.

[d] Overfished is defined here as the case where the mixed-stock or any given substock is fished with exploitation rate greater than is expected to produce MSY.

[e] Trending towards extirpation is defined here as the case where expected equilibrium escapement is less than or equal to zero.

TABLE 4.3: Summary of evaluated models in this analysis. Regression models are described in Section 4.2.1.1 and state-space models are described in Section 4.2.1.2

Model	n_j	Unique σ_j	AR(1)	Recruitment Covariance	Time-Varying Maturity
Regression-Based Models					
LM	10	Yes	No	None	Yes
LME	10	No	No	None	Yes
State-Space Models					
vm	13	No	Yes	Single ρ bounded by [-0.05 − 1)	No
Vm	13	Yes	Yes	Unique $\rho_{i,j}$	No
vM	13	No	Yes	Same as vm	Yes
VM	13	Yes	Yes	Same as Vm	Yes

TABLE 4.4: Prior distributions used for regression-based spawner-recruit parameters as described in Section 4.2.1.1. Prior distributions were identical for the empirical data and simulation-based analyses.

Parameter	Prior
Independent Regression Models	
$\log(\alpha_j)$	Uniform(0, 5)
β_j	Uniform(0, 1)
$\sigma_{R,j}$	Uniform(0, 5)
Mixed-Effect Regression Model	
$\log(\alpha)$	Uniform(0, 10)
σ_α	Uniform(0, 10)
β_j	Uniform(0, 1)
σ_R	Uniform(0, 5)

TABLE 4.5: Prior distributions used for all model parameters in the state-space models. In all cases, priors were selected to be minimally informative while still preventing the sampler from exploring highly unlikely areas of the parameter space. Differences among versions of the state-space model (*e.g.*, SSM-vm and SSM-Vm; Table 4.3) are described by footnotes. Prior distributions were identical for the empirical data and simulation-based analyses.

Parameter	Prior	Description
$U_{\mathrm{MSY},j}$	Uniform(0.01, 0.99)	Exploitation rate that produces MSY
$S_{\mathrm{MSY},j}$	Lognormal(0, 0.001)	Spawner abundance that produces MSY
ϕ	Uniform(-0.99, 0.99)	Lag-1 auto-correlation coefficient
Σ_R^{-1}[a]	Wishart(R, $n_j + 1$)	Inverse covariance matrix for white-noise recruitment process variability
σ_R[b]	Uniform(0, 2)	White-noise recruitment process standard deviation
ρ[b]	Uniform(-0.05, 1)	Correlation in recruitment process variability among substocks
π	Dirichlet($\alpha = [1,1,1,1]$)	Average probability of maturing at each age
$D^{-0.5}$[c]	Uniform(0.03, 1)	Dispersion of brood year-specific maturity schedules
p_y[c]	Dirichlet($\alpha = \pi \cdot D$)	Brood year-specific probability of maturing at each age
U_t	Beta(1,1)	Annual exploitation rate of fully vulnerable substocks

[a] Only for SSM-Vm and SSM-VM

[b] Only for SSM-vm and SSM-vM, Σ_R was constructed using σ_R and ρ as described at the end of Section 4.2.1.2.

[c] Only for SSM-vM and SSM-VM, all p_y took on π for SSM-vm and SSM-Vm.

TABLE 4.6: Dimensions for the Markov Chain Monte Carlo algorithms used in this analysis. Note that the state-space models were sampled much more intensively than the regression models – this was to ensure adequate convergence and effective sample size for inference. Fewer chains were used for the simulation analysis to maximize High Performance Computing efficiency. MCMC diagnostics indicated these settings were adequate for reliable inference; the state-space models fitted to the empirical data were over-sampled to ensure this.

	Regression Models		State-Space Models	
	Empirical	Simulation	Empirical	Simulation
Burn-in	20,000	20,000	50,000	50,000
Post Burn-in	100,000	200,000	800,000	600,000
Thin Interval	50	50	400	100
Chains	10	5	10	5
Total	**1,200,000**	**1,100,000**	**8,500,000**	**3,250,000**
Saved	**20,000**	**20,000**	**20,000**	**30,000**

TABLE 4.7: Estimated population parameters for Kuskokwim River Chinook salmon compared among assessment models, including the regression-based estimators. Only 10 of the 13 substocks had enough data to fit the linear regression model, the three missing stocks were discarded in the calculation of the summaries presented for the state-space models. Numbers shown are posterior medians with 95% credible limits in parentheses. Quantities with a bar and a j subscript denote averages over substocks, those with no subscript are the appropriate reference points for the aggregate of the 10 substocks.

| Parameter | Regression-Based Models | | State-Space Models | | | |
	LM	LME	vm	Vm	vM	VM
$\bar{\alpha}_j$	7.74 $(4.47-20.22)$	4.63 $(3.16-7.45)$	2.98 $(1.94-5.03)$	2.7 $(1.91-3.99)$	3.3 $(2.25-5.01)$	2.82 $(1.99-4.03)$
$\bar{U}_{\mathrm{MSY},j}$	0.58 $(0.47-0.68)$	0.51 $(0.41-0.6)$	0.39 $(0.26-0.52)$	0.37 $(0.25-0.49)$	0.45 $(0.32-0.56)$	0.4 $(0.28-0.51)$
$\bar{S}_{\mathrm{MSY},j}$	2,500 $(1,900-5,700)$	2,600 $(2,100-3,900)$	3,700 $(2,300-7,000)$	4,100 $(2,500-7,700)$	3,500 $(2,700-5,500)$	3,700 $(2,800-5,600)$
$\bar{S}_{\mathrm{MAX},j}$	4,900 $(3,200-22,400)$	5,700 $(3,900-22,900)$	15,900 $(7,300-77,600)$	15,000 $(7,500-63,200)$	9,500 $(5,900-40,800)$	9,900 $(6,100-27,400)$
$\bar{S}_{\mathrm{eq},j}$	6,800 $(5,300-14,400)$	6,700 $(4,200-9,400)$	8,400 $(5,300-15,500)$	9,200 $(5,600-17,000)$	8,400 $(6,400-12,400)$	8,600 $(6,600-12,700)$
$\bar{\sigma}_{R,j}$	0.79 $(0.56-1.14)$	0.53 $(0.44-0.65)$	1.02 $(0.9-1.18)$	1.08 $(0.9-1.33)$	0.85 $(0.75-0.98)$	0.85 $(0.73-1.01)$
U_{MSY}	0.78 $(0.59-0.97)$	0.68 $(0.49-0.84)$	0.43 $(0.2-0.67)$	0.41 $(0.21-0.62)$	0.48 $(0.32-0.63)$	0.45 $(0.29-0.6)$
S_{MSY}	16,200 $(4,000-36,300)$	18,900 $(8,500-32,000)$	25,600 $(8,700-63,900)$	31,200 $(10,000-71,600)$	29,700 $(17,500-45,500)$	33,800 $(20,000-51,300)$

TABLE 4.8: Estimated population parameters and biological reference points for Kuskokwim River Chinook salmon compared among the four evaluated versions of the state-space assessment model assessment models. Unlike in Table 4.7, all 13 of the substocks were included in the caluclation of these summaries. Numbers shown are posterior medians with 95% credible intervals in parentheses. Quantities with a bar and a j subscript denote averages over substock-specific parameters. Reference points with no subscript are the appropriate reference points for the aggregate of the 13 substocks included.

Parameter	vm	Vm	vM	VM
$\bar{\alpha}_j$	2.78 $(1.91 - 4.43)$	2.48 $(1.82 - 3.57)$	3.22 $(2.27 - 4.73)$	2.81 $(2.02 - 4.19)$
$\bar{U}_{\mathrm{MSY},j}$	0.37 $(0.25 - 0.49)$	0.34 $(0.23 - 0.45)$	0.44 $(0.33 - 0.55)$	0.39 $(0.28 - 0.5)$
$\bar{S}_{\mathrm{MSY},j}$	4,100 $(2,700 - 6,900)$	4,200 $(2,600 - 7,900)$	3,800 $(3,000 - 6,000)$	3,600 $(2,800 - 5,900)$
$\bar{S}_{\mathrm{MAX},j}$	16,400 $(8,900 - 65,000)$	18,100 $(9,200 - 73,500)$	10,500 $(6,400 - 45,200)$	10,200 $(6,200 - 37,600)$
$\bar{S}_{\mathrm{eq},j}$	9,200 $(6,100 - 15,200)$	9,300 $(5,800 - 17,300)$	9,000 $(6,900 - 13,500)$	8,500 $(6,600 - 13,300)$
$\bar{\sigma}_{R,j}$	1.02 $(0.9 - 1.18)$	1.23 $(1.05 - 1.48)$	0.85 $(0.75 - 0.98)$	1 $(0.85 - 1.19)$
$\bar{\rho}_{i,j}$	0.21 $(0.1 - 0.36)$	0.18 $(0.09 - 0.31)$	0.28 $(0.15 - 0.44)$	0.18 $(0.09 - 0.32)$
U_{MSY}	0.39 $(0.22 - 0.59)$	0.38 $(0.2 - 0.59)$	0.48 $(0.32 - 0.62)$	0.46 $(0.3 - 0.64)$
S_{MSY}	41,100 $(15,800 - 80,600)$	40,400 $(13,500 - 92,400)$	41,500 $(25,200 - 64,900)$	41,600 $(21,500 - 64,600)$
ϕ	0.32 $(0.12 - 0.49)$	0.32 $(0.14 - 0.49)$	0.27 $(0.06 - 0.44)$	0.3 $(0.1 - 0.48)$
π_1	0.27 $(0.265 - 0.275)$	0.271 $(0.266 - 0.276)$	0.232 $(0.203 - 0.263)$	0.233 $(0.204 - 0.264)$
π_2	0.379 $(0.374 - 0.385)$	0.379 $(0.373 - 0.384)$	0.371 $(0.336 - 0.406)$	0.372 $(0.337 - 0.406)$
π_3	0.327 $(0.322 - 0.333)$	0.327 $(0.321 - 0.332)$	0.359 $(0.325 - 0.394)$	0.358 $(0.324 - 0.392)$
π_4	0.024 $(0.022 - 0.025)$	0.024 $(0.022 - 0.025)$	0.037 $(0.027 - 0.049)$	0.036 $(0.027 - 0.049)$
D	—	—	17.94 $(13.37 - 23.63)$	18.22 $(13.55 - 24.03)$

TABLE 4.9: Comparison of management quantities obtained from the sensitivity analysis of the state-space model with regards to substock vulnerability. The default case had all substocks equally vulnerable; the alternative case is shown in Figure 4.13, and reference points were calculated using the same vulnerability schedules as used in estimation. Estimates shown are from SSM-VM only, the other three models showed similar differences. The biodiversity metrics are expressed as the proportion of substocks expected to be overfished or trending towards extirpation at the mixed-stock MSY.

	v_j Schedule	
Quantity	Default	Alternative
Escapement		
$S^*_{0.1}$	85,300 $(61,200 - 130,300)$	83,500 $(59,500 - 129,000)$
$S^*_{0.3}$	68,800 $(49,700 - 102,400)$	66,700 $(47,200 - 99,700)$
$S^*_{0.5}$	51,300 $(36,300 - 72,600)$	48,800 $(33,400 - 71,900)$
S_{MSY}	41,600 $(21,500 - 64,600)$	41,800 $(23,300 - 67,500)$
Harvest		
$H^*_{0.1}$	17,800 $(4,100 - 39,600)$	20,600 $(5,000 - 45,500)$
$H^*_{0.3}$	27,100 $(11,300 - 49,100)$	30,300 $(12,500 - 55,800)$
$H^*_{0.5}$	33,100 $(16,400 - 56,800)$	36,100 $(17,800 - 62,800)$
MSY	34,900 $(17,900 - 59,600)$	37,500 $(19,000 - 64,800)$
Exploitation Rate		
$U^*_{0.1}$	0.17 $(0.04 - 0.34)$	0.2 $(0.05 - 0.38)$
$U^*_{0.3}$	0.28 $(0.13 - 0.42)$	0.32 $(0.16 - 0.48)$
$U^*_{0.5}$	0.39 $(0.25 - 0.52)$	0.44 $(0.29 - 0.59)$
U_{MSY}	0.46 $(0.3 - 0.64)$	0.5 $(0.33 - 0.68)$
$\bar{U}_{full,t}$	0.32 $(0.42 - 0.35)$	0.35 $(0.48 - 0.41)$
Biodiversity		
$p_{OF,MSY}$	0.62 $(0.38 - 0.85)$	0.62 $(0.38 - 0.85)$
$p_{EX,MSY}$	0.23 $(0 - 0.62)$	0.23 $(0 - 0.54)$

151

TABLE 4.10: Comparison of management quantities obtained from the sensitivity analysis of the state-space model with regards to weighting of age composition data. Estimates shown are from SSM-VM only, the other three models showed similar differences. The biodiversity metrics are expressed as the proportion of substocks expected to be overfished or trending towards extirpation at the mixed-stock MSY.

| Quantity | $ESS_{t,j}$ Scheme | |
	Default	Alternative
Escapement		
$S^*_{0.1}$	85,300 (61,200 − 130,300)	75,600 (56,700 − 104,900)
$S^*_{0.3}$	68,800 (49,700 − 102,400)	59,900 (44,800 − 81,400)
$S^*_{0.5}$	51,300 (36,300 − 72,600)	42,700 (31,200 − 56,600)
S_{MSY}	41,600 (21,500 − 64,600)	35,000 (21,800 − 48,100)
Harvest		
$H^*_{0.1}$	17,800 (4,100 − 39,600)	24,100 (6,200 − 48,100)
$H^*_{0.3}$	27,100 (11,300 − 49,100)	36,200 (16,400 − 59,100)
$H^*_{0.5}$	33,100 (16,400 − 56,800)	44,700 (24,700 − 68,800)
MSY	34,900 (17,900 − 59,600)	46,600 (26,200 − 71,900)
Exploitation Rate		
$U^*_{0.1}$	0.17 (0.04 − 0.34)	0.24 (0.06 − 0.43)
$U^*_{0.3}$	0.28 (0.13 − 0.42)	0.38 (0.19 − 0.53)
$U^*_{0.5}$	0.39 (0.25 − 0.52)	0.51 (0.36 − 0.63)
U_{MSY}	0.46 (0.3 − 0.64)	0.57 (0.41 − 0.71)
$\bar{U}_{full,t}$	0.32 (0.42 − 0.35)	0.32 (0.41 − 0.34)
Biodiversity		
$p_{OF,\mathrm{MSY}}$	0.62 (0.38 − 0.85)	0.62 (0.38 − 0.85)
$p_{EX,\mathrm{MSY}}$	0.23 (0 − 0.62)	0.23 (0.08 − 0.54)

TABLE 4.11: Number of successful model fits and elapsed time for each multi-stock spawner-recruit analysis method from the simulation-estimation exercise. All regression-based methods were fitted in a single JAGS model. Subscripts denote differences in time units.

	# Data Sets		Elapsed Time		
	Attempted	Successful	Minimum	Mean	Maximum
LM + LME[a]	160	160	4.8	10.9	21.6
SSM-vm[b]	160	137	0.8	1.3	2.8
SSM-Vm[b]	160	136	1.2	1.8	3.5
SSM-vM[b]	160	136	1.2	2	4.1
SSM-VM[b]	160	136	1.8	2.9	5.1

Time Units
[a] Minutes
[b] Days

TABLE 4.12: MCMC diagnostic summaries from the simulation-estimation trials. Numbers in the cells represent the percentage of all estimated values that did not meet a diagnostic threshold. For the Brooks-Gelman-Rubin statistic, failure was defined as having a value of 1.1 or greater. For the effective samples, having fewer than 3,000 was considered a failure. According to the diagnostic by Raftery and Lewis (1992), 3,000 effective samples should result in a 99% chance of estimating the 0.025 or 0.975 posterior quantiles within $\pm$ 0.0125 quantile units, and a 83% chance of estimating the posterior median with the same level of precision. Diagnostic summaries are shown only if that parameter was assigned a prior in the model, *e.g.*, regression-based models estimated α_j and β_j as leading parameters, whereas the state-space models estimated $U_{\mathrm{MSY},j}$ and $S_{\mathrm{MSY},j}$ as leading parameters.

| | Brooks-Gelman-Rubin | | | | | | Effective MCMC Samples | | | | | |
| | Regression | | State-space | | | | Regression | | State-space | | | |
Parameter	LM	LME	vm	Vm	vM	VM	LM	LME	vm	Vm	vM	VM
α_j	<1	<1	—	—	—	—	<1	0	—	—	—	—
β_j	<1	0	—	—	—	—	<1	0	—	—	—	—
$U_{\mathrm{MSY},j}$	—	—	0	0	0	0	—	—	<1	4	<1	4
$S_{\mathrm{MSY},j}$	—	—	0	<1	0	0	—	—	<1	4	<1	5
$\sigma_{R,j}$	—	—	0	0	0	0	—	—	0	<1	0	<1
ϕ	—	—	0	0	0	0	—	—	0	0	0	0
π	—	—	0	0	0	0	—	—	0	0	0	0
$\bar{\rho}_{i,j}$	—	—	0	0	0	0	—	—	0	<1	0	0
$R_{y,j}$	—	—	0	<1	<1	<1	—	—	<1	<1	<1	<1
U_t	—	—	0	0	0	0	—	—	0	<1	0	<1

TABLE 4.13: Posterior coverage for key quantities in the simulation-estimation trials. Coverage was calculated as the percentage of all estimated 95% credible intervals across simulated data sets that contained the true quantity. Bold numbers are those that fall greater than 5 percentage points from the optimal coverage.

