

Challenges and Tools in the Assessment and Management of Pacific Salmon Fisheries

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 5, 2019

Keywords: Fisheries management, Bayesian inference, decision analysis

Copyright 2019 by Benjamin A. Staton

Approved by

Matthew J. Catalano
Asheber Abebe, Chair, Professor of Mathematics and Statistics
Lewis G. Coggins, Jr., Affiliate Faculty
Conor McGowan, School of Forestry and Wildlife Sciences

Abstract

The single-index varying coefficient model has received much attention due to its flexibility and interpretability in recent years. This dissertation is mainly concerned with the rank-based estimation and variable selection in single-index varying coefficient models.

In the first part of this dissertation, we consider a rank-based estimation of the index parameter and the coefficient functions for single-index varying coefficient model. The consistency and asymptotic normality of the proposed estimators are established. An extensive Monte-Carlo simulation study demonstrates the robustness and the efficiency of the proposed estimators compared to the least squares estimators. The rank-based approach was motivated by a problem from fisheries ecology where it is used to provide accurate estimates of interspecies dependence along an environmental gradient. We use a real data example to show that the classical approach is highly affected by outliers in response space but not the rank-based method we proposed in this dissertation.

The second part of this dissertation is based on variable selection method for single-index varying coefficient model. We develop a LASSO-type rank-based variable selection procedure to select and estimate coefficient functions. A Monte-Carlo simulation study shows that the proposed method is highly robust and efficient compared to least squares type approaches. Our method can be easily applied to single-index and varying coefficient models since they are special cases of single-index varying coefficient model.

Acknowledgments

This dissertation would not have been possible without the support of many people.

First of all, I wish to take this opportunity to express my sincere gratitude to my supervisor, Dr. Asheber Abebe who was abundantly helpful and offered invaluable assistance, support and guidance. I benefit so much from his deep insight into statistics, his professionalism and his constant encouragement. I feel extremely lucky to have him as my thesis advisor.

I am also deeply grateful to my Co-advisor Dr. Nedret Billor for her invaluable advice and helpful comments. Deepest gratitude are also due to my members of the supervisory committee, Drs. Peng Zeng, Guanqun Cao and Huybrechts F.Bindele without whose knowledge and assistance this study would not have been successful.

I wish to thank all professors who taught me in the Mathematics and Statistics department for their excellent lectures, particularly, Drs. Asheber Abebe, Nedret Billor, Peng Zeng, Ming Liao and Jessica McDonald.

I appreciate the friendship with my classmates and friends, including Dr. Yi Xu, Dr. Zhefeng Jiang, Dr. Zhifeng He, Xuyu Wang, Kefan Xiao, Yu Wang, Mingjie Feng, Ninkai Tang, Dr. Hao Wu, Chi Xu, Hannah Correia, Dr. Jessica Busler, and many others.

Finally, I would like to dedicate this thesis to my parents for their boundless love.

Table of Contents

Abstract	ii
Acknowledgments	iii
List of Figures	vi
List of Tables	vii
Preface	1
1 Introduction	2
2 Development and Evaluation of a Migration Timing Forecast Model for Kuskokwim River Chinook Salmon	10
Abstract	10
2.1 Introduction	11
2.2 Methods	16
2.2.1 Estimates of migration timing	16
2.2.2 Environmental variables	17
2.2.3 Selection of predictive time periods	19
2.2.4 Evaluated forecast models	22
2.2.5 Forecast uncertainty	23
2.2.6 Forecast model selection	24
2.2.7 Retrospective forecast analysis	25
2.2.8 Value of forecast to run size assessments	25
2.2.9 Investigation of a run timing versus run size relationship	26
2.3 Results	27
2.3.1 Subsection of results	27

2.4	Discussion	27
3	Run Timing	28
3.1	Introduction	28
3.2	Methods	28
3.3	Results	28
3.3.1	Subsection of results	28
3.4	Discussion	28
4	Multi-stock Spawner Recruit Analyses	29
4.1	Introduction	29
4.2	Methods	29
4.3	Results	29
4.3.1	Subsection of results	29
4.4	Discussion	29
	Bibliography	30