Quantity	Regression		State-space			
	LM	LME	vm	Vm	vM	VM
Parameters						
α_j	**84**	**68**	95	94	95	95
β_j	**83**	**73**	90	**88**	90	**89**
$U_{\mathrm{MSY},j}$	**84**	**68**	95	94	95	95
$S_{\mathrm{MSY},j}$	**85**	**76**	**85**	**84**	**85**	**85**
$\sigma_{R,j}$	—	—	**44**	93	**45**	95
ϕ	—	—	**87**	97	**88**	97
π	—	—	**12**	**11**	**16**	**16**
$\bar{\rho}_{i,j}$	—	—	**86**	93	**88**	93
Mixed-stock reference points						
$S_{0.1}^*$	99	94	93	92	93	93
$S_{0.3}^*$	99	94	96	96	96	96
$S_{0.5}^*$	94	91	91	90	91	90
S_{MSY}	100	**88**	94	90	94	91
$U_{0.1}^*$	**79**	**38**	96	97	96	96
$U_{0.3}^*$	**66**	**32**	99	99	99	99
$U_{0.5}^*$	**21**	**38**	95	95	96	96
U_{MSY}	**75**	**80**	95	91	95	93
Abundance states						
U_t	—	—	**64**	**64**	**66**	**66**
$R_{y,j}$	—	—	**59**	**58**	**60**	**59**
$S_{t,j}$	—	—	**50**	**49**	**51**	**51**
H_t	—	—	**51**	**51**	**55**	**55**
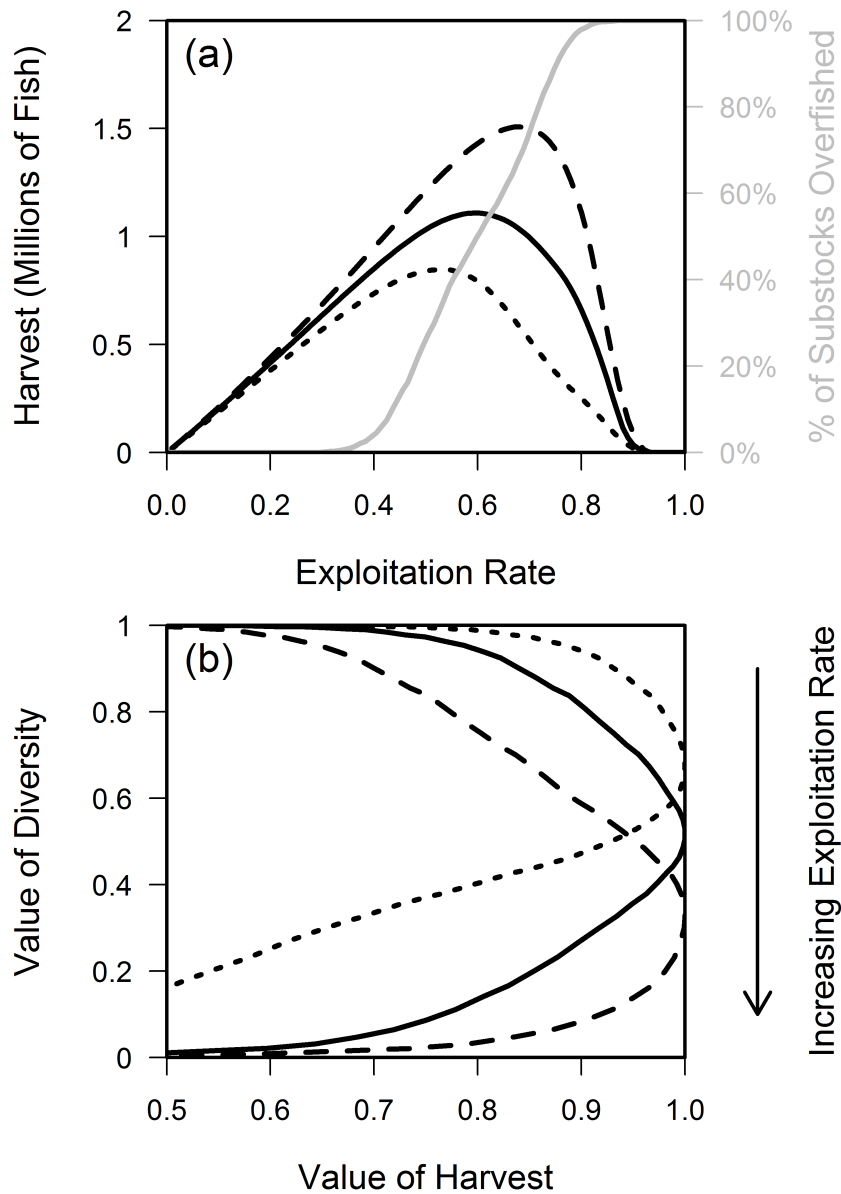
FIGURE 4.1: Visualization of how different types of heterogeneity in substock productivity and size influence the shape of trade-offs in mixed-stock salmon fisheries. Solid black likes are the case where stock types are split evenly among large/small and productive/unproductive stocks. Dotted black likes are the case where all small stocks are productive and all large stocks are unproductive, and dashed lines are the opposite (all big stocks are productive). (a) Equilibrium aggregate harvest and proportion of substocks overfished plotted against the exploitation rate (b) value of the biodiversity objective (0 = all stocks overfished) plotted against the value of harvest (the long term proportion of the aggregate MSY attained). Notice that when all big stocks are productive (dashed lines), the trade-off is steeper, *i.e.*, more harvest must be sacrificed in order to ensure a greater fraction of substocks are not overfished.
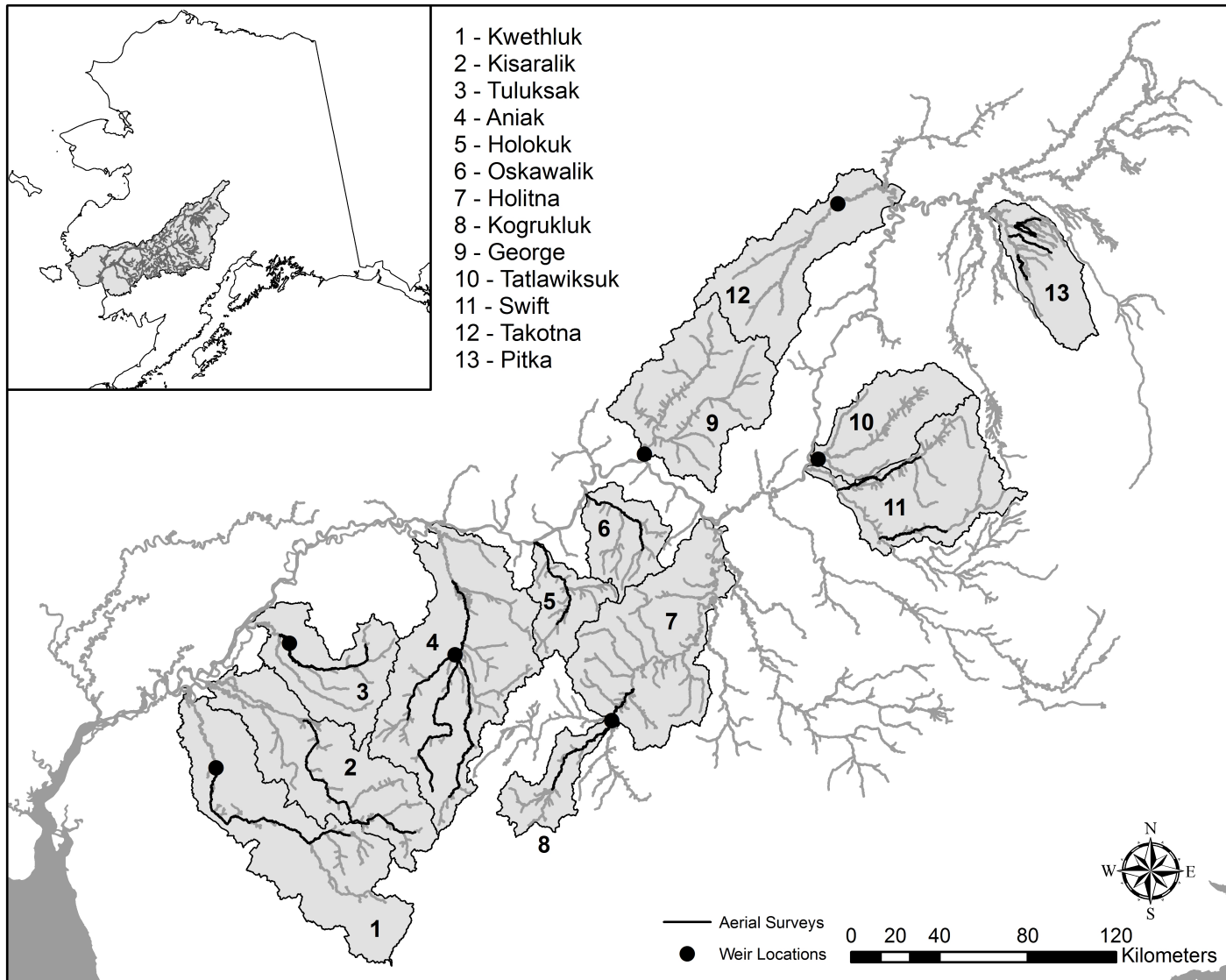
FIGURE 4.2: Map of the Kuskokwim River drainage, with the 13 drainage basins representing unique spawning units (substocks) used in this analysis. Black points show the location of weir projects, black sections of river indicate the reaches flown as part of aerial surveys. Drainages monitored *via* both aerial survey and weir used the weir counts to inform escapement estimates in this analysis, with the exception of the Aniak drainage (#4), for which aerial survey data were much more abundant than available weir data.
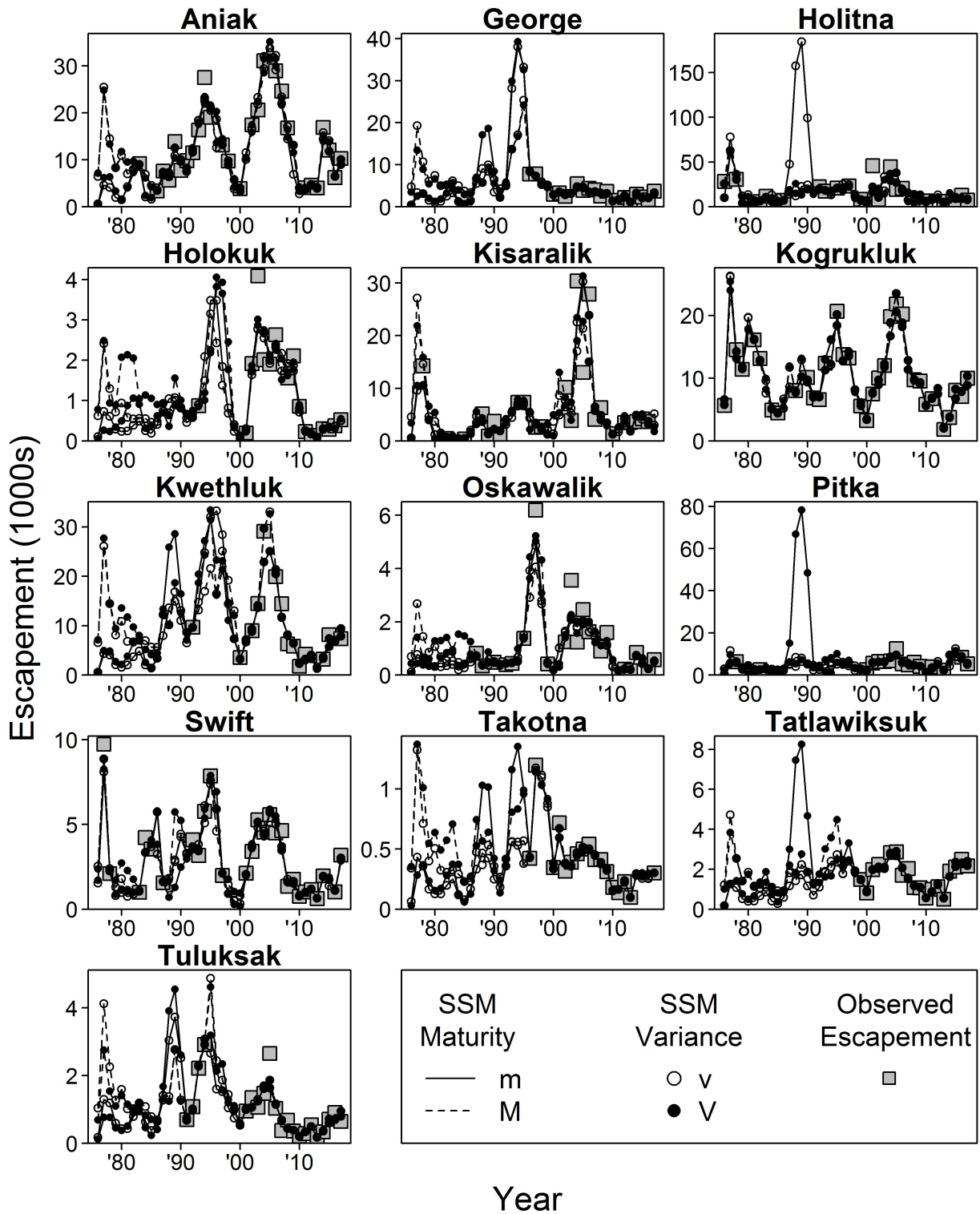
FIGURE 4.3: Observed and fitted escapement time series for each Kuskokwim River substock. Line/symbol types denote the particular state-space model and grey squares denote observed data.
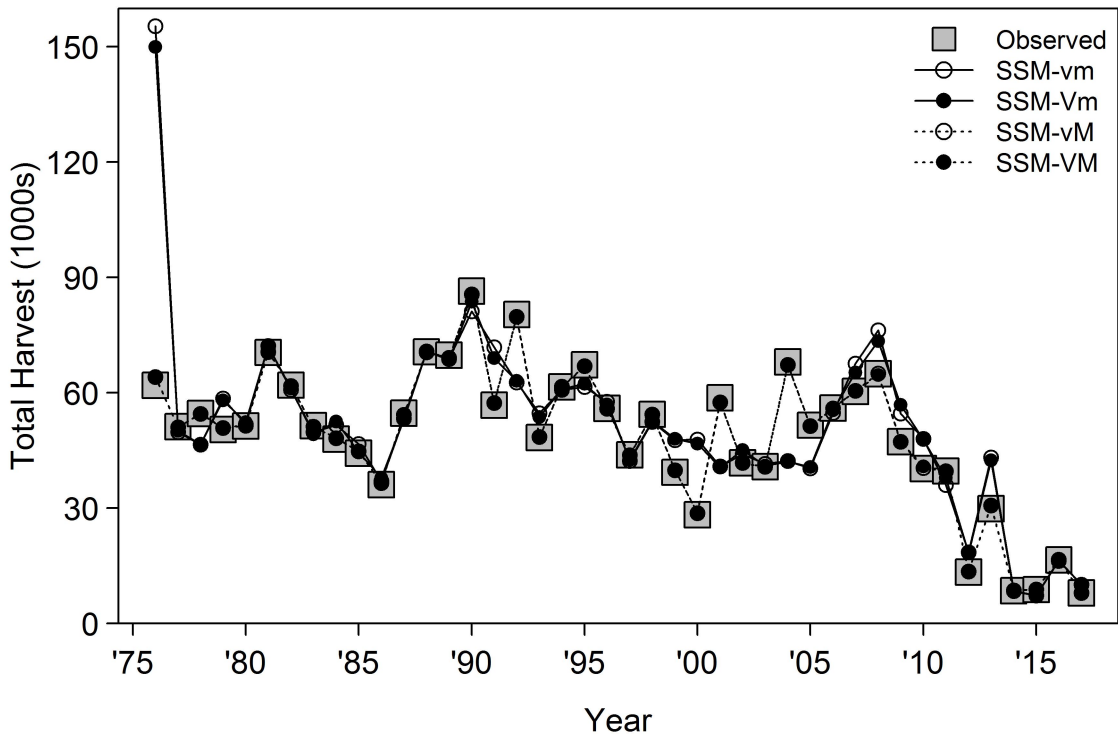
FIGURE 4.4: Observed and fitted harvest time series aggregated across all Kuskokwim River substocks included in this analysis. Line/symbol types denote the particular state-space model and grey squares denote observed data.
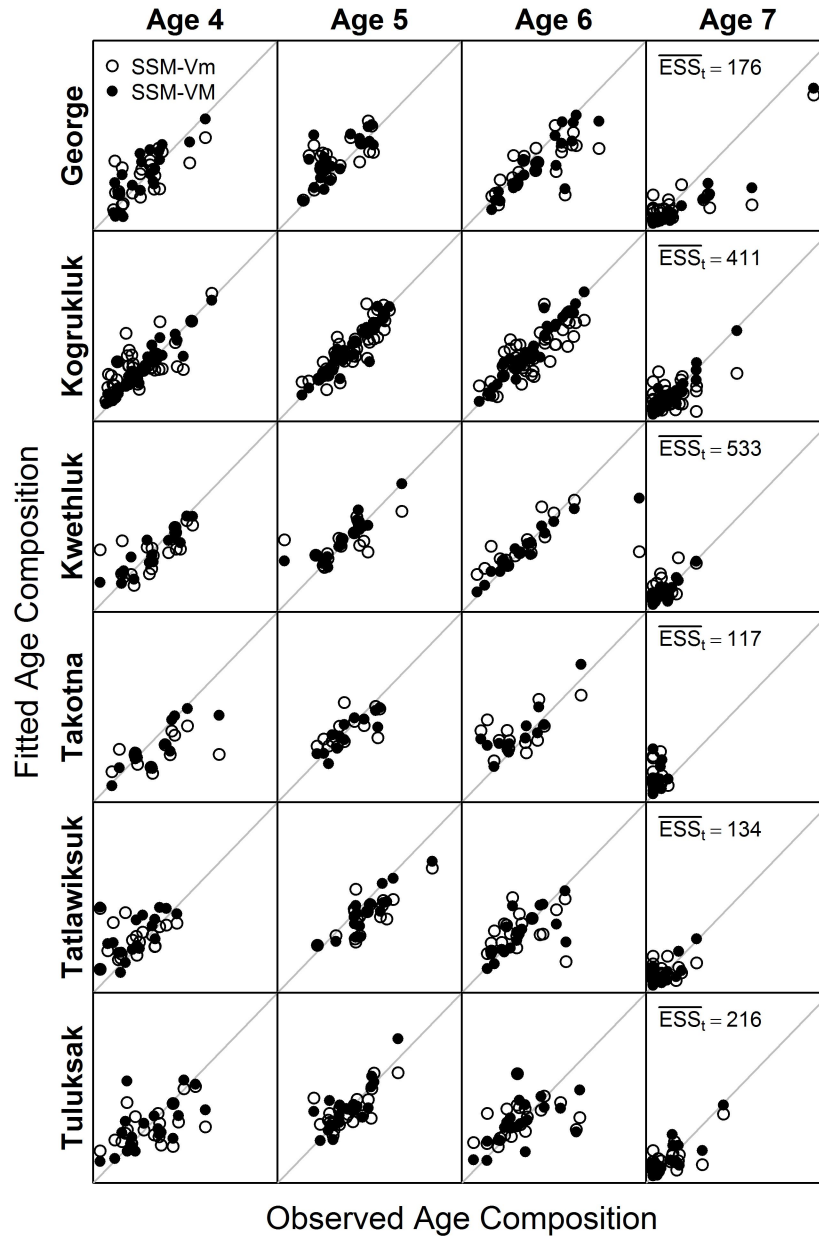
FIGURE 4.5: Observed and fitted age composition for the six weir-monitored substocks. Each scatter plot is the pair of fitted *versus* observed proportion of the escapement in each age each year with data available and the grey line represents the 1:1 perfect fit line. Point types denote two models: SSM-Vm (time-constant maturity; hollow circles) and SSM-VM (time-varying maturity; filled circles). $\overline{ESS}_t$ represents the average number of fish successfully aged each year with data for each substock, which was used as the sample size to weight the data in the multinomial likelihoods that used these data. Panels are scaled to have the same x-axis and y-axis limits within an age across substocks and range from $0-1$ for ages $4-6$ and $0-0.3$ for age 7.
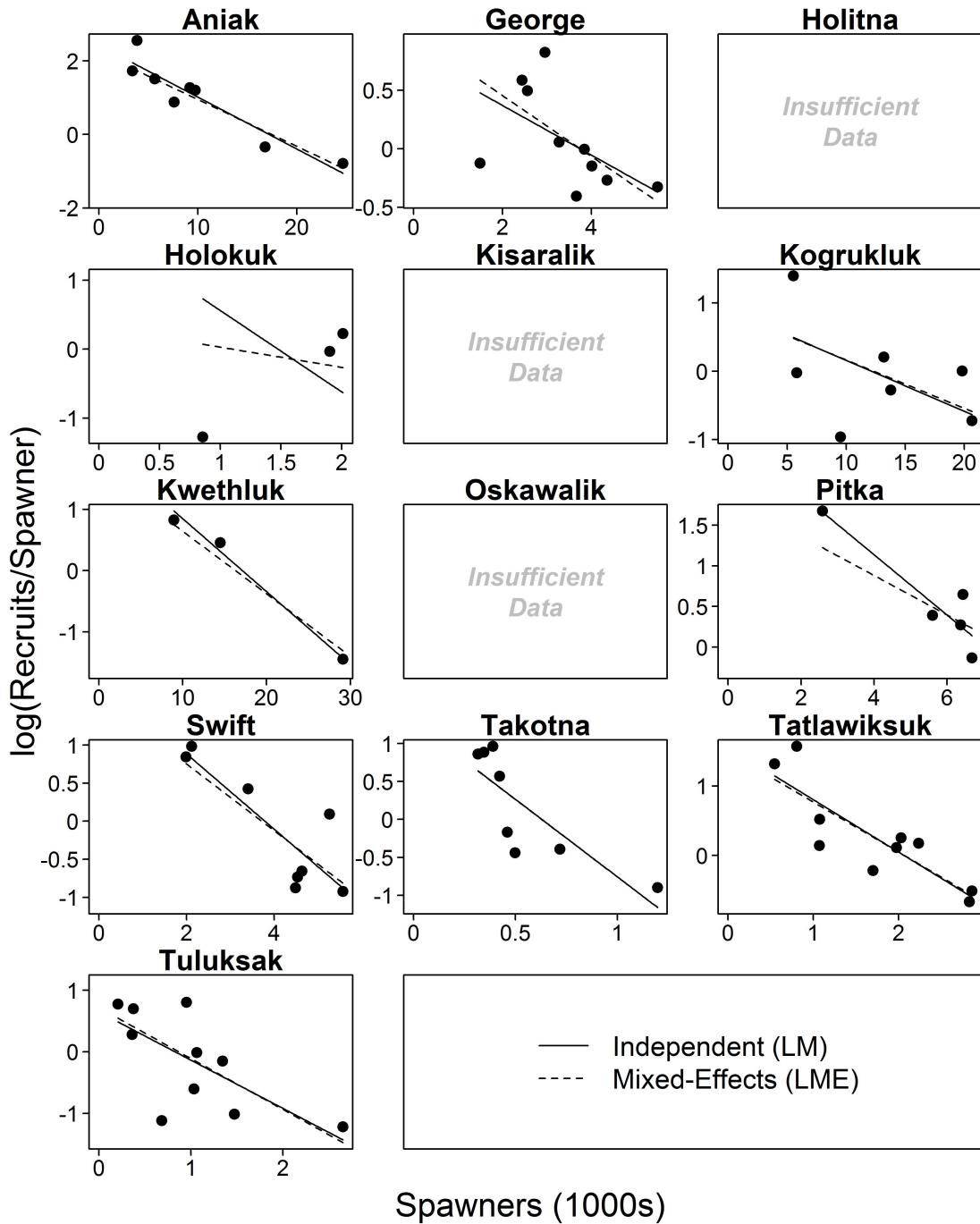
FIGURE 4.6: Fit of the regression approaches to fitting the multi-stock spawner-recruit analysis to the Kuskokwim River substock data. Points are observed log(recruits/spawner) *versus* spawners, solid lines represent the fit for the independent regression models, and dashed lines represent the fit suggested by the model with random intercept effects for each substocks. Three substocks had fewer than three observed data points, rendering fitting a regression line infeasible. Note that a constraint was imposed that maintained $\log(\alpha_j) > 0$ which prevented biologically implausible values, and explains the poor fit for the Holokuk River substock.
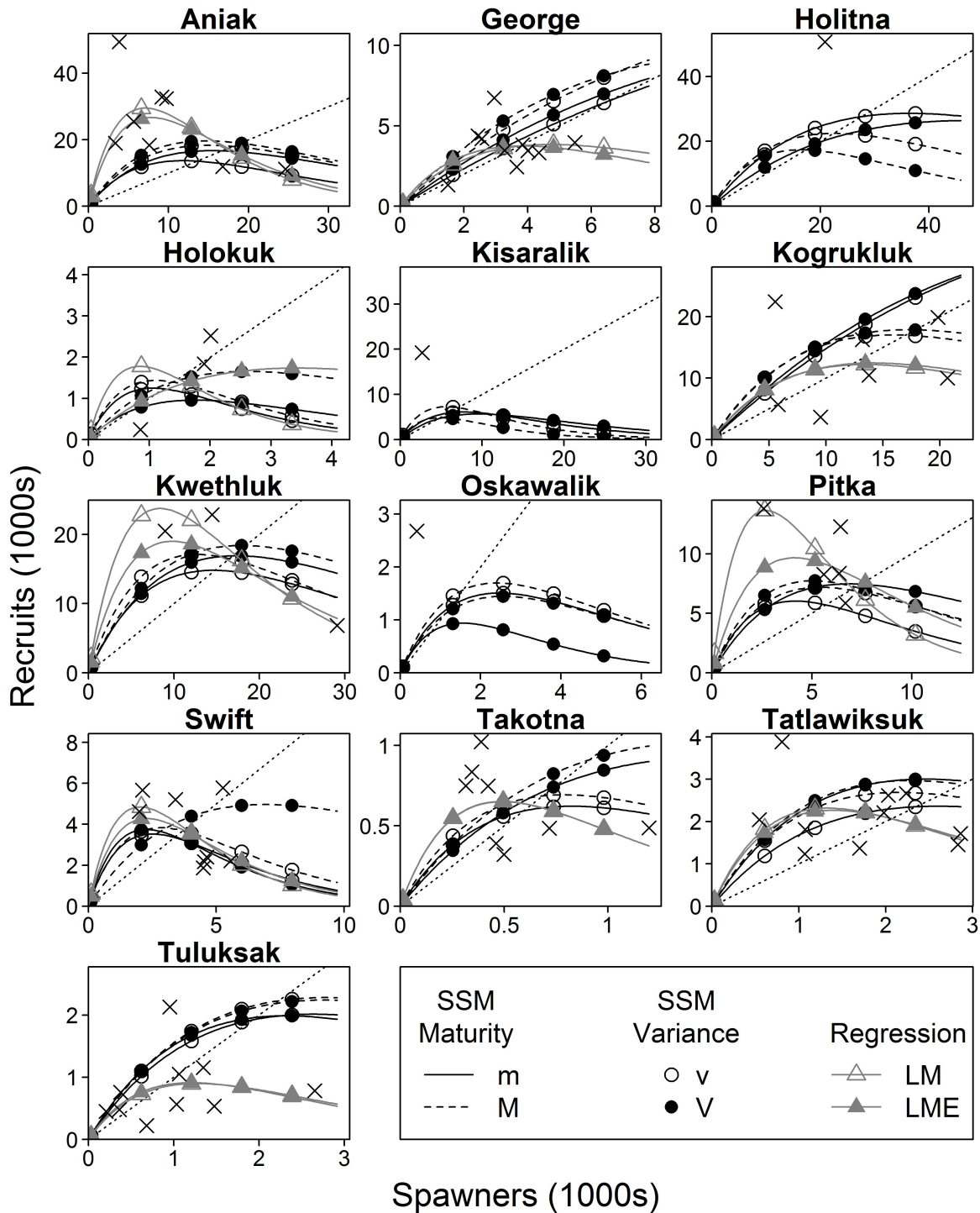
FIGURE 4.7: Fitted spawner-recruit relationships for the 13 substocks monitored in the Kuskokwim River subdrainage included in this analysis. Line and point types correspond to different models; crosses are completely observed spawner-recruit pairs. Note that the regression approaches (grey lines/triangles) fitted only to these data, the state-space models (black lines/circles) fitted to all observations of substock-specific escapement, aggregate harvest, and age composition.
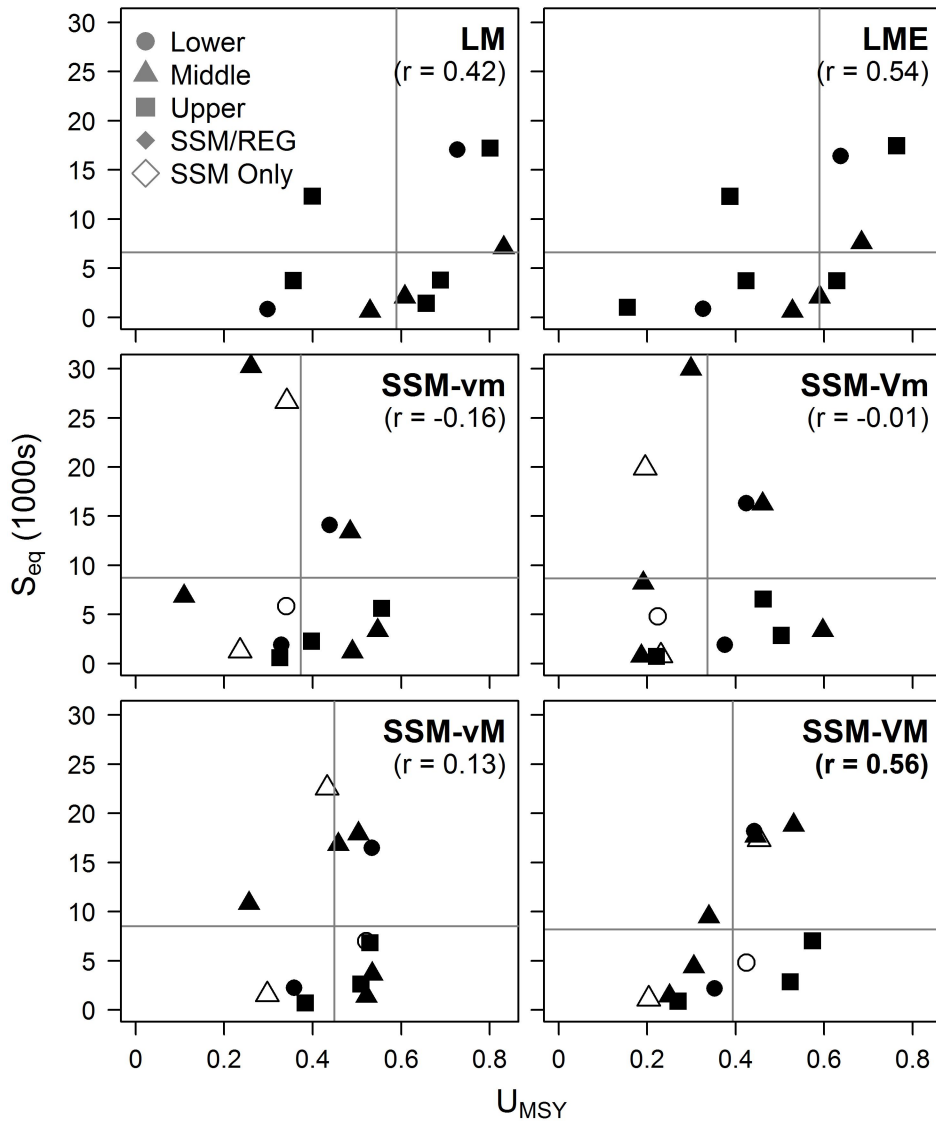
FIGURE 4.8: Relationships between substock size and productivity as estimated by the six estimation approaches in the analysis. Symbol shapes denote the region within the Kuskokwim drainage the substock is located in and hollow symbols in the state-space models are substocks that could not be fitted by the regression approaches. The value in parentheses is Pearson's r correlation coefficient; bold numbers indicate a significant correlation at $\alpha = 0.05$.
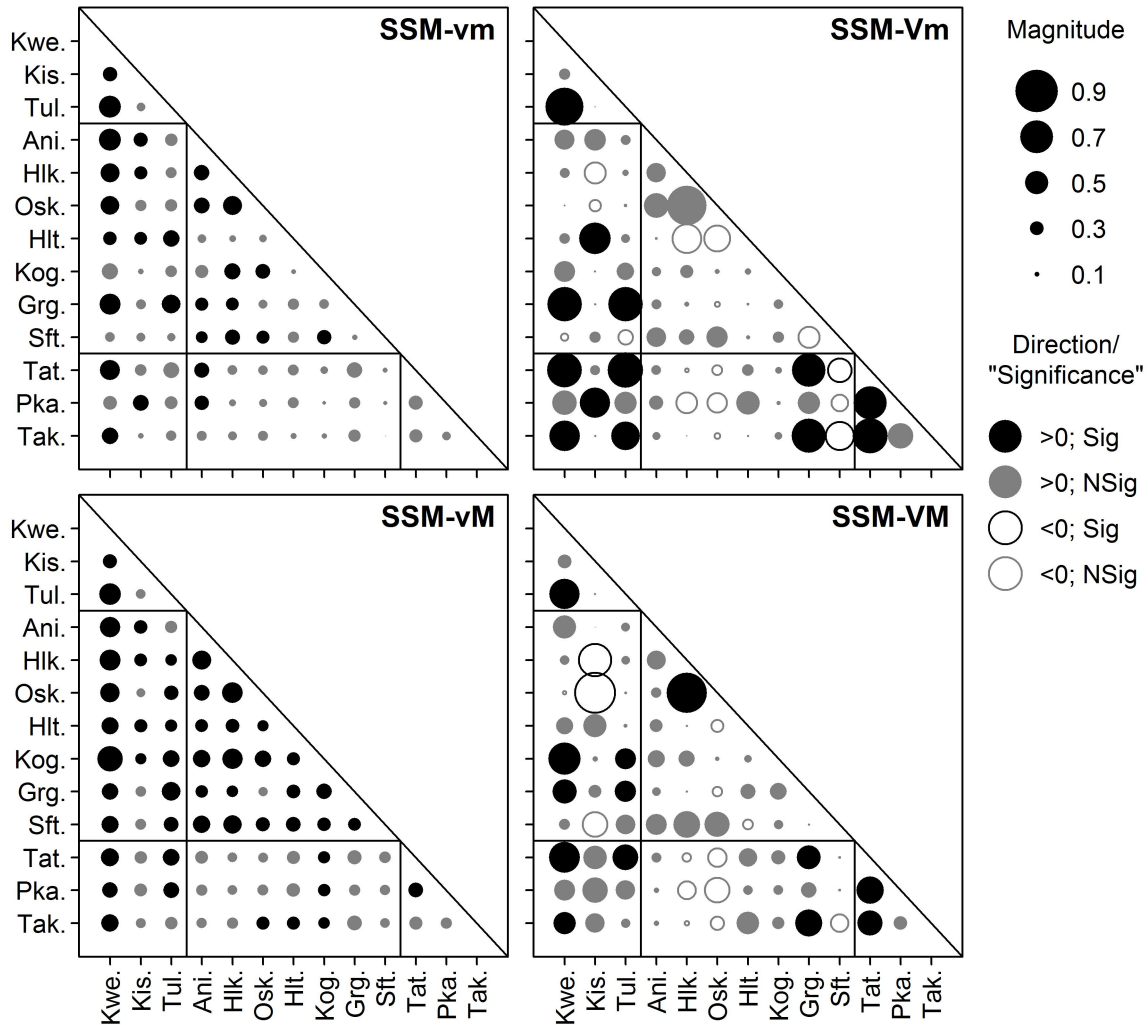
FIGURE 4.9: Correlation coefficients for recruitment residuals for each pair of substocks. The size of each circle represents the magnitude of the correlation, the shade represents significance (whether 95% credible interval included 0), and the fill represents directionality as described in the legend. Substocks are ordered from downriver to upriver on both axes, and vertical/horizontal lines denote the boundaries between lower, middle, and upper river substocks referred to in the text.
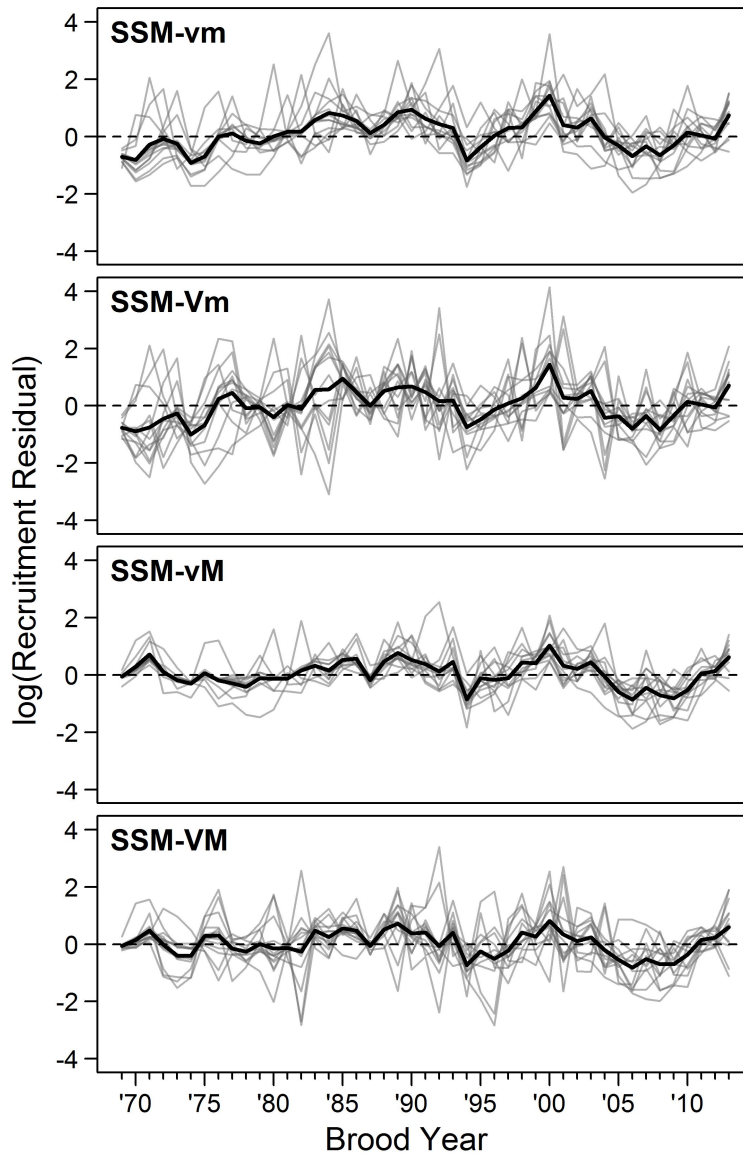
FIGURE 4.10: Time series of recruitment residuals for each substock under each of the four state-space models. Substock-specific time series are represented by grey lines and the average across substocks within a brood year are represented by the thick black line. A dashed line at zero (no error) is provided for reference.
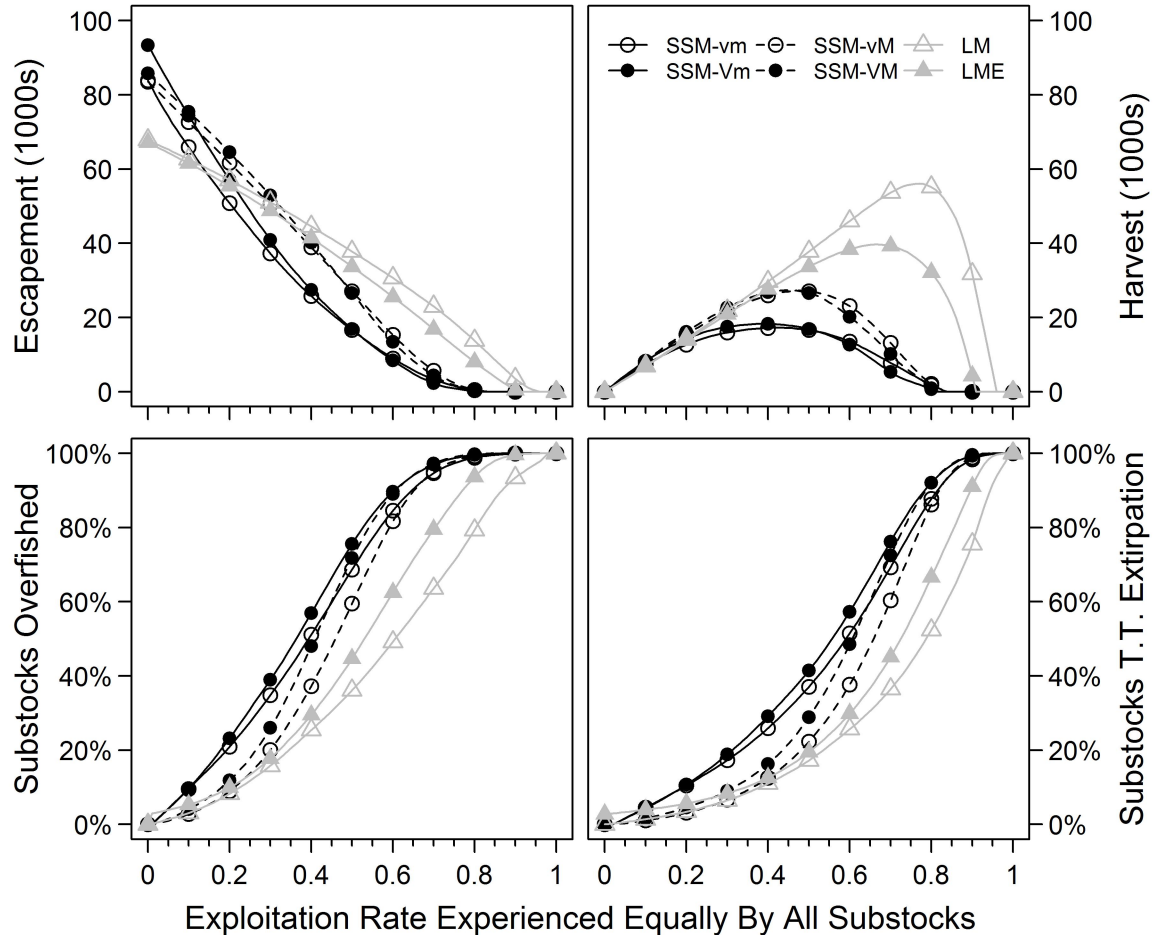
FIGURE 4.11: Visualization of harvest-biodiversity trade-offs based on equilibrium states (escapement and harvest) of the aggregate stock and the percentage of substocks expected to be in an undesirable state as a function of the exploitation rate under the assumption that all substocks are fished at the same rate. Overfished is defined here as $U > U_{\mathrm{MSY},j}$. "T.T." stands for "trending toward", and represents the case where equilibrium escapement would be ≤ 0. To facilitate comparisons with the regression approaches (grey lines/triangles), the three substocks with insufficient data for fitting regression models were excluded from summaries of the state-space models (black triangles/circles).
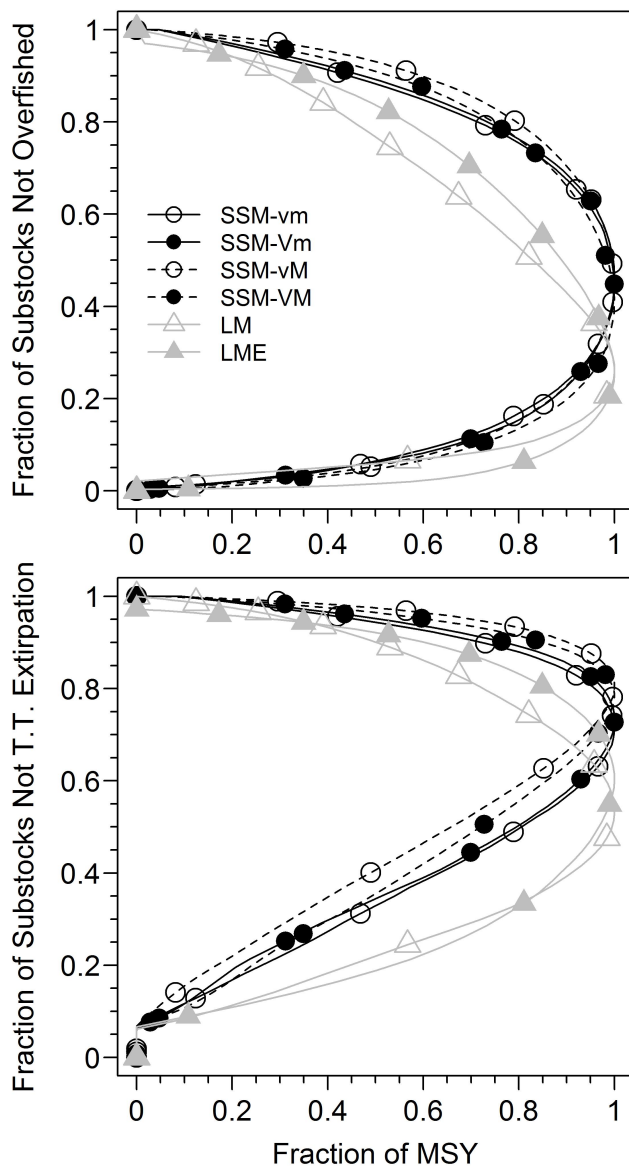
FIGURE 4.12: Alternative (and more direct than Figure 4.11) visualization of harvest-biodiversity trade-offs for monitored Kuskokwim River Chinook salmon substocks. The conditions of overfished and trending towards extinction are the same as defined in Figure 4.11. These figures should be interpreted by determining how the value of the biodiversity objective (*y*-axis; expressed as the fraction of substocks that would not be in the undesirable condition) must be reduced to increase the value of the harvest objective (*y*-axis; expressed as a fraction of the maximum sustainable yield). To facilitate comparisons with the regression approaches (grey lines/triangles), the three substocks with insufficient data for fitting regression models were excluded from summaries of the state-space models (black triangles/circles). All symbols represent increasing exploitation rates in increments of 0.1 as you move down the *y*-axis.

167
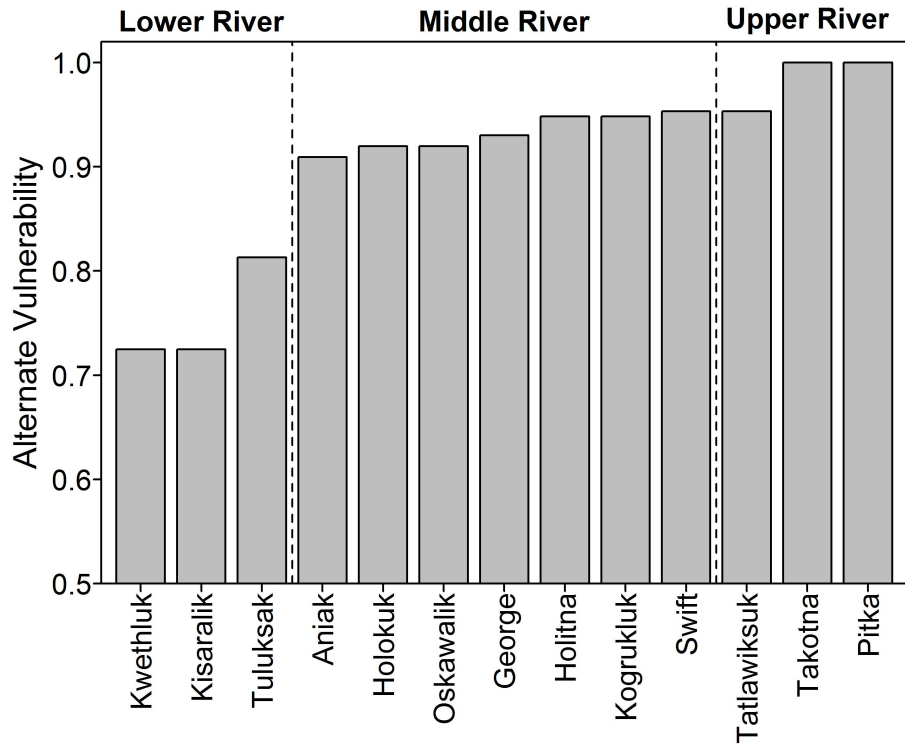
FIGURE 4.13: Alternative vulnerability schedule (v_j) used in a sensitivity analysis of the state-space models. Vulnerabilities were obtained by determining the fraction of total fishing households the individual substocks must swim past in order to reach their respective natal spawning grounds. Substocks are ordered from downstream to upstream.
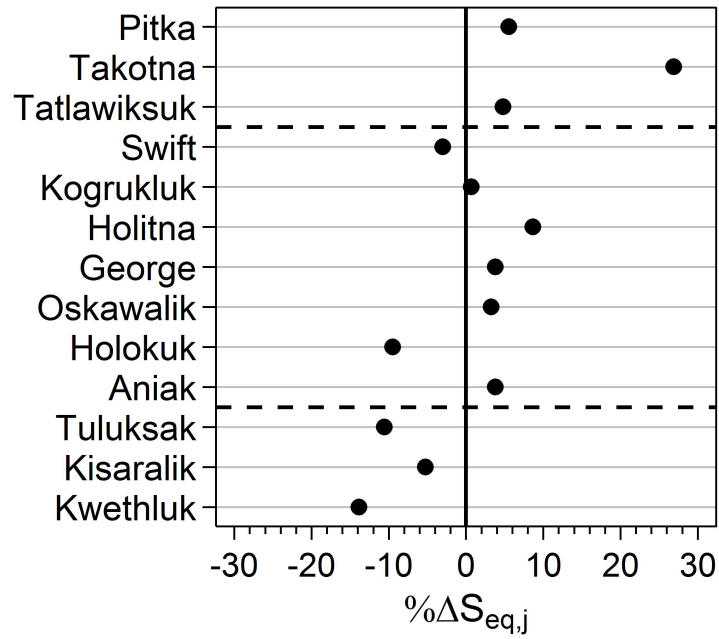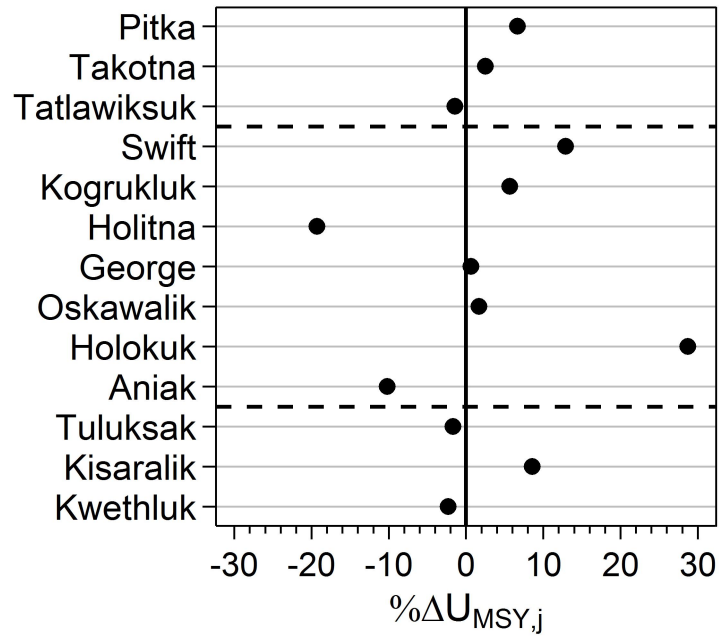
FIGURE 4.14: Percent change in substock-specific population quantities between the alternative and default vulnerability schedules. Positive differences indicate that the alternative estimates were larger than the default estimates. Estimates are displayed for SSM-VM only, all other state-space models returned similar differences.
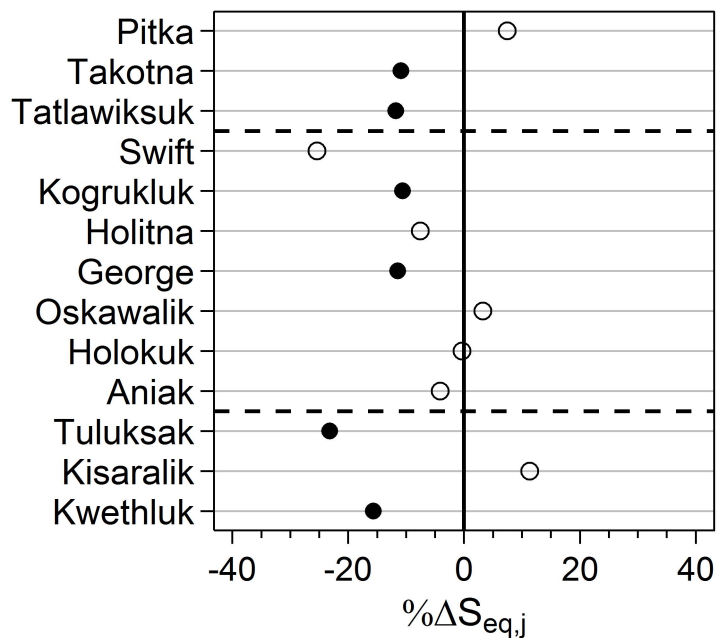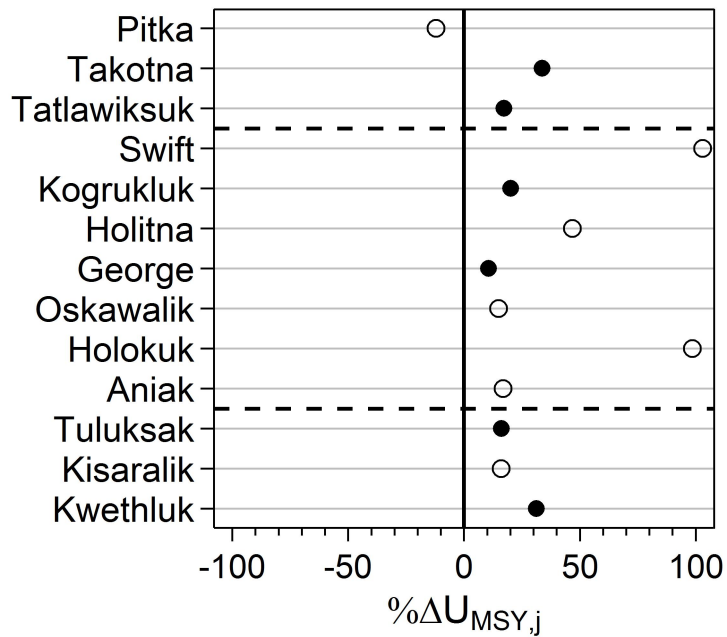
FIGURE 4.15: Percent change in substock-specific population quantities estimates between the alternative and default age composition weighting schemes. Positive differences indicate that the alternative estimates were larger than the default estimates. Filled symbols denote substocks with age composition data. Estimates are displayed for SSM-VM only, all other state-space models returned similar differences.
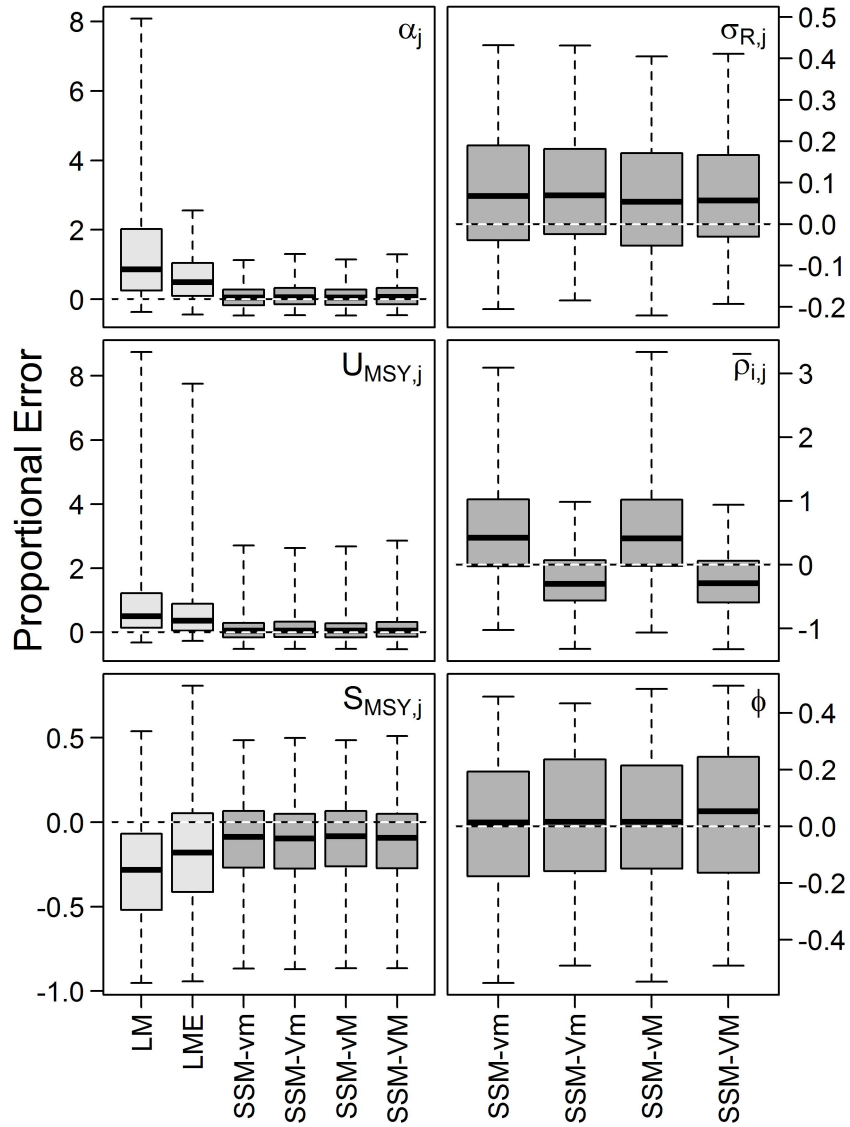
FIGURE 4.16: Central tendency and variability in proportional error for some parameters in the multi-stock spawner-recruit models from the simulation-estimation trials. Point estimates used were posterior medians.
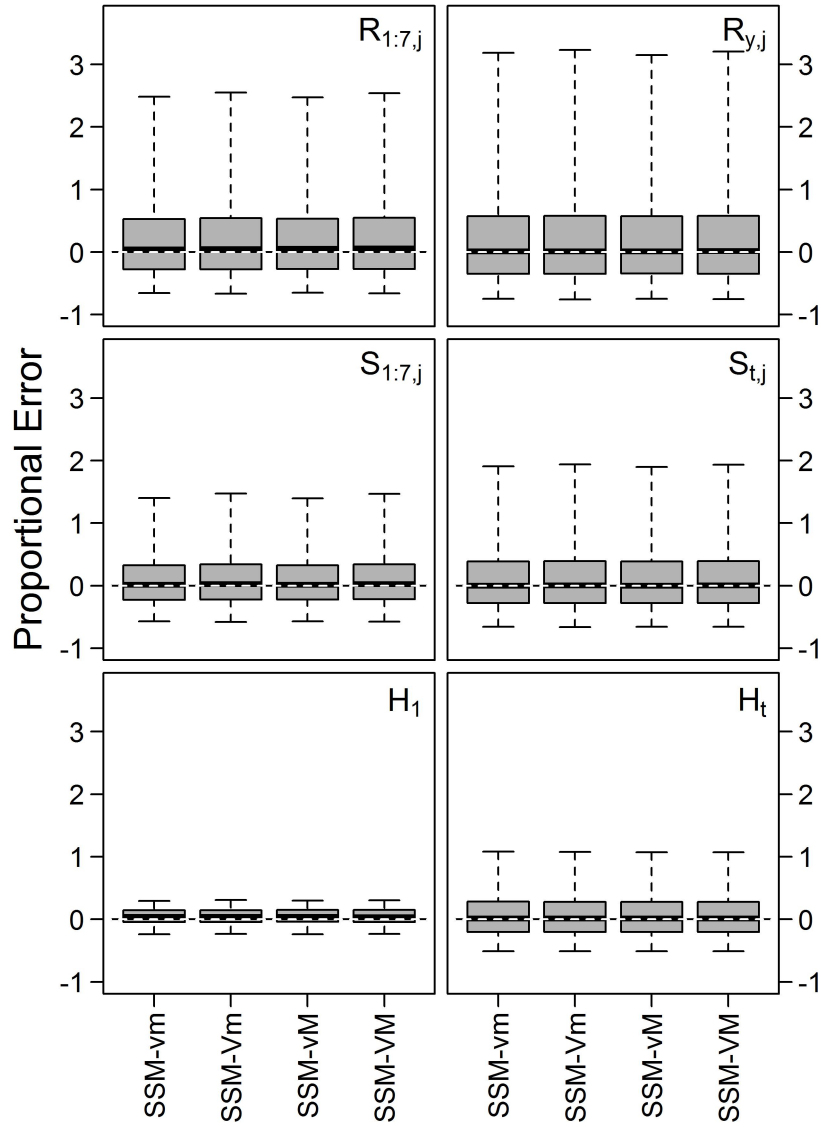
FIGURE 4.17: Central tendency and variability in proportional error for some abundance states in the multi-stock spawner-recruit models from the simulation-estimation trials. Quantities are separated by the early portion of the time series (left panels) and the entire time series (right panels). Point estimates used were posterior medians.
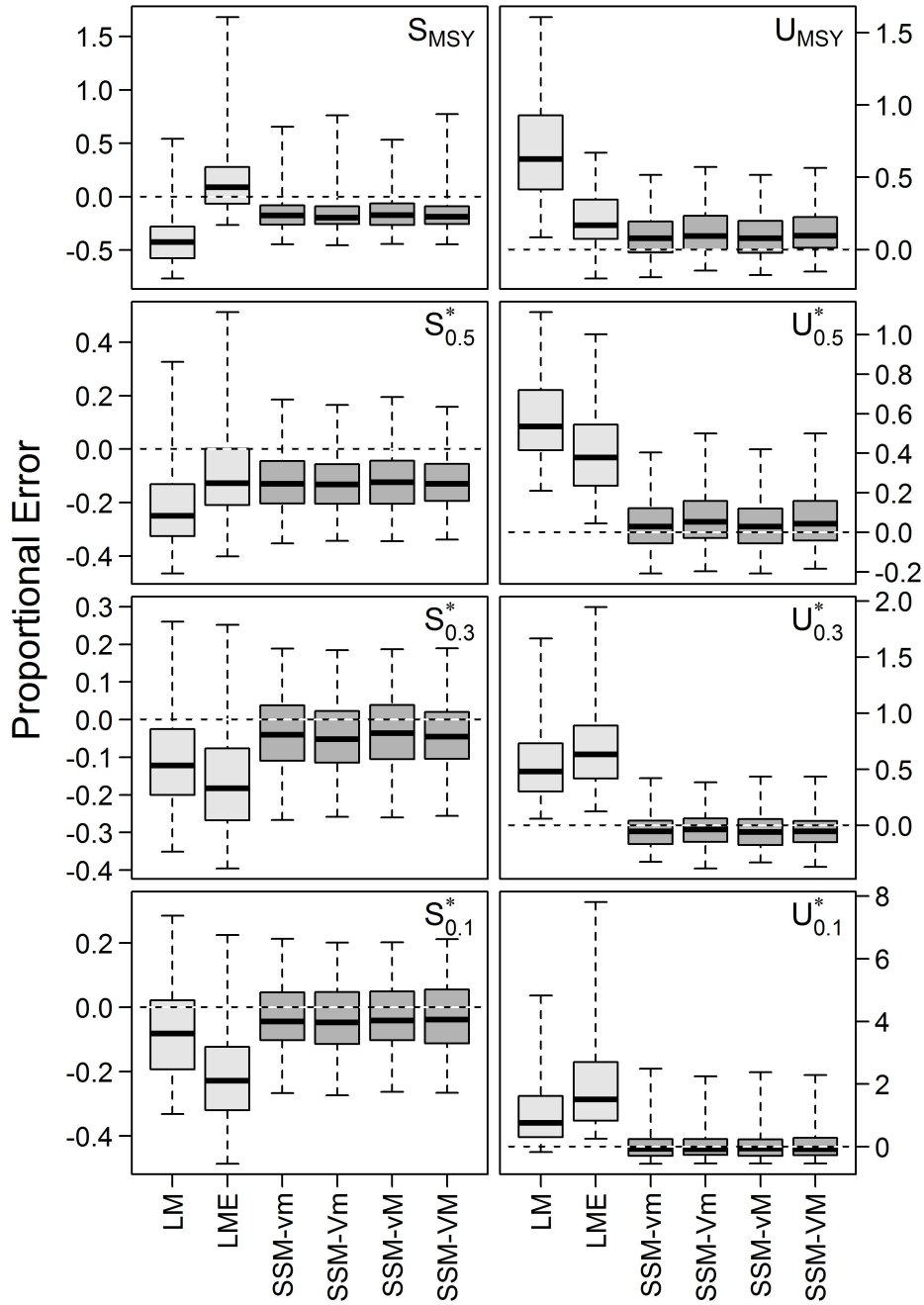
FIGURE 4.18: Central tendency and variability in proportional error for key biological reference points of the aggregate mixed stock. S_p^* and U_p^* are the aggregate escapement and fully vulnerable exploitation rate that would ensure no more than $p \cdot 100\%$ of substocks are overfished, respectively. Point estimates used were posterior medians.