List of Figures

- 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. The structure of these systems is complex (both biologically and sociologically), and capturing that complexity is not the objective of this figure. Rather it is to provide a framework for thinking about where key decisions and uncertainties fall in the big picture. Different resource scientists and managers may have different ways of thinking about this structure; however, this is how I have organized the pieces in my thinking and this proposal. 8

List of Tables

1.1	Here is a table	9
-----	---------------------------	---

Preface

I used bookdown (Xie, 2015) to make this document.

Chapter 1

Introduction

Pacific salmon (*Oncorhynchus* spp.) constitute an integral natural resource in Alaska to subsistence, commercial, and recreational interests. There is a long history of resource development, exploitation, regulation, and dependence on these resources within the region (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 proposed dissertation focuses on the challenges and solutions in the assessment and management of these more remote stocks and the subsistence fisheries that rely heavily upon them.

Wild Pacific salmon represent a fantastic natural resource, which results largely from their unique life history strategy. Pacific salmon exhibit a migratory strategy known as anadromy: they spawn in freshwater where eggs hatch and juveniles rear for several 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 delivers itself to harvesters when the time comes for exploitation.

Like for 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 intrinsically

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 he or she wishes to obtain, getting there is made difficult by not knowing how large the harvestable surplus is, how the stock will respond to harvesting, or that their management action will actually obtain what is desired. 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, I would argue that the science of monitoring, assessment, and prediction in the context of Pacific salmon fisheries is tasked with identifying trade-offs and reducing uncertainty, both of which insert difficulty into the decision-making process as discussed below.

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. At the upper level, long-term decisions are made about the objectives of the resource exploitation. These long-term objectives constitute what could be referred to as fundamental objectives: they are desired endpoints, but do not at all imply how they should be attained. These fundamental objectives often involve notions of sustainability and maintenance of biological diversity and often include social objectives such as maximization and stability of harvest. 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 be obtained by overexploiting weak stock components and possibly reducing 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 purely social and economic and salmon stock assessment scientists should play little-to-no advisory or advocacy roles in making these decisions (Walters and Martell 2004). The Policy for the Management of Sustainable Salmon Fisheries (5 AAC 39.222) 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. Uncertainty at this intermediate level of the management hierarchy (i.e., 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, which are inherently uncertain: data are rarely informative about the shape of the true underlying spawner-recruit relationship but instead provide a probabilistic distribution of expected outcomes at a given stock size (Walters and Martell 2004). Traditionally, it has been thought that these uncertainties can be reduced by more monitoring and the development of 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), true understanding of system behavior comes only from experimentation in management (the concept of “active adaptive management”; Walters 1986). The 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 recruitments must be observed. However, management actions that ensure these observations are made are undesirable to many managers and stakeholders, considering that exploiting a stock down to these low levels is risky (Walters 1986).

At the lower level in the management hierarchy, intra-annual (or in-season) decisions are made regarding how to exploit the current year’s run to attain the long-term objectives. 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. As will be illustrated in this proposed dissertation, these decisions at the intra-annual level of the management hierarchy are often poorly informed by data which often results in indecisiveness, subjectivity, non-transparency, frustration, and missed opportunities.

This proposed dissertation will be partitioned into three chapters, each that expands on and investigates the performance of potential solutions to the aforementioned difficulties in Pacific management in Western Alaska. Each chapter will focus on the Kuskokwim River drainage in Western Alaska, which is characterized by being a large drainage ($>50,000$ km²), harvests are taken by primarily subsistence users who are nearly all native Alaskans, and the primary species of interest being Chinook salmon. Although this proposed dissertation is quite narrow in its geographical and biological focus, a wide range of management issues will be addressed and the developed tools and assessment methods will be evaluated thoroughly. Furthermore, the concepts and tools discussed, developed, and evaluated will be generalizable

to other stocks with similar spatial structures, exploitation characteristics, and population dynamics.