Chapter 5

Conclusions

In writing this dissertation that summarizes my research over the last 3.5 years, I sought to develop, evaluate, and illustrate the application of quantitative tools that may be used to address challenges in salmon management, with particular focus on stocks in western Alaska. Specifically, the tools sought to reduce uncertainty in decision-making (in the case of the run timing forecast in Chapter 2) and inform policy analyses by identifying and measuring the shape and magnitude of trade-offs among competing objectives (for in-season management in Chapter 3 and long-term management in Chapter 4). In a broad sense, I believe I have accomplished this objective. The work presented in this dissertation provides a detailed look at ways quantitative tools can be used in the management of salmon fisheries in western Alaska – some were novel to salmon management, all were novel with respect to the Kuskokwim River. This final chapter serves as my reflection on my doctoral work, including further insights on each of the three primary research projects I completed as well as my time working in Kuskokwim River management system at the Yukon Delta National Wildlife Refuge during the summers of 2016 – 2018.

5.1 Further insights on each project

5.1.1 Chapter 2: Run timing forecasts

My work on developing run timing forecasts for Kuskokwim River Chinook salmon has not proven as fruitful as I would have hoped when we wrote the grant proposal. In each year since its inception following the 2016 season, the forecast model did not provide good forecasts of

D_{50}: both 2017 and 2018 were moderately late runs yet the forecast model suggested the runs would be several days early. Furthermore, a subsequent analysis (Staton and Catalano 2019) strongly suggested that using the forecast model provides no utility for improving perceptions of run size based on test fishery data. The cause for this finding is two-fold. First, although the environmental relationships were present for all evaluated variables (Figure 2.3; Table 2.3), the residual variability was too high to result in accurate and precise forecasts based on them. Second, inter-annual variability in the Bethel Test Fishery catchability (*i.e.*, the fraction of total run captured; the inverse is commonly referred to as "run-per-index"; Flynn and Hilborn 2004) is the dominant cause of uncertainty in using this index for predicting run size for much of the season, not run timing. Staton and Catalano (2019) illustrated that the average CV of abundance predictions from a relationship between historical abundance and cumulative catch per effort each day was approximately 38% early in the season and 30% at the end of the season (Figure 4 therein). So even after the test fishery is done catching Chinook salmon, a large amount of uncertainty still remains in the actual run size (this finding was the same for methods that included and excluded the run timing forecast). Based on the variability of errors made by these predictions, it seems this level of uncertainty is appropriate (also shown in Figure 4 of Staton and Catalano 2019). These findings illustrate that the Bethel Test Fishery is a poor index of run size, and that availability of the run timing forecast did not improve this situation.

Still, the forecasting framework I developed was a statistically rigorous approach to dealing with high-dimensional variable selection over time and space. To my knowledge, the sliding climate window algorithm I employed has not widely been used in ecological problems. This must either be due to the computational costs or because it is not well-known, given it is an intuitive and objective approach to select temporal periods for prediction. After giving a talk on the forecasting approach I developed, a prominent ecologist in my field asked whether I thought this constituted a "data dredge" analysis, with the implication that it

was a bad thing if so. In my view, data dredging is the practice of searching for all possible relationships that significantly support some hypothesis and discarding those that do not, or otherwise developing and confirming ad-hoc hypotheses based on such a search. First, the notion of statistical significance was never used in my analysis: all decisions were made based on actual out-of-sample predictive performance (through forecast cross-validation) or an index of predictive performance (Akaike's Information Criterion). Second, no biological hypotheses were tested in this analysis. Instead, the search was conducted to find the variables most appropriate for forecasting run timing. In these cases, I think it is completely rational to perform as exhaustive of a search as possible, and the approach I developed served as an intuitive and objective means to do this.

It has been brought to my attention that methods exist that may provide better forecasts than the approach I used, which was simple/multiple regression at its core. Specifically, machine learning tools like random forests could show promise for this problem and others like it. Additionally, upon further reflection it is possible that having a continuous forecast of D_{50} could be less useful than a discrete forecast of an early, average, or late run. Such a forecast could be obtained using a multinomial logistic regression (Agresti 2002, Ch. 7), which would provide predicted probabilities that the run will be early, average, or late. In my experience, uncertainty is more appropriately interpreted by managers when presented as the probability of outcomes rather than using uncertainty intervals. There would surely be some loss of resolution (*e.g.*, not all early runs are equally so), but the categories could be selected based on how the D_{50} within each would influence the management inference. For example, there should exist some thresholds of early/late timing that would drastically change the inference from if it was an average run; I would propose that these thresholds be used to delineate the categories.

5.1.2 Chapter 3: In-season MSE analyses

In this stochastic MSE analysis of a large salmon-producing river system in western Alaska, I found several important implications, some that were known *a priori* and some that were not. For example, the finding that more conservative substrategies should be favored in small runs as opposed to large runs makes intuitive sense. However, the finding that management performance was generally most sensitive to run timing in large runs rather than small runs was not necessarily known *a priori*, nor was the finding that good performance can be attained with a wide range of management strategies. This analysis was useful in that it provided an objective basis for strategy comparisons and necessitated critical thinking about the important drivers of system dynamics (*e.g.*, effort responses to fishery conditions) as well as the direct ways in which information influences a particular decision. It is my hope that when presented to fisheries managers and stakeholders in the region, perhaps a more-informed dialog may be had regarding the merits and detriments of candidate management strategies.

The most complex assessed management strategy (Strategy #4; explicit harvest target, selected probabilistically, updated with in-season information) is a computer-based representation of the strategy implemented by the U.S. Fish and Wildlife and the Kuskokwim River Inter-Tribal Fisheries Commission for the portion of the fishery within the Yukon Delta National Wildlife Refuge in the years 2015 – 2018. Each year has differed in the method, rigor, and transparency of (1) initially selecting the season-wide harvest target (H_T), (2) utilizing in-season information to update H_T and the weekly target ($H_{T,w}$), and (3) determining how much fishing opportunity should be allowed conditional on $H_{T,w}$. The overarching structure (*e.g.*, the use of H_T based on an escapement limit threshold; S_L), however, is the same as the simulated strategy and the probabilistic selection of H_T based on risk tolerance (P^*) and S_L was conducted for the first time in 2018.

With the current amount of information available for in-season management, this is among the most complex management strategies that could be used for the Kuskokwim

Chinook salmon harvest control situation, and a key finding of the MSE analysis was that it did not perform overwhelmingly better than far simpler strategies. In the very smallest runs (50,000 – 80,000), it did provide substantially better escapement performance than other policies, but only because the fishery was all but shut down completely. If this is desirable, the schedules used by Strategies #2 and #3 (Figure 3.1) could be altered to be more conservative for forecasts falling in this bin. Additionally, uncertainty in the forecast could be included by weighting schedules by the probability that the run will be in each run size bin according to the pre-season forecast. In the next smallest category (80,000 – 130,000), the gains in escapement utility by using Strategy #4 were negligible in comparison to other simpler strategies, and came at the cost of reducing the harvest utility by approximately half (Figure 3.11). Likewise, it showed no real gains in harvest equity or even substock exploitation utilities, though no assessed strategies strongly influenced these metrics. As a result of these findings, simple schedule-based strategies based solely on a pre-season forecast (as in Strategy #2) or combined with in-river species composition (as in Strategy #3) would perform well in the Kuskokwim and similar systems. Given the finding that many strategies perform similarly, however, perhaps the more important question now becomes which strategy could meet other objectives not included in this analysis (*e.g.*, fishers' desire to know when to expect fishing opportunities far in advance, desire to fish early in the season, *etc.*) and is transparent to the parties involved.

5.1.3   Chapter 4: Multi-stock spawner-recruit analyses

For my last project, I evaluated spawner-recruit methods for mixed-stock salmon fisheries. I developed a novel state-space framework for simultaneously estimating spawner-recruit parameters from the substocks within a larger drainage basin when they are all harvested as a mixed-stock – the method was shown to perform substantially better than simpler regression-based approaches that, though they have long been known for their problems (Ludwig and

Walters 1981; Walters 1985; Walters and Martell 2004), they are still applied to estimate salmon population dynamics parameters and to provide management recommendations (Clark *et al.* 2009; Korman and English 2013).

An alternative assessment approach would be to attempt to fit separate state-space models to the individual substocks as opposed to the more integrated approach I developed. This approach should retain the benefits of (1) not needing to exclude some observations because they do not constitute fully observed brood year pairs and (2) separately modeling the population dynamics and observation processes to (at least partially) handle the time-series and errors-in-variables biases. However, given that not all substocks have age composition data, the individual models would require some additional assumptions in this regard, and the degree of recruitment synchrony among substocks would only be available as part of an ad-hoc analysis of patterns from individual substocks. The more integrated state-space approach I took was a rational use of the available age composition data: only substocks with observed data were fitted for age composition, and they informed the maturity model parameters to be used for other substocks. Additionally, by incorporating a covariance matrix for recruitment variability, the degree of synchrony in substock dynamics could be informed by years with data on multiple substocks, which could be then used to inform the latent recruitment states in cases where there were fewer observations to inform them. Further, as computationally intensive as the model I developed was, this independent state-space model approach would be even more so given independent MCMC sampling would need to be conducted for each substock, which is a primary reason why it was not evaluated here.

Although the state-space model was applied only to the Kuskokwim River data and the simulation was designed to mimic the data collection and substock heterogeneity for that system, the model potentially has wide applicability to other systems. Many salmon populations are harvested in mixed stocks; good examples include: the Columbia River in the northwestern contiguous United States, the Fraser and Skeena River drainages in British

179

Columbia, Canada, the Bristol Bay stocks in southwestern Alaska, and the Yukon River in western Alaska and Canada. All of these systems are composed of several distinct spawning units, some portion of the harvest occurs as a mixed-stock, and should have data time series at least as rich as the Kuskokwim River Chinook salmon data set I used. Granted, in some situations modifications would be necessary for each specific case. For example, in the Bristol Bay fishery for sockeye salmon, some harvest occurs by "intercept fleets", where fishers in one district may also harvest fish from other stocks that are migrating through that area but some harvest also occurs in-river. This kind of situation would require more detailed harvest information separated by mixed- and non-mixed sources, but I believe this would be within the abilities of the state-space model given the data are available.

5.2   Reflections on working for USFWS

In 2016 – 2018, I took on a more-involved role in the management of Kuskokwim River salmon fisheries by working as a Pathways Student[1] with the U.S. Fish at Wildlife Service at the Yukon Delta National Wildlife Refuge based out of Bethel, AK. Between May and August of these years, I abandoned my graduate work[2] and focused on in-season salmon assessment. In this role, I lead several efforts dealing with providing information to managers: (1) compiling, analyzing, and presenting run assessment information in consistent formats[3], (2) developing and applying statistically rigorous approaches to estimating harvest and effort arising from short-duration block-openers[4] and presenting this information to managers, and (3) presenting assessments of the risk of seeing undesirable escapement outcomes conditional

---

[1]Unofficial title: Quantitative Ecologist
[2]And my loving and understanding girlfriend (as of 2016), fiancée (as of 2017), and wife (as of 2018)
[3]Referred to as the "Daily Assessment Updates", over 250 documents were produced using `{rmarkdown}`; examples available upon request.
[4]Documented in Staton and Coggins (2016), Staton and Coggins (2017), Staton (2018d), plus a synthesis manuscript currently in preparation

on different harvest levels[5]. I have enjoyed this role enormously, and it was highly rewarding to apply my quantitative skills to applied management problems, and to see the inferences be used in decision-making. Though there was a distinct boundary between the topics I worked on in-season in this role and my more academic work (*i.e.*, presented in Chapters 2, 3, 4 herein), I certainly used what I learned in each aspect to inform the other. I have no doubt my summers spent in Bethel resulted in better dissertation research, particularly with respect to developing the operating model and candidate management strategies used in Chapter 4, and that I would have been less valuable to the management system without the quantitative training I have received in my graduate program.

---

[5]2018 only; with the aid of a Shiny application I wrote, which has become commonly known as the "P^* model" (https://bstaton.shinyapps.io/BayesTool/).

Appendix A

Parameterization of the Operating Model in Chapter 3

There were two main components of the operating model that needed to be parameterized based on observed information for it to adequately represent the dynamics of the real Kuskokwim River subsistence salmon fishery: biological (abundance, timing, spatial characteristics of the salmon populations, etc.) and sociological (spatial distribution of effort and desired/needed harvest and temporal aspects of the effort dynamics). This appendix details how the Kuskokwim River empirical information was used to parameterize the operating model used in the Chapter 3 analysis.

A.1 Biological quantities

A.1.1 Chinook salmon total abundance

Drainage-wide total Chinook salmon run abundance was informed by Liller *et al.* (2018), which reported estimates in the years $1976 - 2017$ from a maximum likelihood run reconstruction model. The model was fitted to 20 escapement indices, commercial fishery catch per effort, and nine years of drainage-wide estimates of total abundance obtained *via* large-scale mark-recapture experiments. Based on Liller *et al.* (2018), drainage-wide Chinook salmon abundance has varied between 79,238 (in 2012) and 411,724 (in 1994), with a mean of 216,929 and standard deviation of 87,556. A kernel density estimator was fitted to this distribution, and the cumulative density function was obtained to allow sampling of continuous run sizes in accordance with the historical frequency of run sizes (Figure A.1). The distribution was truncated at the smallest and largest runs on record as of $2017 \pm 30,000$ fish.

A.1.2 Chinook salmon substock composition

Substock composition, or the fraction of the aggregate Chinook salmon run that was made up of fish from each substock, was informed by the proportions of radio telemetry-tagged fish that spawned in each region in the years 2015 and 2016 (Smith and Liller 2017a,b). Although telemetry data from 2003 – 2007 were also available, only these two years were used because: (1) they allowed the incorporation of information from lower river fish (as a result of the tagging location; see Section A.1.3.2) and (2) the management of the fishery resulted in less selection of upper river substocks in the harvest because fishing was pushed later in the season than in the 2003 – 2007 block of years, indicating the 2015 – 2016 block of years are more representative of unfished stock composition.

In each run of the operating model, a random Dirichlet vector was drawn with parameter vector equal to [lower = 19, middle = 61, upper = 20], which results in an expectation roughly equal to the average contribution in 2015 and 2016. The use of a Dirichlet distribution with these parameters generated a modest amount of variability around the expected substock composition.

A.1.3 Chinook salmon run timing

A.1.3.1 Aggregate timing

Run timing information for the aggregate Chinook salmon stock was available from the Bethel Test Fishery (Bue and Lipka 2016), which has produced a daily value of catch per effort for each day between June 1 and August 24 for the years 1984 – 2018. The estimates of location (D_{50}) and inverse scale (h) of a logistic function shown in Table 2.1 were used to quantify the timing with which the simulated aggregate Chinook salmon stock migrates through the lower river.

A.1.3.2 Substock-specific timing

The timing of the specific Chinook salmon substocks (*i.e.*, those spawning in lower, middle, and upper river tributaries) were informed by radio telemetry studies (Stuby 2007; Smith and Liller 2017a,b). The tag date and final tributary of each fish was available for the years 2003 – 2007 and 2015 – 2016. In the first block of years, the tagging site was located near Kalskag, which excluded any fish spawning in lower river tributaries. In the second block of years, the tag site was moved near the Johnson River, which allowed the inclusion of fish spawning in the lower river tributaries. Logistic models (2.1) were fitted to the data from each substock and year separately to obtain estimates of the D_{50} for each substock in each year data were available, and differences in D_{50} for the middle river substocks and each of the other substocks were calculated (Table A.1). For parameterizing the run timing of middle river substocks, random values drawn from the aggregate population estimates were used, and random uniform deviations for the lower river and upper river D_{50} were used in accordance with the deviations shown in Table A.1 (*i.e.*, lower river substocks had a D_{50} value that was anywhere between 0 and 3 days later than that of the middle river, and upper river substocks had a value that was between 5 and 10 days earlier than middle river substocks.

A.1.4 Spatial distribution of escapement

Due to the spatial nature of the operating model, it was important to capture the behavior of fish becoming invulnerable to harvest by swimming up a spawning tributary. This aspect was informed using data from the telemetry studies: it was possible to quantify the fraction of all tagged fish making it to a particular reach that ultimately spawned in a tributary with a confluence in that reach in each year. These fractions were averaged across years and the average was used to dictate how many fish from each substock s in reach r on day d would "peel off" from the main-stem into a tributary in that reach on that day. For the aggregate

184

chum/sockeye stock, which does not have this kind of information, the substock structure was removed. These estimates are shown in Table A.2.

A.1.5   Species ratios

Because chum and sockeye salmon lack the abundance data available for Chinook salmon, their daily entry dynamics were modeled using observed species ratios from the Bethel Test Fishery. These data were prepared by taking the catch per effort of chum salmon plus sockeye salmon, and dividing it by the catch per effort of Chinook salmon on each day of each year for which data were available. This represents how many vulnerable chum/sockeye salmon were available for harvest relative to Chinook salmon. Daily values that could not be calculated (*i.e.*, when zero Chinook salmon were caught) were populated with the average value for all years for which a species ratio could be calculated on that same day. These annual time series were highly variable from day to day, likely as a result of sampling variability, so a cubic spline smoother was fitted to remove this variability. The time series of smoothed ratios from all years is shown in Figure A.2.

In each simulated year, one randomly sampled annual time series was selected to generate the daily species composition for that year. To avoid anomalous outcomes, *i.e.*, unlikely combinations of Chinook run timing and abundance matched with very high or low species ratios in the simulation. I investigated two historical variables for covariance with the species ratio: D_{50} and total Chinook salmon run size using a χ^2 test for independence. For each historical year, run timing, run size, and the first date at which a species ratio of 15:1 was observed were categorized into three bins, with endpoints delineated by the 33% and 66% percentiles of each variable. I was interested in whether Chinook salmon runs with different run timing or size tended to coincide with attaining high species ratios earlier or later in the season. If these sorts of patterns were present, they would need to be accounted for in the simulation.

The first date of 15:1 ratios and Chinook salmon run timing had more non-independence ($\chi^2 = 11, df = 4, p = 0.027$) than Chinook salmon run size ($\chi^2 = 1.84, df = 4, p = 0.765$). This indicated that species ratios could be drawn independently with regards to the simulated Chinook salmon run size, but not the simulated run timing. As shown in Table A.3, the probability of having early high ratios has been historically highest in early Chinook runs. Late Chinook salmon runs tended to occur in years that had later dates of 15:1 ratio attainment. These patterns were incorporated in the operating model by first sampling the run timing for that simulated year, then assigning it to a category, then sampling a ratio category with probability equal to the appropriate column in Table A.3. Finally, a year was randomly selected from the approximately 10 years in that same category, and the daily species ratios that year were used to drive the species composition time series in that simulated year.

A.2   Sociological quantities

A.2.1   Needed salmon harvest by river reach

Herein, the term "minimally needed salmon harvest" salmon harvest refers to the amount of salmon that would satisfy the very basics of the subsistence needs of fishers in the drainage – without meeting this level it is reasonable to assume the fishing population is experiencing hardship. "Maximally needed salmon harvest" represents the salmon harvest that would completely meet subsistence needs (*i.e.*, if as many fish could be harvested as desired). The Alaska Board of Fisheries has produced ranges for each species, termed the "Amounts Reasonably Necessary for Subsistence" (ANS) and represents the drainage-wide range of harvest by species needed to sustain subsistence fishers each year. These ANS ranges are 67,200 – 109,800 for Chinook salmon and 73,400 – 175,100 for chum+sockeye salmon. In this analysis, the lower bound of the ANS range was used to specify minimally needed salmon harvest by species, and the upper bound of the range was used to specify maximally needed

salmon harvests. Maximally needed amounts were used to drive the dynamics of the effort model and the midpoint between the minimal and maximal needs was used to measure the attainment of management objectives.

However, these ANS values are only available for the entire drainage – they are not partitioned to individual villages. For this analysis, a minimal and maximal value was needed for the villages located within each reach. The drainage-wide totals were thus partitioned by calculating the average fraction that villages in each reach have harvested of the drainage-wide total harvest by species. Hamazaki (2011) present year-, species- and village-specific salmon harvests for the period (1990 – 2009), and data through 2015 can be found in Carroll and Hamazaki (2012), Shelden *et al.* (2014), Shelden *et al.* (2015), Shelden *et al.* (2016a), and Shelden *et al.* (2016b). Only years 1990 – 2000 were included for the spatial distribution of salmon need because stakeholders provided input during meetings that indicated the restrictions in recent years make the harvest proportions non-representative and that the earlier years are more reflective of how harvest should be distributed. The partitioned values by species are shown in Table A.4.

A.2.2 Maximum daily effort by river reach

A key aspect of the sociological component to the operating model was the spatial distribution of maximum fishing effort, *i.e.*, the greatest number of boat days that can be exerted by villages in each reach when the fishery is open. This maximum effort was altered as the simulated salmon season progressed based on the effort response submodel. The important characteristic to capture is the proportion of all effort that is attributable to each reach, *i.e.*, the scale of effort is not important as the efficiency of any one unit can be adjusted by altering the q parameter. To determine how effort should be apportioned to each reach, a simple index of effort for each village and year was devised based on the number of reported fishing households residing in each village. The Alaska Department of Fish and Game has collected

this information since 1990, and it is presented in the same studies that quantified subsistence harvest patterns: Hamazaki (2011), Carroll and Hamazaki (2012), Shelden *et al.* (2014), Shelden *et al.* (2015), Shelden *et al.* (2016a), and Shelden *et al.* (2016b). The data were reported as the number of households that "usually fish" and the number of households that "do not usually fish" as surveyed each year (as well as the number of "unknown" fishing status households). First, any unknown households were apportioned to the other two categories by assuming the information was missing at random: *e.g.*, if 60% of the known-status fishing households belonged to the "usually fishes" category in a village in a year, then 60% of the unknown households were apportioned to "usually fishes" and 40% to "does not usually fish". The effort index was calculated for each village as $1 \times \#$ usually $+ 0.5 \times \#$ not usually. Village-specific index values were summed across villages within each reach and year, the annual proportion belonging in each reach was calculated, and the average in each reach across years was obtained.

TABLE A.1: Differences among D_{50} for tagged fish destined for lower or upper river tributaries and those destined for middle river tributaries. These estimates were used to inform Chinook salmon substock-specific run timing.

Year	Lower	Upper
2003		-2.0
2004		-9.5
2005		-4.9
2006		-7.8
2007		-2.5
2015	-0.7	-10.7
2016	2	-9.8

TABLE A.2: Spatial distribution of escapement in the operating model. The number in each cell represents $\psi_{r,s}$: the fraction of fish from a substock that make it to a reach and survive the fishery that ultimately escape and spawn in a tributary with a confluence with the main-stem Kuskokwim located in that reach. These estimates were obtained from radio telemetry studies as described in Section A.1.4, and the chum/sockeye salmon estimates were obtained by removing the substock structure from the Chinook salmon data.

		Chinook Salmon			
Reach #	Tributaries in Reach	Lower	Middle	Upper	Chum/Sockeye
Lower River					
4	Kwethluk	65.3%	0%	0%	12.4%
5	Kasigluk, Kisaralik	80.1%	0%	0%	6%
6	Tuluksak	100%	0%	0%	1.7%
Middle River					
9	Aniak	0%	28.1%	0%	24.6%
10	Owhat	0%	0.5%	0%	0.4%
11	Holokuk, Sue Creek, Veahna	0%	3.7%	0%	3.4%
12	Oskawalik	0%	2.7%	0%	2.4%
13	Crooked Creek, George	0%	6%	0%	4.8%
15	Vreeland, Holitna	0%	77.3%	0%	64.6%
16	Stony	0%	32.8%	0%	25.8%
17	Swift, Tatlawiksuk	0%	100%	0%	55.9%
Upper River					
20	Selatna, Black	0%	0%	6%	6%
22	Takotna	0%	0%	17.5%	17.5%
24	Middle Fork	0%	0%	94%	94%
26	South Fork, East Fork	0%	0%	100%	100%

TABLE A.3: Non-independence of historically observed Chinook salmon run timing and the date at which the species ratio of 15:1 chum+sockeye:Chinook was attained. Columns sum to one and represent the empirical probability of observing a ratio type in each of the three categories along the rows conditional on a run timing scenario in the columns. Independence would have all cells equal to 33.3% – note that early high ratios tend to occur in years with early Chinook salmon runs, and *vice versa*.

	Chinook Salmon Run Timing		
Ratio Category	**Earliest 33%**	**Middle 33%**	**Latest 33%**
Earliest 33%	66.7%	11.1%	20%
Middle 33%	33.3%	44.4%	20%
Latest 33%	0%	44.4%	60%

TABLE A.4: Key sociological quantities used in the operating model, broken down by spatial area (reach; numbers are in order from downstream to upstream). Each reach is 35 km in main-stem river length. Effort ($E_{\mathrm{MAX},r}$) is expressed as the maximum number of boats fishing per day in reach r. The % columns represent the average fraction of the total harvest by species that was harvested by villages within each reach over the period $1990 - 2000$. Harvest values have been rounded to the nearest 100 for ease of presentation, but the total row represents the sum of non-rounded quantities. Although these data were available through 2015, region stakeholders indicated that the recent years have been contaminated by harvest restrictions, and that these earlier years would be more representative.

Reach #	Villages in Reach	Effort	Chinook Salmon			Chum/Sockeye Salmon		
			%	Min.	Max.	%	Min.	Max.
Lower River								
1	Tuntutuliak, Eek	42	7.6%	5,100	8,300	6.2%	4,600	10,900
2	Atmautluak, Kasigluk, Nunapitchuk	74	11%	7,400	12,000	13.6%	10,000	23,900
3	Napakiak, Napaskiak, Oscarville, Bethel	415	40.5%	27,200	44,500	34.1%	25,000	59,600
4	Kwethluk, Akiachak	74	17.2%	11,600	18,900	15.2%	11,200	26,600
5	Akiak	18	4.3%	2,900	4,800	4.9%	3,600	8,600
6	Tuluksak	21	3.9%	2,600	4,300	4.4%	3,300	7,800
Middle River								
8	Lower Kalskag, Upper Kalskag	33	5.1%	3,400	5,600	4.2%	3,100	7,400
9	Aniak	46	4.2%	2,800	4,600	4.6%	3,400	8,100
10	Chuathbaluk	9	1.3%	900	1,400	2.1%	1,600	3,700
13	Crooked Creek	9	1%	600	1,100	1.5%	1,100	2,600
14	Red Devil	6	0.3%	200	400	1.1%	800	2,000
15	Sleetmute	12	1.1%	800	1,200	2.1%	1,500	3,600
16	Lime Village, Stony River	10	0.7%	500	700	4%	3,000	7,000
Upper River								
22	McGrath, Nikolai, Takotna, Telida	42	1.7%	1,100	1,800	1.9%	1,400	3,300
Total		**800**	**100%**	**67,200**	**109,800**	**100%**	**73,400**	**175,100**
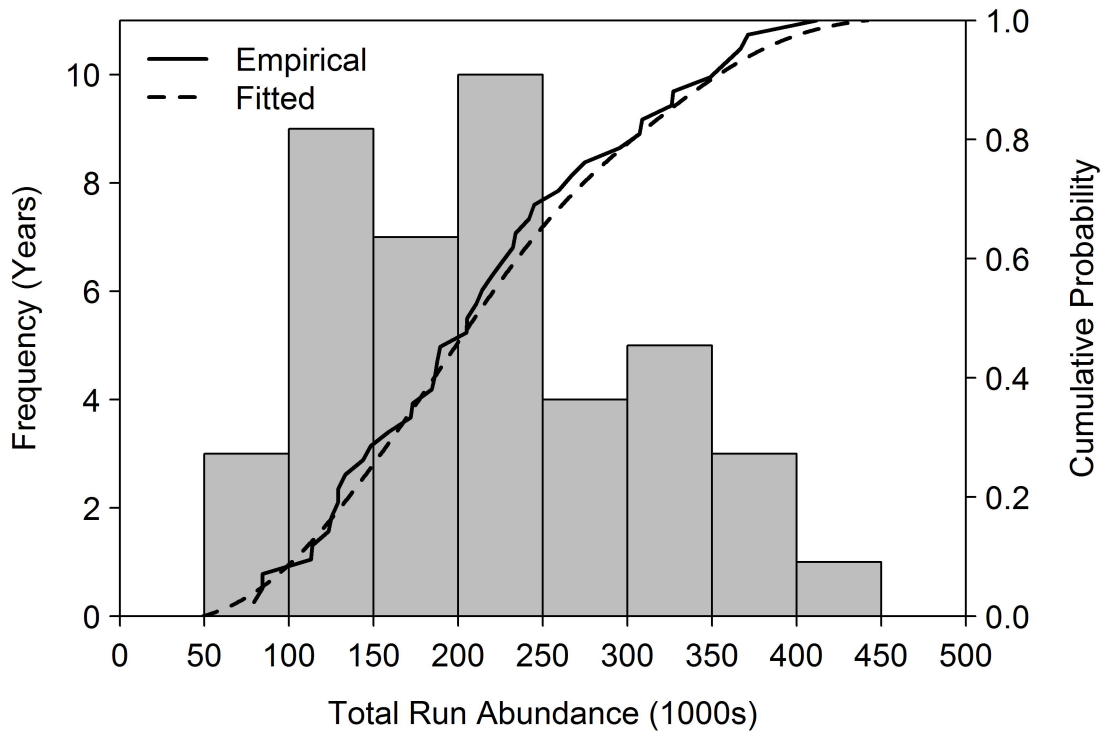
FIGURE A.1: Distribution of total drainage-wide run size for Kuskokwim River Chinook salmon, as presented in Liller *et al.* (2018). This distribution was used to generate the run size of the aggregate Chinook salmon populations entering the fishery system in a simulated year. The secondary *y*-axis represents the probability of a run falling below a given run size according to the historical frequency of run sizes; where the solid line shows the empirical cumulative distribution function and the dashed line shows one obtained by fitting a kernel density smoother to the empirical data. The fitted distribution was used for simulation to prevent the same 42 run size values from being replicated in the analysis.
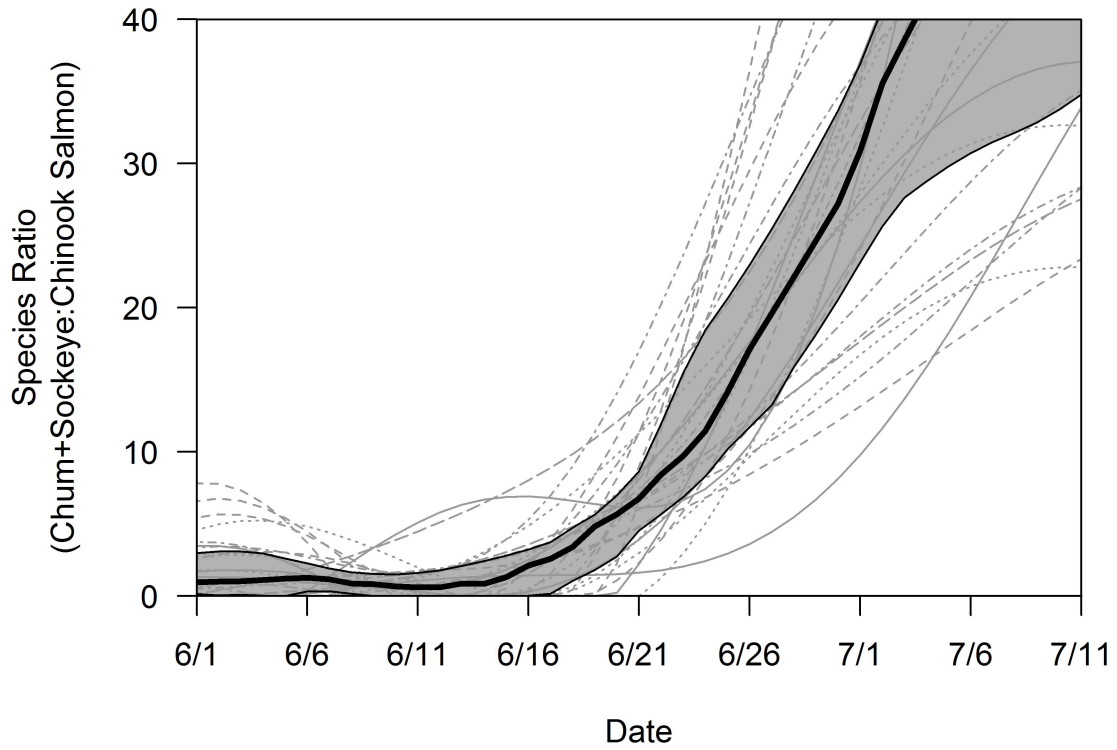
FIGURE A.2: Smoothed species ratios of chum+sockeye:Chinook salmon as detected by the Bethel Test Fishery. Individual grey lines represent separate years from 1984 – 2017, the grey region represents the central 50% of all smoothed ratios on each day and the thick black line represents the daily median. Only this time period is shown because at ratios larger than 20, the differences in the influence of chum/sockeye salmon on Chinook salmon harvest by the subsistence fishery are negligible.

Appendix B

Validation of the Operating Model in Chapter 3

For any simulation model used in the context of management strategy evaluation, the reliability of inferences drawn is conditional on the ability of the model components to capture the important behavioral properties of the real system. Here, a brief validation is provided that the fishery component of the operating model did in fact provide a reasonable representation of the real system for the case where the fishery is unrestricted.

First, it is important that the model be able to replicate the relationship between total Chinook salmon run size and total Chinook salmon subsistence harvest. Capturing this pattern was important to ensure that the fishery would not inadvertently harvest an unrealistically large or small amount of fish in different run sizes than would typically occur, which would confound the inference regarding strategy performance. As shown in Figure B.1, this historical relationship has been quite noisy for the observed historical time series, though an increasing pattern has emerged: in general, more fish have been harvested in years with large runs than years with small runs. It was found that by tuning the catchability (q) and effort response coefficients, this pattern could be reproduced quite well. Additionally, the scale and variability of modeled chum/sockeye harvests were also similar to the historically observed distribution (Figure B.2) – this was not critical given chum/sockeye harvests did not inform any objectives, but the agreement contributes more evidence that the effort response model was adequately calibrated.

The next behavior of interest was the spatiotemporal distribution of harvest. Because in-river salmon fisheries are sequential, fish harvested in one area are invulnerable to harvest (and escapement) in upriver areas. It also means that communities in downriver communities

may finish fishing earlier in the season because they are the first to experience favorable fishing conditions (*i.e.*, high in-river abundance and resulting catch rates; in the Kuskokwim River drying weather also plays an important role). If the timing of harvest was not captured adequately, this would be an indication that the effort response coefficients were improperly tuned and could result in unrealistic conclusions. The patterns and variability in the day of the year at which various percentiles of Chinook salmon harvest were attained by reach compared between observed data and the modeled outcomes are shown in Figure B.3. It seems that the patterns and variability in harvest timing were reasonably well-captured, particularly for downriver reaches. Reaches 14, 15, 16 and 22 seemed to have had the largest deviations between observed and modeled patterns, but given communities in these reaches harvest a negligible amount of Chinook salmon in comparison to the downriver villages (Figure B.4), this finding is not concerning.