Chapter 2 will work at the intra-annual 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. As will be shown, uncertainty in run timing makes the interpretation of in-season data incredibly difficult. For example, consider a case in which an abundance index has produced high counts early in the season when compared to the historical average. This observation is consistent with at least two run scenarios: (1) a small run with early run timing or (2) a large run with average timing. There is a large discrepancy in the amount of harvestable surplus suggested by each of these discrete scenarios, leading to a large amount of uncertainty about how the fishery should be executed. The overall objective of Chapter 1 will be to develop and evaluate the reliability of a run timing forecast model for Kuskokwim River Chinook salmon. A secondary goal of Chapter 1 will be to formally assess the utility of having access to the run timing forecast model in terms of reducing uncertainty and bias in run size indices used in intra-annual management decisions.

Chapter 2 will again address 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). This analysis will seek to evaluate several harvest control tactics to identify strategies that perform well at attaining pre-defined objectives (e.g., meeting the escapement goal, distributing harvest equally across villages and stock components, etc.) across a range of biological states (e.g., run size, stock composition, and run timing). This analysis is needed because while the fixed escapement strategy seems simple to execute, actually doing so is made difficult largely due to uncertainty regarding the size of the incoming run (i.e., the amount of harvestable surplus is not known). Additionally, there may be a set of tactics that perform well at limiting harvest in low run size years but

doing so in a “fair way”, where the burdens of shortages aren’t carried primarily by a certain subset of resource users. 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, it could prove very useful to managers and decision making within the region.

Chapter 3 will move up the hierarchy to the second level and will seek 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%) 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 are quantified, some understanding of the variation in substock productivity and carrying capacity is required. The multi-stock assessment framework developed in Chapter 3 will be 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 3 will be to evaluate the performance of a range of assessment strategies (in terms of accuracy and precision) across a range of biological and data quality conditions. Figure 1.1. Table 1.1.

LEVEL OF MANAGEMENT HIERARCHY	EXAMPLES	SOURCES OF UNCERTAINTY
Level 1: Overarching Policies and Objectives of Resource Management	<ul style="list-style-type: none"> • Ensure Sustainability • Maximize Harvest • Stabilize Harvest • Maximize Profit • Maximize Participation 	<ul style="list-style-type: none"> • Conflicting objectives • Relative weight of objectives • Utility to society
Level 2: Inter-annual Management Strategies to Obtain Objectives	<ul style="list-style-type: none"> • Constant escapement • Constant exploitation rate • Adaptive exploitation 	<ul style="list-style-type: none"> • Stock productivity • Stock status • Drivers of stock change • Shape/magnitude of trade-offs
Level 3: Intra-annual Tactics to Implement Management Strategy	<ul style="list-style-type: none"> • Means objectives • Management triggers • Time and area restrictions • Gear restrictions • Limited participation 	<ul style="list-style-type: none"> • Harvestable surplus • Uninformative data • Fisher behavior

Figure 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. The structure of these systems is complex (both biologically and sociologically), and capturing that complexity is not the objective of this figure. Rather it is to provide a framework for thinking about where key decisions and uncertainties fall in the big picture. Different resource scientists and managers may have different ways of thinking about this structure; however, this is how I have organized the pieces in my thinking and this proposal.

$$y_i = \beta_0 + \beta_1 x_1 \quad (1.1)$$

Table 1.1: Here is a table

	speed	dist
	Min. : 4.0	Min. : 2.00
	1st Qu.:12.0	1st Qu.: 26.00
	Median :15.0	Median : 36.00
	Mean :15.4	Mean : 42.98
	3rd Qu.:19.0	3rd Qu.: 56.00
	Max. :25.0	Max. :120.00

Chapter 2

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

Abstract

Annual variation in adult salmon migration timing makes the interpretation of in-season assessment data difficult, leading to much in-season uncertainty in run size. We 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 multidimensional 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 2016, the null model had the lowest mean absolute error (2.64 days), although the model-averaged forecast performed as well or better than the null model in the majority of retrospective years. 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. The availability of the run timing forecast was not found to increase overall accuracy of in-season run assessments in relation to the null model, but was found to substantially increase the precision of these assessments, 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 by 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 remains the most precise piece 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 (Cooperman et al., 2010; Cooke et al., 2008; 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 (Quinn et al., 2000; Anderson and Beer, 2009; O'Malley et al., 2010) 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 sub-stock 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 (Keefer et al., 2008; Anderson and Beer, 2009). 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., sub-stocks that must travel farther in-river to reach spawning grounds enter freshwater earlier [Clark et al., 2015]; sub-stocks 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. 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 paper, 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 observations). 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 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 used to forecast run timing with some level of accuracy and certainty.

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 (*O. tshawytscha*) that used air temperature, sea surface temperature, and ice cover to predict the day at which the 15th and 50th percentiles of the run passed a test fishery index location. Model predictions fit the

observed data well (nearly always within seven days, usually within three days), although out-of-sample predictive ability was not presented (Mundy and Evenson, 2011). 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 changes in abundance of genetically distinct populations, which had their own distinct migration timing, affected overall run timing of the 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 efforts.

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 14th and July 2nd. 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 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.,

2017). 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 (previous year's run size estimate \pm approximately 20%) and (2) an in-river drift gillnet test fishery operated in Bethel, AK which has been implemented using consistent methods since 1984. 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 (1) trust a pre-season run size forecast for the majority of the season or (2) take their chances at deciding if the run is early or late 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 (i.e., the unwarranted opening or closing the fishery resulting in severe under- or over-escapement). No published run timing forecast models currently exist for Kuskokwim River Chinook salmon but given the potential utility of independent run timing estimates for interpretation of in-season data, the development and evaluation of such a model is needed. 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), 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.

We present 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, and (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 represented 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, we used daily catch-per-unit-effort (CPUE) data from the Bethel Test Fishery (BTF) operated by the Alaska Department of Fish and Game (ADFG), which spans 1984 – 2016. The raw data were daily CPUE beginning on 1 June and ending 24 August each year. The cumulative sum of these daily CPUE 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 CPUE of the form:

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

where $p_{d,t}$ is the predicted cumulative proportion on Julian day 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 CPUE 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. Uncertainty in

these parameter estimates was not further considered in the analysis as the uncertainty was negligible. Further, by using the BTF daily values to infer the location and shape of year-specific logistic timing curves, we made the assumption that these data provided an accurate representation of daily run strength within a year (influence of weather conditions or harvest on sampling was negligible).

2.2.2 Environmental variables

Environmental variables to be assessed for forecasting performance were chosen based on:

1. previously established association with salmon run timing,
2. availability for the Kuskokwim River during the years for which BTF index observations exist (1984 – 2016), and
3. availability for use in a pre-season forecast model (i.e., available no later than June 10th 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., 1997; <http://research.jisao.washington.edu/pdo/PDO.latest.txt>). The index is produced by taking the first principal component of monthly SST anomalies in the northern Pacific Ocean, after removing any global trends due to any systematic change over time (Mantua et al., 1997). Thus, for each year of the data set, a single monthly value was available for PDO. In the

selection of relevant time periods (Section 2.3), we considered monthly PDO values from January to June of each year, as 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 (http://akclimate.org/acis_data). These data were available as daily means for each day of each year in the 1984 – 2016 data set. Daily air temperature values from the start of January to mid-June of each year were considered in the selection procedures.

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; <http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html>). 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 1) as that is the area that 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 create 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 these areas with stronger signals may 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 2.2.4. Because of the time-intensive nature of the selection procedure described in Section 2.2.3, considered time periods for gridded variables were shortened based on preliminary correlation analyses showing excluded times had little or no correlation with D_{50} . Daily means for SST from mid-April to mid-June and SIC from mid-February to mid-May in each year were considered in the predictive variable selection techniques.