The final important characteristic was the spatial distribution of end-of-season harvest. Accurately representing this component of the system would further indicate model adequacy. Figure B.4 shows a comparison of the proportion of total drainage-wide Chinook salmon subsistence harvest attributable to communities in each reach between observed and modeled outcomes. While the overall pattern was fully captured, there were moderate deviations between the model and observations in reaches 2, 3, and 4.
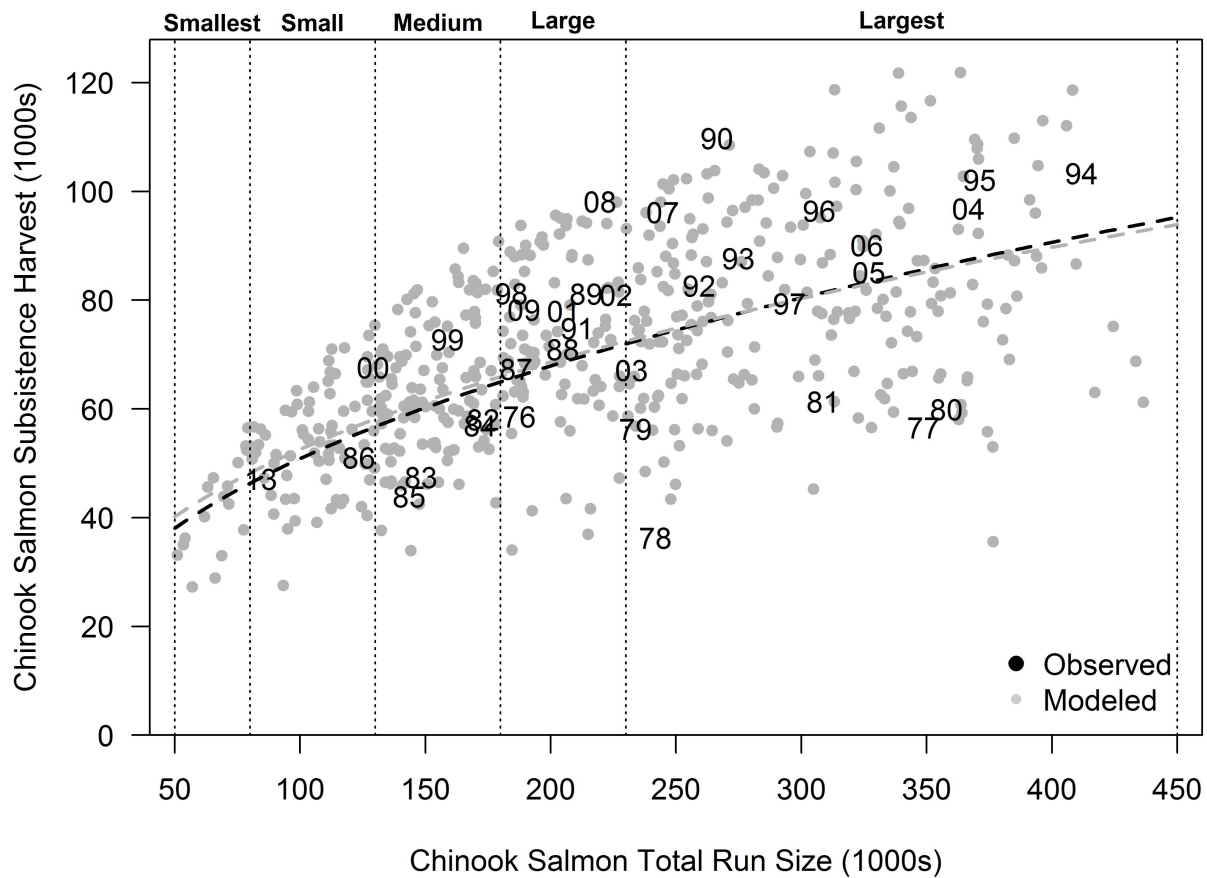
FIGURE B.1: Observed and modeled Chinook salmon subsistence harvest as a function of total Chinook salmon run size. Individual black numbers are historical realizations in years with no harvest restrictions on the subsistence salmon fishery. Individual grey dots are modeled outcomes, each representing a hypothetical salmon run with different random subpopulation compositions, run timing, and species ratios. Fitted models display close agreement between the average simulated and observed harvest outcomes across the range of run sizes. Vertical dotted lines show the boundaries of important run size strata used in this analysis.

197
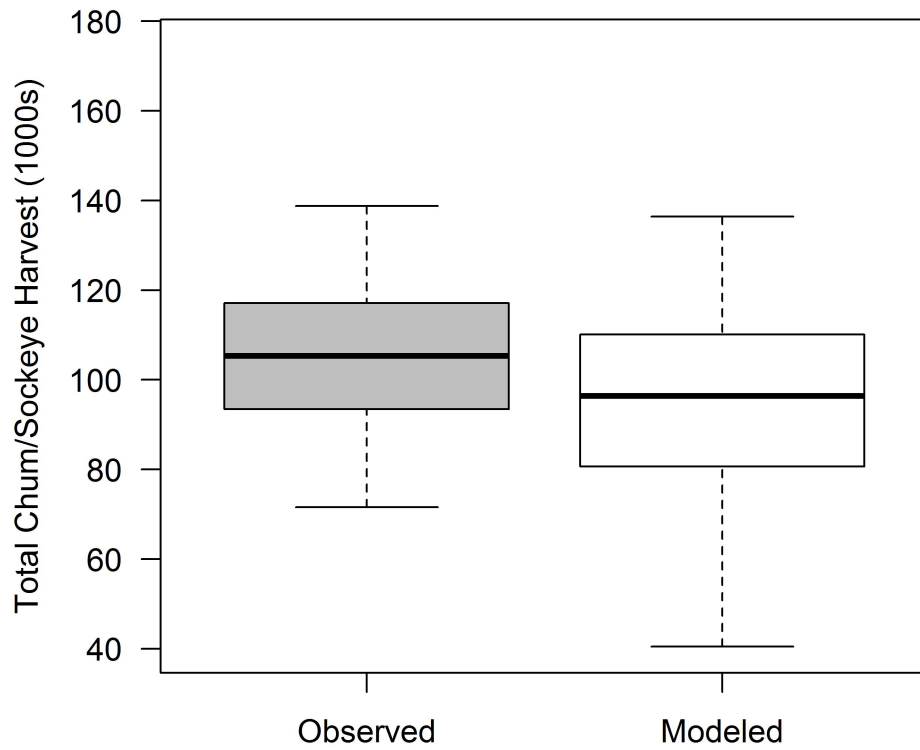
FIGURE B.2: Comparison of the inter-annual distribution of observed and modeled chum/sockeye salmon harvests by all villages located in the Kuskokwim River.
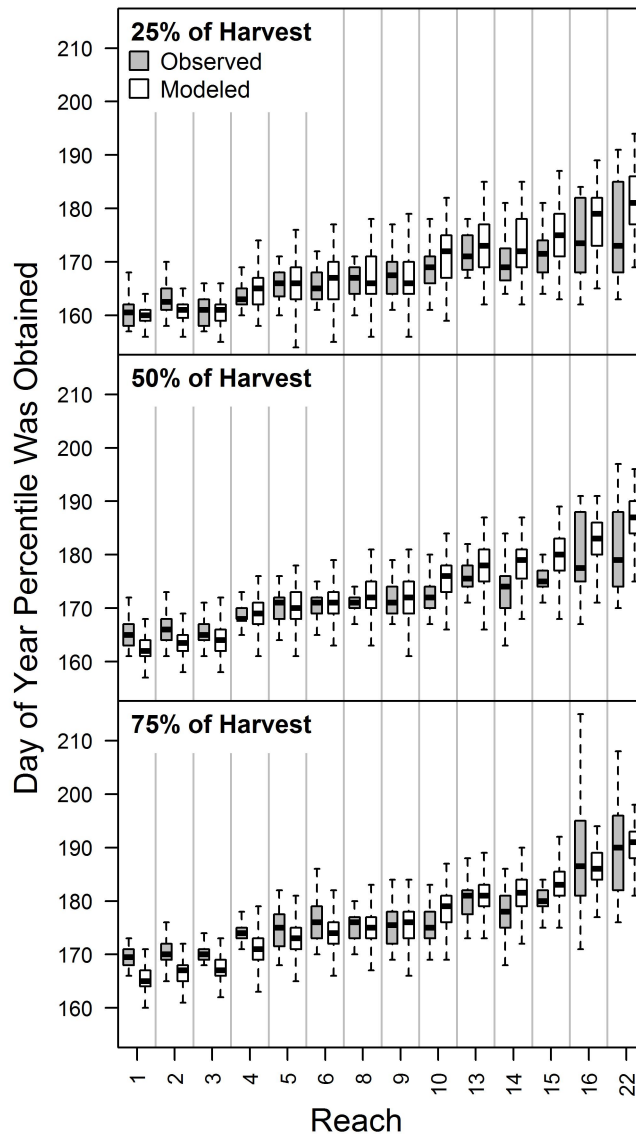
FIGURE B.3: Comparison of the day of the year at which various percentiles of Chinook salmon harvest were attained by reach between observed and modeled outcomes. Variability in the observed boxplots is due to inter-annual variability in run size and timing and represents between-simulation variability for the modeled outcomes. Reach numbers are ordered from downriver to upriver. Note that not all reaches contain communities that harvest salmon.
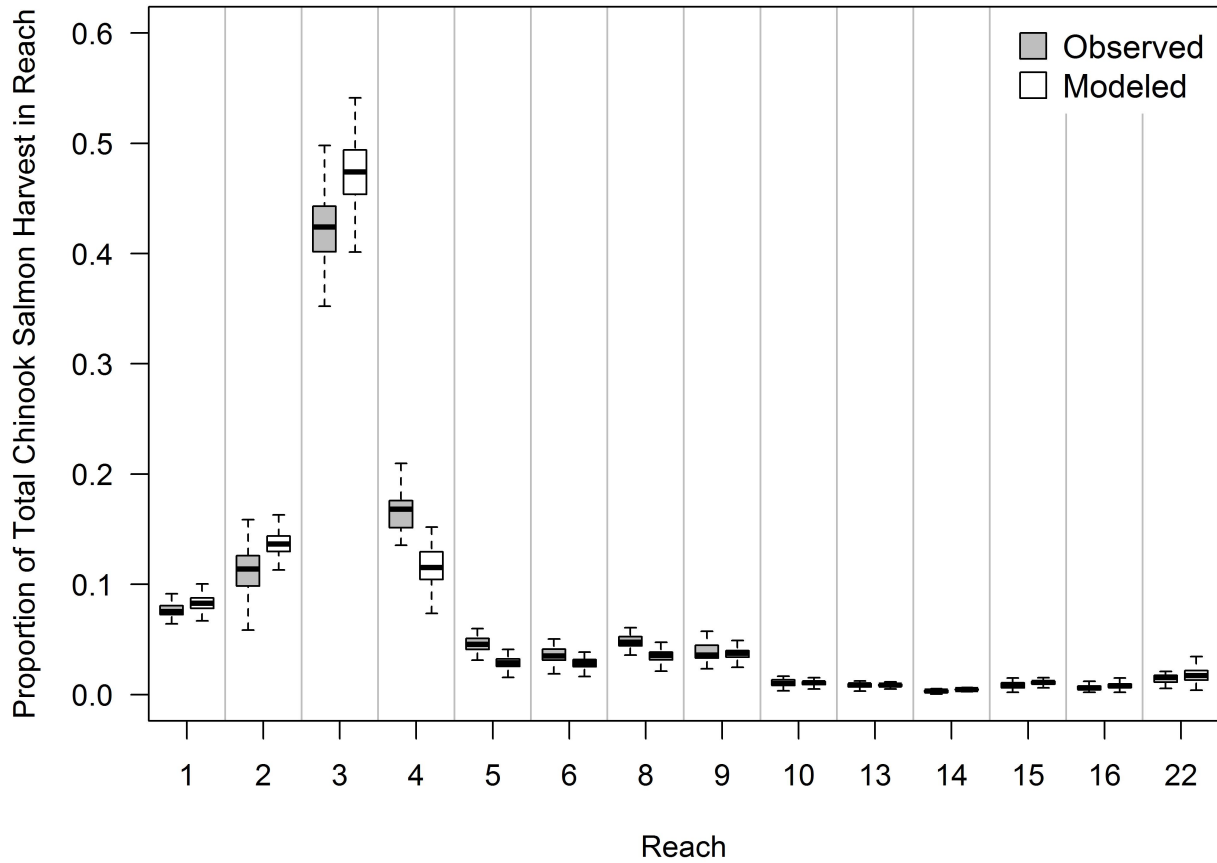
FIGURE B.4: Comparison of the proportion of total drainage-wide Chinook salmon subsistence harvest attributable to communities in each reach between observed and modeled outcomes. Variability in the observed boxplots is due to inter-annual variability, and represents between-simulation variability for the modeled outcomes. Reach numbers are ordered from downriver to upriver. Note that not all reaches contain communities that harvest salmon.

Appendix C

Model Code for Age-Structured Multi-Stock State-Space Spawner-Recruit Model

This appendix presents the JAGS model code for the full state-space model (SSM-VM) in Chapter 4. Variable names used here are intended to be as similar as possible to the symbology used in the main text.

```
model {

  ### PRIORS FOR POPULATION DYNAMICS PARAMETERS ###
  phi ~ dunif(-0.99, 0.99)
  for (j in 1:nj) {
    U_msy[j] ~ dunif(0.01, 0.99)
    log_S_msy[j] ~ dnorm(0, 0.001) %_% I(1, 12)
    S_msy[j] <- exp(log_S_msy[j])
    alpha[j] <- exp(U_msy[j])/(1 - U_msy[j])
    log_alpha[j] <- log(alpha[j])
    beta[j] <- U_msy[j]/S_msy[j]
    log_resid_0[j] <- 0
  }

  ### PRIORS FOR WHITE-NOISE RECRUITMENT VARIABILITY ###
  Tau_R[1:nj,1:nj] ~ dwish(R_wish[1:nj,1:nj], df_wish)
  Sigma_R[1:nj,1:nj] <- inverse(Tau_R)
  for (j in 1:nj) {
    sigma_R[j] <- sqrt(Sigma_R[j,j])
  }
  for (i in 1:nj) {
    for (j in 1:nj) {
      rho_mat[i,j] <- Sigma_R[i,j]/(sigma_R[i] * sigma_R[j])
    }
  }
```

```
### RECRUITMENT PROCESS MODEL ###
for (j in 1:nj) {
  # EXPECTATIONS FOR FIRST A_MAX BROOD YEARS
  R_eq[j] <- log_alpha[j]/beta[j]
  R0[j] <- R_eq[j]
  log_R0[j] <- log(R0[j])
  log_R_mean1[1,j] <- log_R0[j]
  R_mean1[1,j] <- R0[j]
  log_R_mean2[1,j] <- log_R_mean1[1,j] + phi * log_resid_0[j]
  for (y in 2:a_max) {
    R_mean1[y,j] <- R0[j]
    log_R_mean1[y,j] <- log_R0[j]
    log_R_mean2[y,j] <- log_R_mean1[y,j] + phi * log_resid[y-1,j]
  }

  # EXPECTATIONS FOR REMAINING BROOD YEARS
  for (y in (a_max+1):ny) {
    R_mean1[y,j] <- S[y-a_max,j] *
      exp(log_alpha[j] - beta[j] * S[y-a_max,j])
    log_R_mean1[y,j] <- log(R_mean1[y,j])
    log_R_mean2[y,j] <- log_R_mean1[y,j] + phi * log_resid[y-1,j]
  }
}

# LATENT RECRUITMENT STATES
for (y in 1:ny) {
  log_R[y,1:nj] ~ dmnorm(log_R_mean2[y,1:nj], Tau_R[1:nj,1:nj])
  for (j in 1:nj) {
    R[y,j] <- exp(log_R[y,j])
    log_resid[y,j] <- log_R[y,j] - log_R_mean1[y,j]
  }
}

### MATURITY PROCESS MODEL ###
# PRIORS
prob[1] ~ dbeta(1, 1)
prob[2] ~ dbeta(1, 1)
prob[3] ~ dbeta(1, 1)
D_scale ~ dunif(0.03, 1)
```

```
# DIRICHLET HYPERPARAMETERS
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]
D_sum <- 1/D_scale^2

# BROOD YEAR MATURITY VECTORS
for (a in 1:na) {
  dir_alpha[a] <- D_sum * pi[a]
  for (y in 1:ny) {
    g[y,a] ~ dgamma(dir_alpha[a], 1)
    p[y,a] <- g[y,a]/sum(g[y,1:na])
  }
}


### APPORTION RECRUITS TO CALENDAR YEAR RUNS ###
for (j in 1:nj) {
  for (t in 1:nt) {
    for (a in 1:na) {
      N_taj[t,a,j] <- R[t+na-a,j] * p[t+na-a,a]
    }
  }
}


### HARVEST AND ESCAPEMENT PROCESS MODELS ###
for (t in 1:nt) {
  U[t] ~ dbeta(1,1)
  for (j in 1:nj) {
    N[t,j] <- sum(N_taj[t,1:na,j])
    S[t,j] <- N[t,j] * (1 - U[t] * v[j])
    H[t,j] <- N[t,j] * (U[t] * v[j])
  }
  H_tot[t] <- sum(H[t,1:nj])
  log_H_tot[t] <- log(H_tot[t])
}
```

```
### CALCULATE AGE COMPOSTION FOR SUBSTOCKS WITH DATA ###
for (i in 1:n_age_stocks) {
  for (t in 1:nt) {
    for (a in 1:na) {
      q[t,a,i] <- N_taj[t,a,age_stocks[i]]/N[t,age_stocks[i]]
    }
  }
}
### LIKELIHOOD FOR HARVEST OBSERVATIONS ###
for (t in 1:nt) {
  H_tot_t_obs[t] ~ dlnorm(log_H_tot[t], tau_H_obs[t])
}


### LIKELIHOOD FOR ESCAPEMENT OBSERVATIONS ###
# the escapement data are vectorized
# S_obs_t stores the year observed
# S_obs_j stores the substock observed
for (i in 1:S_obs_n) {
  log_S[i] <- log(S[S_obs_t[i], S_obs_s[i]])
  S_obs[i] ~ dlnorm(log_S[i], tau_S_obs[i])
}


### LIKELIHOOD FOR AGE COMPOSITION OBSERVATIONS ###
for (i in 1:n_age_stocks) {
  for (t in 1:nt) {
    x_tas_obs[t,1:na,i] ~ dmulti(q[t,1:na,i], ESS_ts[t,i])
  }
}
}

### END OF MODEL ###
```

Appendix D

Preparation of Data for Fitting Spawner-Recruit Models to Substocks of Kuskokwim River

Chinook Salmon in Chapter 4

D.1 Overview of data needs

All data for this analysis are available to the public, and came primarily from the Arctic-Yukon-Kuskokwim Database Management System (AYKDBMS)[1] maintained by the Alaska Department of Fish and Game (ADF&G). Cases in which other data sources were necessary are highlighted in the description, *e.g.*, the telemetry data needed to perform the expansion of aerial survey counts described in Section D.2 below.

This analysis required three primary data sources:

(1) Estimates of annual escapement to each of the substocks included.

(2) Estimates of annual harvest. Linear regression models (Section 4.2.1.1) required harvest apportioned to each substock, the state-space models (Section 4.2.1.2) required only total aggregate harvest summed across all substocks included.

(3) Estimates of annual age composition (*i.e.*, the fraction of the run each year made up of each age) for all substocks that have had it collected.

Any of these data sources could have missing years.

D.2 Substock escapement

Escapement count data for this analysis were informed predominately by the ADF&G Kuskokwim River salmon escapement monitoring program, the details of which have been

---

[1]http://www.adfg.alaska.gov/CommFishR3/WebSite/AYKDBMSWebsite/Default.aspx

most-recently documented in Head and Smith (2018). The data set available spanned 20 different escapement monitoring projects (six weirs and 14 aerial surveys) and 42 calendar years from 1976 – 2017, though monitoring projects differed in the date they were initialized (Figure D.1). For substocks monitored *via* weir, observed escapement for substock j in year t ($S_{obs,t,j}$) was taken to be the total estimated weir passage each year. Substocks monitored *via* aerial survey needed special care, however. Surveys have been flown only once per year on a relatively small fraction of each tributary system (Figure 4.2), resulting in these data being indices of escapement rather than estimates of total escapement to the subdrainage. This analysis required estimates of total escapement to each substock however, because this would allow calculation of biological reference points that are expressed in terms of the scale of the population (*e.g.*, the spawner abundance that is expected to produce maximum recruitment; $S_{\text{MAX},j}$), rather than as a rate (*i.e.*, $U_{\text{MSY},j}$).

The approach used to estimate total escapement from single-pass aerial surveys involved two main steps:

(1) Mapping the distribution of detected telemetry-tagged Chinook salmon against distribution of the aerial survey counts. This comparison allowed for a spatial expansion to estimate how many salmon would have been counted had the entire tributary been flown. Radio telemetry studies targeting Chinook salmon in the Kuskokwim River have been conducted intermittently since the early 2000s, and data were used from the years 2003 – 2007 and 2015 – 2016. These studies tagged fish migrating through the middle or lower river and (among other things) located them at the end of the season using aerial telemetry (Stuby 2007; Smith and Liller 2017a,b).

(2) Obtaining and applying a temporal correction factor for the problem of counting a dynamic pool at one point in its trajectory. This correction factor was based on the relationship between paired weir and aerial counts on $n = 3$ of the systems in the analysis.

D.2.1 Spatial expansion

The core of the the spatial expansion estimator was the assumption:

$$\frac{A_{f,t,i}}{T_{f,t,i}} = \frac{A_{u,t,i}}{T_{u,t,i}}, \tag{D.1}$$

where the quantities A and T represent fish and tags, respectively, in flown (A_f and T_f) and unflown (A_u and T_u) reaches in year t and for aerial survey monitoring project i. This assumption states that the ratio of actual spawners per one tagged spawner is the same between flown and unflown river sections at the time of the aerial index count and the aerial telemetry flights. Equation (D.1) and can be rearranged as:

$$A_{u,t,i} = A_{f,t,i} \frac{T_{u,t,j}}{T_{f,t,i}}. \tag{D.2}$$

If $T_{u,t,i}$ is further assumed to be a binomial random variable with time-constant success parameter p_i, then:

$$T_{u,t,i} \sim \text{Binomial}(p_i, T_{u,t,i} + T_{f,t,i}). \tag{D.3}$$

Here, p_i represents the probability that a tagged fish in the spawning tributary monitored by project i was outside of the survey flight reach at the time of the aerial telemetry flight. When (D.3) is rearranged to put p_i on the odds scale, then:

$$\psi_i = \frac{p_i}{1 - p_i}. \tag{D.4}$$

The odds value ψ_i can be substituted for the division term in (D.2) which gives:

$$A_{u,t,i} = A_{f,t,i}\psi_i. \tag{D.5}$$

To obtain the total number of fish that would have been counted had the entire subdrainage been flown ($\hat{A}_{t,i}$), the components can be summed:

$$\hat{A}_{t,i} = A_{f,t,i} + A_{u,t,i}. \tag{D.6}$$

Substitution of (D.5) into (D.6) and factoring gives the estimator:

$$\hat{A}_{t,i} = A_{f,t,i}(1 + \psi_i). \tag{D.7}$$

The spatial expansion model was integrated with the temporal expansion model described below into a single model fitted in the Bayesian framework fitted using JAGS (Plummer 2017). This allowed for seamless propagation of uncertainty (in ψ_i) from the expansion above to the next step: a temporal expansion.

D.2.2 Temporal Expansion

A temporal expansion model was necessary to convert from the one-pass index scale to the substock total annual escapement scale. The temporal expansion employed here operated by first regressing $n = 16$ observations of paired weir count (W_i) and spatially expanded aerial counts ($\hat{A}_i$; given by (D.7)) on the same tributary systems ($n = 3$) in the same years:

$$\begin{aligned} W_i &= \beta_0 + \beta_1 \hat{A}_i + \varepsilon_i, \\ \varepsilon_i &\overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma_W^2) \end{aligned} \tag{D.8}$$

The estimated coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ (Table D.2) were then applied to tributary systems with an aerial count but not a weir count:

$$S_{obs,t,j} = \hat{\beta}_0 + \hat{\beta}_1 \hat{A}_{t,j} \tag{D.9}$$

The fitted relationship is shown in Figure D.2. For substocks that had both weirs and aerial surveys, the weir count was used as $S_{obs,t,j}$ as opposed to using the expansion in (D.9) and the coefficient of variation (CV) representing observation uncertainty for the state-space models was set at 5%, which assumed annual escapement counts made at weirs are made with little measurement error. For substocks monitored solely *via* aerial survey, the posterior mean value of $S_{obs,t,j}$ was used as the escapement count that year, and the posterior CV was calculated for use as the observation uncertainty passed to the state-space models. In some cases, multiple aerial survey projects were considered to monitor a distinct portion of a larger substock (*e.g.*, three aerial survey projects count the three forks of the Aniak system; Figure 4.2, Table D.1). For this reason, expansions were conducted for each aerial survey project independently and then summed to obtain the estimates for that substock.