2.2.3 Selection of predictive time periods

Climatic variables are frequently associated with biological quantities for the purpose of prediction, however, 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 time window (i.e., reliably more accurate) exists but is 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 window algorithm (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, this method evaluates all possible windows (subject to certain restrictions) over which to average for use in the forecast model. As input constraints, the sliding window algorithm used in this analysis required: (1) the start date of the first window to be evaluated, (2) the end date of the last window to be evaluated, and (3) the minimum window size of a candidate window. It then proceeded to evaluate all possible

combinations of window averages (using only consecutive days). For example, if the start date of the first possible window was set to 1 May, the minimum window size was set to five days, and the end date of the last possible window was set to 30 May, the window algorithm would start by averaging the daily values between 1 May and 5 May for each year separately and would evaluate the predictive performance of that variable using forecast cross-validation (described below). Next, the algorithm would begin to increase the window size by one day on each iteration, averaging daily values and evaluating the predictive performance for 1 May to 6 May, then 1 May to 7 May, and so on, until it reached the largest window starting with 1 May (which would be 1 May to 30 May). Following completion of all windows starting with 1 May, the algorithm would proceed to evaluate all possible windows starting with 2 May, then 3 May, and so on, until it reached windows starting with 26 May. At this point, there would only be one window to be evaluated, since the minimum number of days to include in a candidate window was set at five days. In the analysis, five days was used as the minimum window size for all variables except PDO (which had monthly values), in which case the minimum window size was set such that a single month could constitute a window. Example R code for how the sliding window algorithm was implemented is provided in Appendix A.

All windows were evaluated using a time series forecast cross-validation procedure, 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 available from years $1, \dots, t$. It then continued for all $t = m, \dots, n - 1$, 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 in D_{50} was calculated based on all forecasted years as $|y_{t+1} - \hat{y}_{t+1}|$, and yearly forecast errors were averaged to obtain mean absolute error (\overline{AE}) which was used as the measure of model performance in window selection. The window with the lowest (\overline{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 it does not make sense for the model to need to predict (for example) year 2010 from years 1984 – 2009 and 2011 – 2016, but rather it would always need to predict year $t + 1$ from all previously-collected data. Example R code for how the forecast cross-validation was conducted is provided in Appendix A.

Although a recent R package has been released to conduct sliding climate window analyses (“climwin”; Bailey and van de Pol, 2015), the package can currently only assess candidate windows using Akaike’s Information Criterion (AIC) or k -fold cross-validation. Because the more specialized forecasting cross-validation technique was required for this analysis, custom R code was written to select the optimal window for each variable (Appendix A).

When forecasting $D_{50,t+1}$ from training data from $1, \dots, t$, 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 2.7), 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 sliding window algorithm was used to select the range of monthly (PDO) or daily (Bethel air temperature) values to include in the predictive climate window for each year in the analysis. For SST and SIC which contained a series of 50 and 49 grid cells, respectively, each with unique daily values, the sliding window algorithm 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 2.2.4.

2.2.4 Evaluated forecast models

Linear regression 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 (i.e., intercept only) model to the full (i.e., global) model (all four variables 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 based on AIC was used to obtain a single forecast D_{50} for each year (Burnham and Anderson, 2002). Under this approach, each grid cell g received an AIC_c score:

$$AIC_{c,g} = n \log(\hat{\sigma}_g^2) + 2K + \frac{2K(K+1)}{n-K-1}, \quad (2.2)$$