D.3 Aggregate harvest

Harvest estimates for the Kuskokwim River were available at the drainage-wide scale only, and were obtained each year by subtracting the drainage-wide estimates of total run and escapement (Liller *et al.* 2018). Because the escapement data used here did not encompass all the substocks within the Kuskokwim River system, it was necessary to remove some portion of the total harvest that was produced by stocks not included in this analysis. First, the observed exploitation rate of the drainage-wide Kuskokwim River Chinook salmon stock ($U_{obs,t}$) was calculated by dividing the total harvest by the total run each year. Then, the assumption was made that monitored and unmonitored substocks have received the same exploitation rates, in which case total harvest accounted for in this analysis harvest could be obtained as:

$$H_{obs,t} = \frac{S_{obs,t}U_{obs,t}}{1 - U_{obs,t}}, \tag{D.10}$$

which can be derived from the definition of the exploitation rate $\left(U = \frac{H}{S+H}\right)$. This step was embedded within the same Bayesian model that encompassed the spatial and temporal aerial survey expansions such that uncertainty in these steps could be propagated through the entire analysis. The posterior mean value of $H_{obs,t}$ was used as the observed total harvest data, and the posterior CV was retained for use as the observation error attributed to this data source.

Note that $S_{obs,t}$ and $H_{obs,t}$ do not have j subscripts denoting particular substocks: this indicates that they are aggregate quantities summed across all substock components. In cases where substock-specific harvest was required (*i.e.*, in reconstructing the substock-specific brood tables for fitting regression relationships; Appendix D.5), $H_{obs,t,j}$ was obtained using (D.10) by substituting $S_{obs,t,j}$ in for $S_{obs,t}$.

D.4 Age composition

Age composition data were necessary to reconstruct brood tables for age-structured salmon populations (see Appendix D.5). Age data used in this analysis came from the ADF&G standardized age, sex, and length sampling program operated at the weir projects. All sampled fish that were not aged successfully were discarded as were samples corresponding to the rare ages of 3 and 8 such that only fish successfully aged as between 4 and 7 were included. It is possible that older or younger fish may have the systematic tendency to return early or late in the run, and this could introduce biases if age sampling was not conducted proportionally to fish passage throughout the season. To adjust for this possibility, a weighted-average scheme was applied to obtain the age composition estimates for each substock and year with data. Daily age samples were stratified into two-week strata and strata-specific proportions-at-age were calculated. These strata-specific age compositions were then averaged across strata within a year and stock weighted by the number of Chinook salmon estimated to have passed the weir in each stratum. The total number of fish successfully aged for each year and

substock was retained for data-weighting purposes for the state-space models, which (unlike the regression-based approaches) internally reconstructed the brood tables.

D.5 Brood table reconstruction

An important consideration in the use of the regression-based method (Section 4.2.1.1) is in how the $\mathrm{RPS}_{y,j}$ data are obtained for salmon stocks that return at more than one age (like Chinook salmon), especially given the assumption of no observation error. Only the states $S_{y,j}$ are ever directly observed (and generally with some level of error); $R_{y,j}$ is observed (for Chinook salmon) over four calendar years as not all fish mature and make the spawning migration at the same age. Thus, in order to completely observe one $\mathrm{RPS}_{y,j}$ outcome, escapement must be monitored in year y and escapement, harvest, and age composition must be monitored in the subsequent years $y + 4$, $y + 5$, $y + 6$, and $y + 7$. It is evident that missing one year of sampling (which is common; Figure D.1) can lead to issues with this approach. Only completely observed $\mathrm{RPS}_{y,j}$ data were used for this analysis of reconstructing brood tables for regression analysis, with the exception of missing age composition data. For substocks with no age composition data (*i.e.*, those monitored *via* aerial survey), the average age composition each year across substocks that have data was used to reconstruct $\mathrm{RPS}_{y,j}$, but was provided only for years with escapement sampling for substock j. Only substocks with ≥ 3 completely observed pairs of $\mathrm{RPS}_{y,j}$ and $S_{y,j}$ were included for model fitting, given each fitted line was dictated by two parameters.

D.6 Results

D.6.1 Escapement

The escapement project that received the largest spatial expansion factor $(1 + \hat{\psi}_j)$ was the Holitna River, with an expansion factor of 4.78 (4.04 – 5.73; 95% equal-tailed credible limits,

Table D.1). Given the small length of surveyed stream relative to the size of the Holitna subdrainage (Figure 4.2; substock #7), this large estimate makes intuitive sense. The aerial survey project that required the smallest spatial expansion was the Salmon Fork of the Aniak River (1.04; 1.01 – 1.14; Table D.1); note that this project captures nearly all of this tributary of the Aniak system (western-most fork of drainage #4; Figure 4.2). The average spatial expansion factor across all aerial survey projects was 1.78 (1.67 – 1.93).

In terms of the temporal expansion, the estimate for the primary expansion coefficient ($\hat{\beta}_1$) was 2.3 (1.76 – 2.85; Table D.2). This estimate indicates that spatially corrected aerial survey counts needed to be scaled up by a factor of 2.3 in order to be consistent with the total annual escapement counts made at weirs in years and subdrainages that had paired counts. This estimate makes intuitive sense given ADF&G has structured the timing of aerial survey flights to coincide with the peak of the escapement timing arrival curve (Head and Smith 2018), which should occur approximately halfway through the run, indicating the aerial counts would need to be doubled to account for the second half of the run. The freely estimated intercept of the temporal expansion ($\hat{\beta}_0$) had a value of 1.9 (-60.71 – 62.4), indicating that if no fish are counted by an aerial survey, very little escapement actually occurred (Table D.2). For the substocks monitored *via* aerial survey that required these expansions, the average annual observation CV was estimated to be 18% (it was set at 5% for weir-monitored substock escapement).

Based on the scale of the summed escapement estimates from this analysis relative to the drainage-wide scale (which includes monitored and unmonitored substocks; Liller *et al.* 2018), approximately half (57%; 47% – 66%) of the Chinook salmon aggregate population is accounted for by the $n_j = 13$ substocks included here (Figure D.3).

D.6.2 Harvest

The average annual exploitation rate used in this analysis was the same as for the aggregate Chinook salmon population: 0.42 with a minimum and maximum value of 0.13 and 0.62 in 2017 and 1988, respectively. The average annual harvest attributable to substocks included in this analysis was nearly 50,000, which is approximately 57% of the historical average harvest from all substocks in the Kuskokwim River. The average annual observation CV for total harvest was estimated to be 8%.

D.6.3 Age composition

Age composition data were available for six substocks included in this analysis (*i.e.*, those monitored *via* weir projects: George, Kogrukluk, Kwethluk, Takotna, Tatlawiksuk, and Tuluksak; Appendix D.4). The number of years observed per substock depended on the starting year of the project: each year with an escapement count from a weir had associated age data but not all projects have been operating since 1976. The average annual multinomial sample size ($ESS_{t,j}$) varied among substocks (range 117 – 553) and was generally related to the average escapement counted at each project. Across all substocks, the average annual age composition was 27%, 38%, 33%, and 2% for ages 4, 5, 6, and 7, respectively, though there was a high degree of inter-annual variability.

TABLE D.1: The estimated spatial expansion factors for the various aerial survey projects described in Appendix D.2.1. $\hat{p}_i$ represents the average fraction of telemetry tags that were detected outside of index flight reaches, which was used as the basis for determining the multiplier $(1 + \hat{\psi}_i)$ needed to correct the aerial count for not flying the entire subdrainage. In cases where multiple projects were flown to count fish within one substock (*e.g.*, the Aniak, see Figure 4.2 substock #4), the expanded project counts were summed to obtain an estimate for the total substock, as indicated by the footnotes.

Aerial Survey	$\hat{p}_i$	$1 + \hat{\psi}_i$
Kisaralik	0.59	2.46
	$(0.42 - 0.75)$	$(1.72 - 4.04)$
Salmon (Aniak)[a]	0.04	1.04
	$(0.01 - 0.12)$	$(1.01 - 1.14)$
Aniak[a]	0.41	1.71
	$(0.37 - 0.47)$	$(1.58 - 1.87)$
Kipchuk[a]	0.09	1.1
	$(0.04 - 0.17)$	$(1.04 - 1.21)$
Holokuk	0.37	1.59
	$(0.23 - 0.53)$	$(1.3 - 2.12)$
Oskawalik	0.44	1.79
	$(0.29 - 0.6)$	$(1.4 - 2.52)$
Holitna	0.79	4.78
	$(0.75 - 0.83)$	$(4.04 - 5.73)$
Cheeneetnuk[b]	0.25	1.34
	$(0.16 - 0.38)$	$(1.18 - 1.61)$
Gagaryah[b]	0.08	1.08
	$(0.02 - 0.19)$	$(1.02 - 1.24)$
Salmon (Pitka Fork)[c]	0.4	1.66
	$(0.3 - 0.5)$	$(1.42 - 2.01)$
Bear[c]	0.05	1.05
	$(0 - 0.22)$	$(1 - 1.28)$
Upper Pitka Fork[c]	0.62	2.62
	$(0.48 - 0.75)$	$(1.92 - 4)$

Substocks assessed with multiple aerial survey projects
[a] Aniak Substock
[b] Swift Substock
[c] Pitka Substock

TABLE D.2: The estimated temporal expansion parameters for converting spatially expanded aerial counts to estimates of subdrainage-wide escapement abundance each year.

Parameter	Estimate
$\hat{\beta}_0$	1.9 (-60.71 − 62.4)
$\hat{\beta}_1$	2.3 (1.76 − 2.85)
$\hat{\sigma_W}$	4992.15 (3376.54 − 7565.08)
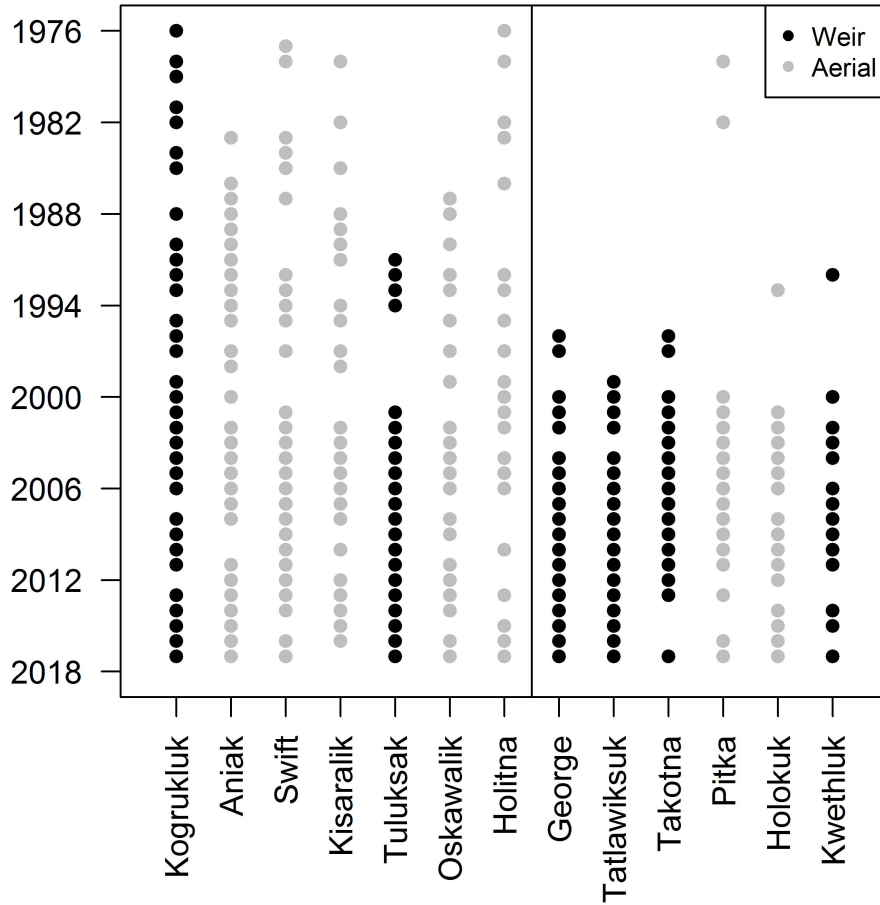
FIGURE D.1: The frequency of escapement sampling for each substock monitored in the Kuskokwim River. Black points indicate years that were sampled for substocks monitored with a weir and grey points indicate years sampled for substocks monitored with aerial surveys. The vertical black line shows a break where > 50% of the years were monitored for a stock.
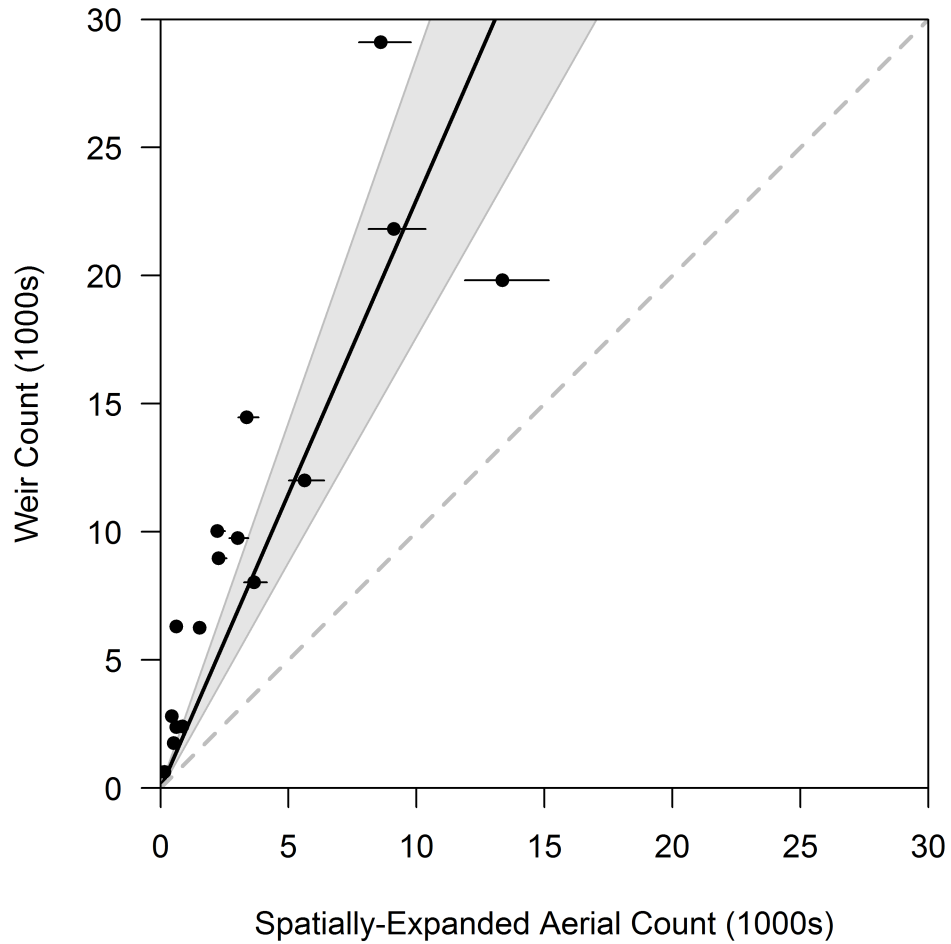
FIGURE D.2: The relationship between spatially expanded aerial survey estimates and weir counts during the same years and substocks as described by (D.8). Notice the uncertainty expressed in the predictor variable; this was included in the analysis by incorporating both the spatial (Appendix D.2.1) and temporal (Appendix D.2.2) expansions in a single model fitted using Bayesian methods.
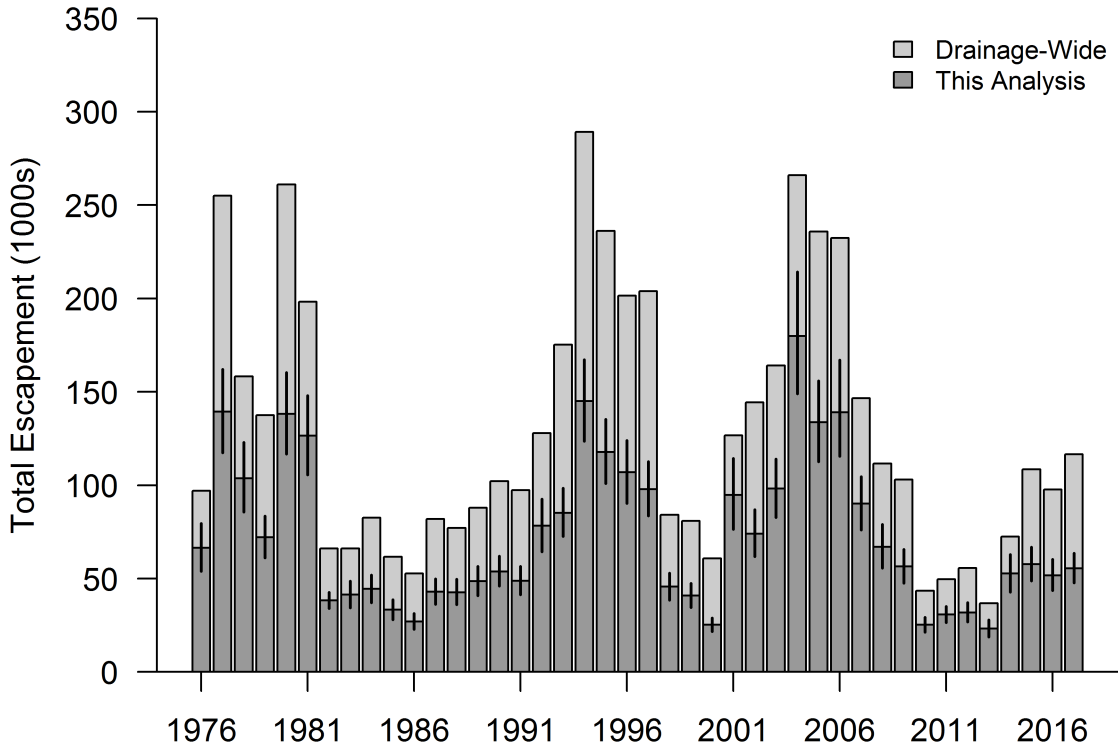
FIGURE D.3: Estimated Chinook salmon escapement for substocks within the Kuskokwim River drainage. "Drainage-wide" refers to the aggregate population estimates provided by a maximum likelihood run reconstruction model. "This analysis" refers to the estimated portion of the aggregate escapement included in this analysis (not all tributaries that produce Chinook salmon in the Kuskokwim River have been monitored; Figure 4.2).

Simulation of Substock- and Year-Specific Maturity Schedules

In the operating model of the simulation-estimation component of Chapter 4, maturity variability was modeled with more complexity than assumed by the most complex estimation model. This was deliberate and allowed assessment of estimation models that did not completely capture the complexity of the processes used to generate the data. Maturity was simulated to vary on average by substock (*i.e.*, some substocks would tend to mature at younger or earlier ages). Additionally, brood-year specific random maturity schedules were generated for each substock, though were simulated to be highly synchronous among substocks. I was unaware of a way to model such correlated dynamics using the Dirichlet random process, so I generated it using a hierarchical linear modeling approach. Note that some of the notation in this appendix uses symbols with different meanings in the main text and other appendices.

E.1 Single substock, single year example

First, define a vector of $n_a - 1$ coefficients:

$$\mathbf{X} = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 \end{pmatrix}$$

and a design matrix with rows and columns equal to $n_a - 1$:

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Then, combine them into a linear predictor on the logit scale:

$$\text{logit}(\boldsymbol{\psi}) = \mathbf{X}\boldsymbol{\gamma}$$

The vector $\boldsymbol{\psi}$ contains the probability of maturity at each age, conditional on not having matured at any previous age for the first $n_a - 1$ possible ages-at-maturation. The elements of $\boldsymbol{\psi}$ can be converted to marginal probabilities of maturing at each age (elements of $\boldsymbol{p}$):

$$p_a = \begin{cases} \psi_a & \text{if } a = 1 \\ \psi_a(1 - \sum_{a=1}^{a-1} p_a) & \text{if } 1 < a < n_a - 1 \\ 1 - \sum_{a=1}^{a-1} p_a & \text{if } a = n_a \end{cases}$$

These marginal probabilities can be then used to apportion recruitments occurring from brood year y to the various calendar years of observation t.

E.2   Extention to multiple substocks and years

Alterations were made to the γ_0 parameter to be specific to each substock and year combination. Stock-level effects were randomly sampled:

$$\varepsilon_j \sim \text{N}(0, \sigma_j),$$

then brood year- and substock-specific random effects were sampled:

$$\varepsilon_{y,1:n_j} \sim \text{MVN}(\varepsilon_{1:n_j}, \Sigma_y),$$

to obtain the year- and substock-specific parameter:

$$\gamma_{0,y,j} = \gamma_0 + \varepsilon_{y,j}.$$

The following parameter values were used: $\gamma_0 = -1.4$, $\gamma_1 = 1.4$, $\gamma_2 = 4$, and $\sigma_j = 0.25$. The covariance matrix Σ_y was constructed with standard deviations for each substock equal to 0.2 and correlation equal to 0.9. These settings resulted in an average vector maturation probabilities of 0.2, 0.4, 0.37, and 0.03 for ages 4, 5, 6, and 7, respectively. An example is shown in Figure E.1.
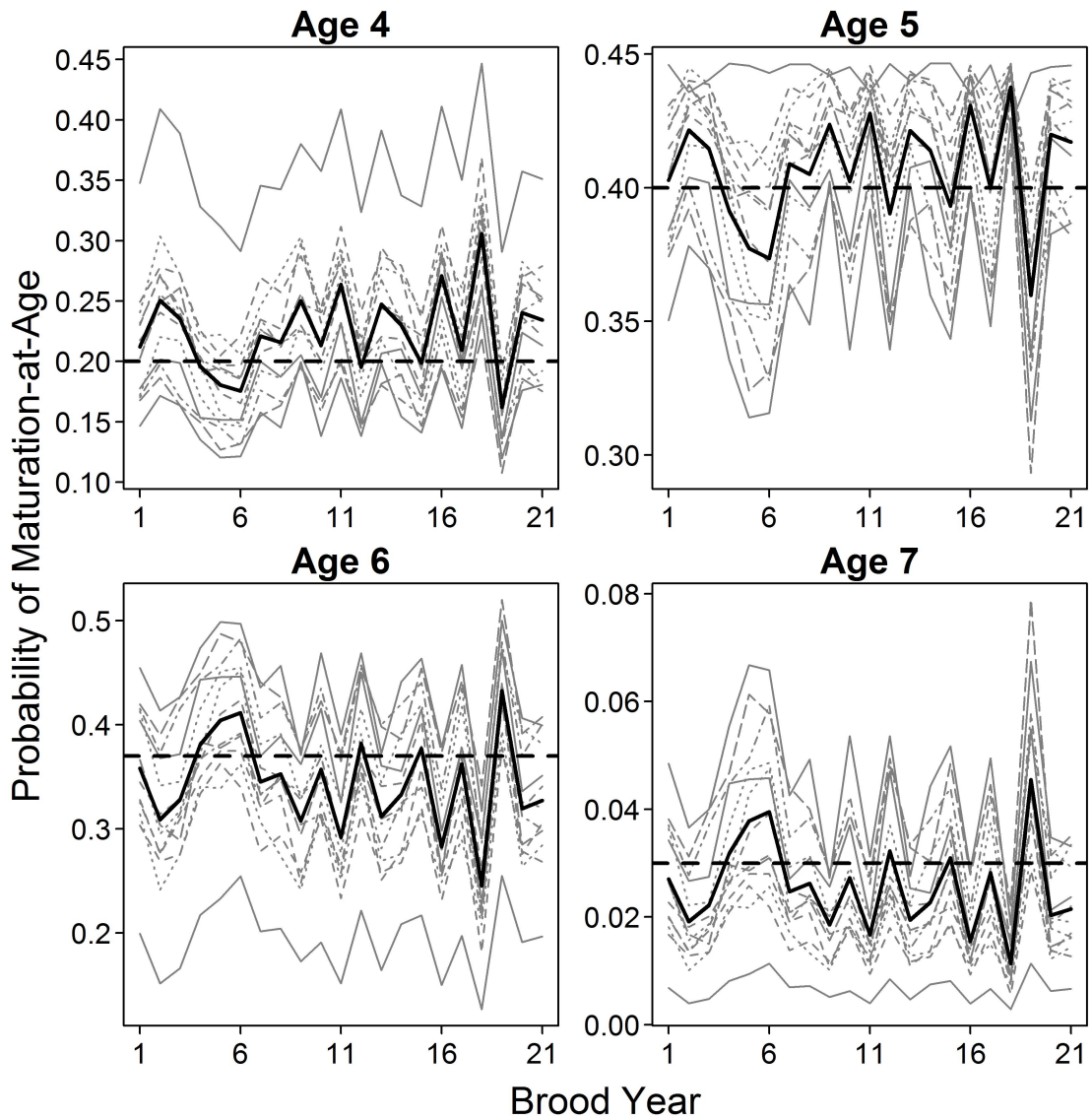
FIGURE E.1: Simulated probability of maturation-at-age for 13 simulated substocks over time. The horizontal dashed line represents maturity without substock or year random effects and the black solid represents the average across substocks. Note the highly correlated patterns in year-specific variability.

Bibliography

Adkison, M. D. and Cunningham, C. J. 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(10):1518–1526.

Adkison, M. D. and Peterman, R. M. 2000. Predictability of Bristol Bay, Alaska, sockeye salmon returns one to four years in the future. *North American Journal of Fisheries Management*, 20(1):69–80.

Agresti, A. 2002. *Categorical Data Analysis*. John Wiley & Sons, Inc., second edition.

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Anderson, J. J. and Beer, W. N. 2009. Oceanic, riverine, and genetic influences on spring Chinook salmon migration timing. *Ecological Applications*, 19(8):1989–2003.

Arlot, S. and Celisse, A. 2010. A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4(0):40–79.

Beer, J. J. 2007. Appendix 7: Run timing of adult Chinook salmon passing Bonneville Dam on the Columbia River. In *Columbia River Salmon Passage (CRiSP) Model Monitoring and Evaluating Support Annual Report, October 1, 2006 - September 30 2007*. U.S. Department of Energy, Bonneville Power Administration, Oregon Division of Fish and Wildlife, Portland, OR.

Blair, G. R., Rogers, D. E., and Quinn, T. P. 1993. Variation in life history characteristics and morphology of sockeye salmon in the Kvichak River system, Bristol Bay, Alaska. *Transactions of the American Fisheries Society*, 122.

Bolker, B. M. 2008. *Ecological Models and Data in R*. Princeton University Press. ISBN 978-0691125220.

Bromaghin, J. F. 2005. A versatile net selectivity model, with application to Pacific salmon and freshwater species of the Yukon River, Alaska. *Fisheries Research*, 74(1-3):157–168.

Brooks, S. P. and Gelman, A. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434.

Bue, B. G., Schaberg, K. L., Liller, Z. W., and Molyneaux, D. B. 2012. Estimates of the historic run and escapement for the Chinook salmon stock returning to the Kuskokwim River, 1976-2011. Fishery Data Series 12-49, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS12-49.pdf [last accessed 2/20/2019].

Bue, D. G. and Lipka, C. G. 2016. Characterization of the 2011 salmon run in the Kuskokwim River based on the test fishery at Bethel. Fishery Data Series 16-05, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska. gov/FedAidPDFs/FDS16-05.pdf [last accessed 2/20/2019].

Burger, C. V., Wilmot, R. L., and Wangaard, D. B. 1985. Comparison of spawning areas and times for two runs of Chinook salmon (*Oncorhynchus tshawytscha*) in the Kenai River, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 42(4):693–700.

Burnham, K. P. and Anderson, D. R., editors 2002. *Model Selection and Multimodel Inference: a practical information-theoretic approach*. Springer New York.

Butterworth, D. S. 2007. Why a management procedure approach? Some positives and negatives. *ICES Journal of Marine Science*, 64(4):613–617.

Carney, J. M. and Adkison, M. D. 2014a. Evaluating the performance of two salmon management strategies using run reconstruction. *North American Journal of Fisheries Management*, 34(1):159–174.

Carney, J. M. and Adkison, M. D. 2014b. Using model simulations to compare performance of two commercial salmon management strategies in Bristol Bay, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(6):814–823.

Carroll, H. C. and Hamazaki, T. 2012. Subsistence salmon harvests in the Kuskokwim Area, 2010. Fishery Data Series 12-38, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS12-38.pdf [last accessed 2/20/2019].

Catalano, M. J. and Jones, M. L. 2014. A simulation-based evaluation of in-season management tactics for anadromous fisheries: Accounting for risk in the Yukon River fall chum salmon fishery. *North American Journal of Fisheries Management*, 34(6):1227–1241.

Clark, R. A., Bernard, D. R., and Fleischman, S. J. 2009. Stock-recruitment analysis for escapement goal development: A case study of Pacific salmon in Alaska. In Krueger, C. C. and Zimmerman, C. E., editors, *Pacific Salmon: Ecology and Management of Western Alaska's Populations*, American Fisheries Society Symposium 70, pages 743–757, Bethesda, MD.

Clark, S. C., Tanner, T. L., Sethi, S. A., Bentley, K. T., and Schindler, D. E. 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(4):829–836.

Cooke, J. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. *ICES Journal of Marine Science*, 56(6):797–810.

Cooke, S. J., Hinch, S. G., Farrell, A. P., Patterson, D. A., Miller-Saunders, K., Welch, D. W., Donaldson, M. R., Hanson, K. C., Crossin, G. T., Mathes, M. T., Lotto, A. G., Hruska, K. A., Olsson, I. C., Wagner, G. N., Thomson, R., Hourston, R., English, K. K.,

Larsson, S., Shrimpton, J. M., and der Kraak, G. V. 2008. Developing a mechanistic understanding of fish migrations by linking telemetry with physiology, behavior, genomics and experimental biology: An interdisciplinary case study on adult Fraser River sockeye salmon. *Fisheries*, 33(7):321–339.

Cooley, R. A. 1963. *Politics and Conservation: the Decline of the Alaska Salmon.* Harper and Row.

Cooperman, M. S., Hinch, S. G., Crossin, G. T., Cooke, S. J., Patterson, D. A., Olsson, I., Lotto, A. G., Welch, D. W., Shrimpton, J. M., Kraak, G. V. D., and Farrell, A. P. 2010. Effects of experimental manipulations of salinity and maturation status on the physiological condition and mortality of homing adult sockeye salmon held in a laboratory. *Physiological and Biochemical Zoology*, 83(3):459–472.

de Valpine, P. and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. *Ecological Monographs*, 72(1):57–76.

DeFilippo, L. B., Schindler, D. E., Ohlberger, J., Schaberg, K. L., Foster, M. B., Ruhl, D., and Punt, A. E. 2018. Recruitment variation disrupts the stability of alternative life histories in an exploited salmon population. *Evolutionary Applications*, 12(2):214–229.

Eiler, J. H., Evans, A. N., and Schreck, C. B. 2015. Migratory patterns of wild Chinook salmon *Oncorhynchus tshawytscha* returning to a large, free-flowing river basin. *PLOS ONE*, 10(4):e0123127.

Fall, J. A., Godduhn, A., Halas, G., Hutchinson-Scarbrough, L., Jones, B., Mikow, E., Sill, L., Trainor, A., Wiita, A., and Lemons, T. 2018. Alaska subsistence and personal use salmon fisheries 2015 annual report. Technical Paper 440, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/techpap/TP440.pdf [last accessed 2/20/2019].

Fleischman, S. J., Catalano, M. J., Clark, R. A., and Bernard, D. R. 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.

Flynn, L. and Hilborn, R. 2004. Test fishery indices for sockeye salmon (*Oncorhynchus nerka*) as affected by age composition and environmental variables. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(1):80–92.

Fried, S. M. and Hilborn, R. 1988. Inseason forecasting of Bristol Bay, Alaska, sockeye salmon (*Oncorhynchus nerka*) abundance using Bayesian probability theory. *Canadian Journal of Fisheries and Aquatic Sciences*, 45(5):850–855.

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. *Ecology*, 84(11):2809–2815.

Habib, E. 2012. On the decomposition of the Schutz coefficient: An exact approach with an application. *Electronic Journal of Applied Statistical Analysis*, 5(2).

Hamazaki, T. 2008. "When people argue about fish, the fish disappear." Fishery closure "windows" scheduling as a means of changing the Chinook salmon subsistence fishery pattern: is it an effective tool? *Fisheries*, 33(10):495–501.

Hamazaki, T. 2011. Reconstruction of subsistence salmon harvests in the Kuskokwim Area, 1990 – 2009. Fishery Manuscript Series 11-09, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FMS11-09.pdf [last accessed 2/20/2019].

Hamazaki, T., Evenson, M. J., Fleischman, S. J., and Schaberg, K. L. 2012. Spawner-recruit analysis and escapement goal recommendation for Chinook salmon in the Kuskokwim River drainage. Fishery Manuscript Series 12-08, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FMS12-08.pdf [last accessed 2/20/2019].

Hamon, T. R., Foote, C. J., Hilborn, R., and Rogers, D. E. 2000. Selection on morphology of spawning wild sockeye salmon by a gill-net fishery. *Transactions of the American Fisheries Society*, 129(6):1300 – 1315.

Hasler, A. D. and Scholz, A. T. 1983. *Olfactory Imprinting and Homing in Salmon.* Springer Berlin Heidelberg.

Head, J. M. and Smith, N. J. 2018. Salmon escapement monitoring in the Kuskokwim Area, 2017. Fishery Data Series 18-11, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS18-11.pdf [last accessed 2/20/2019].

Hendry, A. P. and Quinn, T. P. 1997. Variation in adult life history and morphology among Lake Washington sockeye salmon (*Oncorhynchus nerka*) populations in relation to habitat features and ancestral affinities. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(1):75–84.

Hilborn, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. *Journal of the Fisheries Research Board of Canada*, 36(12):1477–1489.

Hilborn, R. and Walters, C. J. 1992. *Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty.* Springer US.

Hinch, S. G., Cooke, S. J., Farrell, A. P., Miller, K. M., Lapointe, M., and Patterson, D. A. 2012. Dead fish swimming: a review of research on the early migration and high premature mortality in adult Fraser River sockeye salmon *Oncorhynchus nerka. Journal of Fish Biology*, 81(2):576–599.

Hodgson, S., Quinn, T. P., Hilborn, R., Francis, R. C., and Rogers, D. E. 2006. Marine and freshwater climatic factors affecting interannual variation in the timing of return migration to fresh water of sockeye salmon (*Oncorhynchus nerka*). *Fisheries Oceanography*, 15(1):1–24.

Hulson, P.-J. F., Hanselman, D. H., and Quinn, T. J. 2011. Determining effective sample size in integrated age-structured assessment models. *ICES Journal of Marine Science*, 69(2):281–292.

Hyun, S.-Y., Sharma, R., Carlile, J. K., Norris, J. G., Brown, G., Briscoe, R. J., and Dobson, D. 2012. Integrated forecasts of fall Chinook salmon returns to the Pacific northwest. *Fisheries Research*, 125-126:306–317.

Keefer, M. L., Peery, C. A., and Caudill, C. C. 2008. Migration timing of Columbia River spring Chinook salmon: Effects of temperature, river discharge, and ocean environment. *Transactions of the American Fisheries Society*, 137(4):1120–1133.

Keefer, M. L., Peery, C. A., Jepson, M. A., and Stuehrenberg, L. C. 2004. Upstream migration rates of radio-tagged adult Chinook salmon in riverine habitats of the Columbia River basin. *Journal of Fish Biology*, 65(4):1126–1141.

Kellner, K. 2017. *jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses.* R package version 1.4.9.

Kennedy, B. P., Kawachi, I., and Prothrow-Stith, D. 1996. Income distribution and mortality: cross sectional ecological study of the Robin Hood index in the United States. *BMJ*, 312(7037):1004–1007.

Korman, J. and English, K. 2013. Benchmark analysis for Pacific salmon conservation units in the Skeena watershed. A report to Pacific Salmon Foundation.

Liller, Z. W., Hamazaki, H., Decossas, G., Bechtol, W., Catalano, M., and Smith, N. 2018. Kuskokwim River Chinook salmon run reconstruction model revision – executive summary. Regional Information Report 3A.18-04, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/RIR.3A.2018.04.pdf [last accessed 2/20/2019].

Linderman, J. C. and Bergstrom, D. J. 2009. Kuskokwim management area: Salmon escapement, harvest, and management. In Krueger, C. C. and Zimmerman, C. E., editors, *Pacific Salmon: Ecology and Management of Western Alaska's Populations*, American Fisheries Society Symposium 70, pages 541–599, Bethesda, MD.

Ludwig, D. and Walters, C. J. 1981. Measurement errors and uncertainty in parameter estimates for stock and recruitment. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(6):711–720.

Ludwig, D. and Walters, C. J. 1985. Are age-structured models appropriate for catch-effort data? *Canadian Journal of Fisheries and Aquatic Sciences*, 42(6):1066–1072.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C. 2017. A Pacific interdecadal oscillation with impacts on salmon production. *Bulletin of the American Meterological Society*, 78(6):1069–1079.

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(2-3):311–319.

Mundy, P. R. and Evenson, D. F. 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(6):1155–1164.

Neter, J., Kutner, M. H., Nachtsheim, C., and Wasserman, W. 1996. *Applied linear statistical models*. Irwin.

Newman, K. B., Buckland, S. T., Morgan, B. J. T., King, R., Borchers, D. L., Cole, D. J., Besbeas, P., Gimenez, O., and Thomas, L. 2014. *Modelling Population Dynamics*. Springer New York.

O'Malley, K. G., Ford, M. J., and Hard, J. J. 2010. Clock polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B: Biological Sciences*, 277(1701):3703–3714.

Plummer, M. 2017. *JAGS Version 4.3.0 User Manual*.

Prager, M. H., Porch, C. E., Shertzer, K. W., and Caddy, J. F. 2003. Targets and limits for mangement of fisheries: A simple probability-based approach. *North American Journal of Fisheries Management*, 23:349–361.

Punt, A. E., Butterworth, D. S., de Moor, C. L., Oliveira, J. A. A. D., and Haddon, M. 2014. Management strategy evaluation: Best practices. *Fish and Fisheries*, 17(2):303–334.

Quinn, T. P., Unwin, M. J., and Kinnison, M. T. 2000. Evolution of temporal isolation in the wild: genetic divergence in the timing of migration and breeding by introduced Chinook salmon populations. *Evolution*, 54(4):1372–1385.

R Core Team 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A. E. and Lewis, S. 1992. How many iterations in the Gibbs sampler? In *Bayesian Statistics 4*, pages 763–773. Oxford University Press.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G. 2007. Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, 20(22):5473–5496.

Ricker, W. E. 1954. Stock and recruitment. *Journal of the Fisheries Research Board of Canada*, 11(5):559–623.

Rose, K. A., Cowan, J. H., Winemiller, K. O., Myers, R. A., and Hilborn, R. 2001. Compensatory density dependence in fish populations: Importance, controversy, understanding and prognosis. *Fish and Fisheries*, 2(4):293–327.

Salinger, D. H. and Anderson, J. J. 2006. Effects of water temperature and flow on adult salmon migration swim speed and delay. *Transactions of the American Fisheries Society*, 135(1):188–199.

Schaberg, K. L., Liller, Z. W., Molyneaux, D. B., Bue, B. G., and Stuby, L. 2012. Estimates of total annual return of Chinook salmon to the Kuskokwim River, 2002–2007. Fishery Data Series 12-36, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS12-36.pdf [last accessed 2/20/2019].

Schindler, D. E., Armstrong, J. B., and Reed, T. E. 2015. The portfolio concept in ecology and evolution. *Frontiers in Ecology and the Environment*, 13(5):257–263.

Schindler, D. E., Hilborn, R., Chasco, B., Boatright, C. P., Quinn, T. P., Rogers, L. A., and Webster, M. S. 2010. Population diversity and the portfolio effect in an exploited species. *Nature*, 465(7298):609–612.

Schnute, J. T. and Kronlund, A. R. 2002. Estimating salmon stock-recruitment relationships from catch and escapement data. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(3):433–449.

Schutz, R. R. 1951. On the measurement of income inequality. *The American Economic Review*, 41(1):107–122.

Shelden, C. A., Hamazaki, T., Horne-Brine, M., Chavez, R., and Frye, R. 2015. Addendum edition: Subsistence salmon harvests in the Kuskokwim Area, 2013. Fishery Data Series 15-22, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS15-22.pdf [last accessed 2/20/2019].

Shelden, C. A., Hamazaki, T., Horne-Brine, M., Dull, I., and Frye, R. 2016a. Subsistence salmon harvests in the Kuskokwim Area, 2014. Fishery Data Series 16-49, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS16-49.pdf [last accessed 2/20/2019].

Shelden, C. A., Hamazaki, T., Horne-Brine, M., and Roczicka, G. 2016b. Subsistence salmon harvests in the Kuskokwim Area, 2015. Fishery Data Series 16-55, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS16-55.pdf [last accessed 2/20/2019].

Shelden, C. A., Hamazaki, T., Horne-Brine, M., Roczicka, G., Thalhauser, M. J., and Carroll, H. C. 2014. Subsistence salmon harvests in the Kuskokwim Area, 2011 and 2012. Fishery Data Series 14-20, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS14-20.pdf [last accessed 2/20/2019].

Shertzer, K. W., Prager, M. H., and Williams, E. H. 2010. Probabilistic approaches to setting acceptable biological catch and annual catch targets for multiple years: Reconciling methodology with national standards guidelines. *Marine and Coastal Fisheries*, 2(1):451–458.

Smith, N. J. and Liller, Z. W. 2017a. Inriver abundance and migration characteristics of Kuskokwim River Chinook salmon, 2015. Fishery Data Series 17-22, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS17-22.pdf [last accessed 2/20/2019].

Smith, N. J. and Liller, Z. W. 2017b. Inriver abundance and migration characteristics of Kuskokwim River Chinook salmon, 2016. Fishery Data Series 17-47, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS17-47.pdf [last accessed 2/20/2019].

Staton, B. 2018a. *codaTools: Utilities for Dealing with mcmc.lists*. R package version 0.1.1 available at https://github.com/bstaton1/codaTools.

Staton, B. 2018b. *FitSR: Functions to Fit and Summarize SRA Models*. R package version 0.1.2 available at https://github.com/bstaton1/FitSR.

Staton, B. 2018c. *SimSR: Tools for Simulating Pacific Salmon Spawner-Recruit Dynamics*. R package version 0.1.2 available at https://github.com/bstaton1/SimSR.

Staton, B. A. 2018d. In-season harvest and effort estimates for 2018 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. Available at: https://www.fws.gov/uploadedFiles/2018KuskokwimRiverSalmonSubsistenceHarvestReport.pdf [last accessed 2/20/2019].

Staton, B. A. and Catalano, M. J. 2019. Bayesian information updating procedures for Pacific salmon run size indicators: Evaluation in the presence and absence of auxiliary migration timing information. *Canadian Journal of Fisheries and Aquatic Sciences*. Preprint as of 4/16/2019; DOI: https://doi.org/10.1139%2Fcjfas-2018-0176.

Staton, B. A., Catalano, M. J., Farmer, T. M., Abebe, A., and Dobson, F. S. 2017a. Development and evaluation of a migration timing forecast model for kuskokwim river chinook salmon. *Fisheries Research*, 194:9–21.

Staton, B. A., Catalano, M. J., and Fleischman, S. J. 2017b. From sequential to integrated Bayesian analyses: Exploring the continuum with a Pacific salmon spawner-recruit model. *Fisheries Research*, 186:237–247.

Staton, B. A. and Coggins, Jr., L. G. 2016. In-season harvest and effort estimates for 2016 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. Available at: https://www.fws.gov/uploadedFiles/2016KuskokwimRiverSubsistenceSalmonHarvest.pdf [last accessed 2/20/2019].

Staton, B. A. and Coggins, Jr., L. G. 2017. In-season harvest and effort estimates for 2017 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. Available at: https://www.fws.gov/uploadedFiles/2017KuskokwimRiverSubsistenceSalmonHarvest.pdf [last accessed 2/20/2019].

Stuby, L. 2007. Inriver abundance of Chinook salmon in the Kuskokwim River, 2002 – 2006. Fishery Data Series 07-93, Alaska Department of Fish and Game, Anchorage, AK.

Su, Z. and Adkison, M. D. 2002. Optimal in-season management of pink salmon (*Oncorhynchus gorbuscha*) given uncertain run sizes and seasonal changes in economic value. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(10):1648–1659.

Su, Z. and Peterman, R. M. 2012. Performance of a Bayesian state-space model of semelparous species for stock-recruitment data subject to measurement error. *Ecological Modelling*, 224(1):76–89.

Templin, W. D., Smith, C. T., and D. Molyneaux, L. W. S. 2014. Genetic diversity of Chinook salmon from the Kuskokwim River. USFWS Office of Subsistence Management, Fisheries Resource Monitoring Program, Final Report 01-070, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/static/fishing/pdfs/research/fis01-070_final_report.pdf [last accessed 2/20/2019].

van de Pol, M., Bailey, L. D., McLean, N., Rijsdijk, L., Lawson, C. R., and Brouwer, L. 2016. Identifying the best climatic predictors in ecology and evolution. *Methods in Ecology and Evolution*, 7(10):1246–1257.

Walters, C., English, K., Korman, J., and Hilborn, R. 2018. The managed decline of British Columbia's commercial salmon fishery. *Marine Policy*. Pre-print as of 2/18/2019; DOI: 10.1016/j.marpol.2018.12.014.

Walters, C. J. 1985. Bias in the estimation of functional relationships from time series data. *Canadian Journal of Fisheries and Aquatic Sciences*, 42(1):147–149.

Walters, C. J. 1986. *Adaptive Management of Renewable Resources*. Blackburn Press.

Walters, C. J. and Buckingham, S. 1975. A control system for intraseason salmon management. Working Paper WP-75-028, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Walters, C. J. and Hilborn, R. 1976. Adaptive control of fishing systems. *Journal of the Fisheries Research Board of Canada*, 33(1):145–159.

Walters, C. J. and Holling, C. S. 1990. Large-scale management experiments and learning by doing. *Ecology*, 71(6):2060–2068.

Walters, C. J., Lichatowich, J. A., Peterman, R. M., and Reynolds, J. D. 2008. Report of the Skeena independent science reivew panel. A report to the Canadian Department of Fisheries and Oceans and the British Columbia Ministry of the Environment.

Walters, C. J. and Martell, S. J. D. 2004. *Fisheries Ecology and Management*. Princeton University Press.

Wolfe, R. J. and Spaeder, J. 2009. People and salmon of the Yukon and Kuskokwim drainages and Norton Sound in Alaska: Fishery harvests, culture change, and local knowledge system. In Krueger, C. C. and Zimmerman, C. E., editors, *Pacific Salmon: Ecology and Management of Western Alaska's Populations*, American Fisheries Society Symposium 70, pages 349–379, Bethesda, MD.