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 ΔAIC_c score, representing its relative performance in comparison to the best grid cell:

$$\Delta_g = AIC_{c,g} - AIC_{c,min}, \quad (2.3)$$

where $AIC_{c,min}$ is the minimum AIC_c across all grid cells. Model weights were then calculated as:

$$w_g = \frac{e^{-0.5\Delta_g}}{\sum_j^G e^{-0.5\Delta_j}}, \quad (2.4)$$

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

$$\hat{y}_{t+1} = \sum_g^G w_g \hat{y}_{g,t+1}, \quad (2.5)$$

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

2.2.5 Forecast uncertainty

In addition to forecast accuracy, forecast precision is also of great importance. For models that did not require AIC_c model-averaging across grid cells, the following equation was used to produce a forecast standard error (SE):

$$SE = \hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i^n (x_i - \bar{x})^2}} \quad (2.6)$$

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 equation was used to produce prediction SE:

$$SE = \sum_g^G w_g \sqrt{SE_g^2 + (\hat{y}_{g,t+1} - \hat{y}_{t+1})^2} \quad (2.7)$$

where SE_g is the prediction SE from grid cell g calculated using equation (2.6). 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.

2.2.6 Forecast model selection

Given 16 forecast models, it is impossible to know which will perform the best at forecasting 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 equations (2.2) - (2.7), 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:

$$ASE = \sum_i^{16} w_i \sqrt{\left(\frac{t_{df_i, 1-\alpha/2}}{z_{1-\alpha/2}}\right)^2 SE_i^2 + (\hat{y}_{i,t+1} - \hat{y}_{t+1})^2}, \quad (2.8)$$

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 (i.e., standard normal) distribution. The level $\alpha = 0.05$ was used in all cases.

2.2.7 Retrospective forecast analysis

The analysis was conducted in a retrospective forecast framework starting in 1994. 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 2.2.6, and each was evaluated for predictive accuracy. This process was repeated annually until the present (i.e., out-of-sample predictions made for 1995 – 2016), 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.

2.2.8 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, and is the end-of-season cumulative CPUE observed in the BTF. In-season predictions of EOS were made for each year t , model i , and day d in the season with:

$$\widehat{\text{EOS}}_{d,t,i} = \frac{\text{CCPUE}_{d,t}}{\hat{p}_{d,t,i}}, \quad (2.9)$$

where $\text{CCPUE}_{d,t}$ is the cumulative CPUE caught at the BTF through day d in forecasting year t , $\hat{p}_{d,t,i}$ 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 equation (2.1). Uncertainty in the run timing forecast model was propagated to $\text{EOS}_{d,t,i}$ using the delta method (Bolker, 2008). As the parameter h is also unknown in-season, the mean of all historical h_t was used as the point estimate, and the variance and correlation with $D_{50,t}$ was included in the covariance matrix supplied to the delta method. Accuracy was assessed using proportional bias $(\widehat{\text{EOS}}_{d,t,i} - \text{EOS}_t)/\text{EOS}_t$ and precision was assessed using the SE of EOS on day d . Both accuracy and precision measures were compared between the null and the model-averaged forecast model on 15 June, 30 June, 15 July, and 30 July each year a forecast was available (1995 - 2016). Using the null model to obtain \hat{p}_d is essentially equivalent to what is currently done for predicting $\widehat{\text{EOS}}_d$, which is an index of run size in the Kuskokwim River (i.e., in the absence of an environmental variable forecast model for D_{50}).

2.2.9 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). The forecast absolute errors in each year were then compared using a two-tailed paired t -test using $\alpha = 0.05$.

2.3 Results

We found some stuff

2.3.1 Subsection of results

2.4 Discussion

Chapter 3

Run Timing

3.1 Introduction

Here's chapter 3. It's about in-season simulation models for management strategy evaluations.

3.2 Methods

We did some stuff

3.3 Results

We found some stuff

3.3.1 Subsection of results

3.4 Discussion

Chapter 4

Multi-stock Spawner Recruit Analyses

4.1 Introduction

Here's chapter 4. It's about the multi-stock SRA estimation problem.

4.2 Methods

We did some stuff.

4.3 Results

We found some stuff

4.3.1 Subsection of results

4.4 Discussion

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.
- 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.
- Cooley, R. (1963). *Politics and conservation: the decline of the Alaska salmon*. Harper and Row.
- Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.