Modeling Population Collapse and Recovery in Herring

John T. Trochta

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Trevor A. Branch, Chair

Andre E. Punt

Timothy E. Essington

Program Authorized to Offer Degree:

School of Aquatic and Fishery Sciences

University of Washington

Abstract

Modeling Population Collapse and Recovery in Herring

John T. Trochta

Chair of the Supervisory Committee:

Associate Professor Trevor A. Branch

School of Aquatic and Fishery Sciences

Population collapse in forage fish occurs both naturally and due to overfishing, and is a challenge to sustainable fisheries management. Sustained low abundance can result in prolonged fishery closures and impact the abundance of other species via predation or competition. The time taken to recover from collapse is determined by uncertain factors that control population dynamics and can widely vary between populations. Herring (*Clupea* spp.) are a major group of forage fishes with numerous populations throughout the Northern Hemisphere, that have sustained commercial fisheries for centuries and indigenous fishers for far longer, and support ecologically and economically valuable species including various pinnipeds, whales, seabirds, and predatory fishes. Herring populations across the world have shown varying durations of population collapse since the start of industrial fishing while hypotheses of the underlying factors have been

underdetermined across and within individual populations. The Pacific herring (*Clupea pallasii*) population in Prince William Sound, Alaska is a modern enduring example of prolonged population collapse whose population dynamics remain largely uncertain and unpredictable despite intensive monitoring and modeling.

In this dissertation, I explore and evaluate factors that potentially influence population collapse and recovery within herring. My overarching goal is to better inform the population dynamics of herring and more specifically improve the Bayesian stock assessment model of Prince William Sound herring. In Chapter 1, I conducted a meta-analysis on time series collected for 64 populations worldwide to statistically characterize population collapse and recovery in herring and model predictors of recovery times in adult biomass and recruitment. After collapse, herring populations recovered in 11 years on average, with a few populations remaining collapsed for multiple decades. Amongst populations, recovery time duration did not coincide with fishery closures, which occurred at low abundance in most Pacific herring populations but no Atlantic herring populations. Faster recovery in biomass was best associated with higher average recruitment and higher oceanographic variability in both sea surface height anomalies and sea surface temperatures.

In Chapter 2, I modeled ecological factors impacting natural mortality and recruitment in Prince William Sound herring using a custom-built Bayesian age-structured stock assessment model. Support for individual factors was evaluated using multiple Bayesian model selection criteria and alternative modeling assumptions about the ecological data representing these factors. There was strongest evidence for effects on herring natural mortality from pink salmon abundance in Prince William Sound had the most broad and consistent support. Statistical support differed by the type of selection criteria, model assumptions regarding covariates, and

time period modeled, resulting in generally weak evidence for most individual effects and the suggestion that results are sensitive to model flexibility.

In Chapter 3, I developed a novel modeling framework and conducted a simulation study to test the usefulness of age-specific antibody, or seroprevalence, data in assessing the impact of disease-associated mortality on herring, for use in stock assessment models. Viral hemorrhagic septicemia virus (VHSV) in Prince William Sound herring is used as a case study due to its association with fish kills and well-established ecological principles for its epizootiology from extensive monitoring of VHSV in herring populations. I found that incorporating seroprevalence data within stock assessment can accurately inform infection history and disease mortality and improve population estimates. The first real application of age-specific VHSV seroprevalence is demonstrated with the Prince William Sound herring stock assessment. While motivated from VHSV in herring, these models can be easily adapted to different host populations and pathogens and I present advice for future applications of disease data within stock assessment.

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgements

This work is really the product of many who have supported me throughout my PhD. First and foremost, I acknowledge my advisor Trevor Branch. His guidance, patience, and encouragement were crucial to my growth and instilled the confidence and discipline I needed to become a scientist. He has been the most influential and accommodating mentor during my PhD, in addition to being an amazing human being. Thank you Trevor.

I am also grateful to my committee members for their feedback and advice. Andre Punt pushed the depths of my thinking, while Tim Essington pushed the breadth of it. I thank them both for comments that led to a more intimate knowledge of models and ecology. Thank you to Ole Shelton who not only pressed these boundaries of my understanding as well, but also had my future in mind throughout this process. I greatly valued our conversations that brought into focus the contributions of my research to my career and the field of science at large.

I thank many people associated with the Prince William Sound Science Center and Alaska Department of Fish and Game. Thanks to Scott Pegau for his technical knowledge of Prince William Sound herring biology and monitoring, and feedback on my research. Thanks to Paul Hershberger and Maya Groner for their expertise on fish disease and disease models; they have been wonderful collaborators I hope to have the pleasure of working with in the future. Thanks to Sherri Dressel for in-depth conversations on the stock assessment especially as it relates fisheries management, and the opportunity to present my work at ADF&G. I also appreciate the feedback, conversations, and shared data from many others in the Alaskan herring circles without whom my research and herring knowledge would not have been possible.

The people at SAFS are an amazing group of talented scientists and upstanding individuals, and I am privileged to say I have worked with them. To all members of the Branch

lab over the years including Melissa Muradian, Cole Monnahan, Peter Kuriyama, Merrill Rudd, Caitlin Allen-Akselrud, Lewis Barnett, David McGowan, Stephanie Thurner, and Bia Dias; I am grateful that I shared my experience with you and witnessed your own growth into brilliant scientists I admire. I also thank these people and many others in the SAFS community with whom I made lifelong friendships and memories that I will cherish for the rest of my life.

I must also thank those who started me on the path to graduate school and ultimately becoming a scientist. I am grateful to Colleen Mouw who first took me on as an undergraduate research assistant studying physical variability in Lake Superior at Michigan Tech. Her confidence in me and investment of time, resources, and personal effort into my development at such an early stage in my career still inspires me. I also thank all those who helped me to grow in my early research at Michigan Tech including Gary Fahnenstiel, Audrey and Dave Ciochetto, Brice (and Erika and Angie) Grunert, and Angela Yu.

Lastly, I would like to thank the Exxon Valdez Oil Trustee Council for funding my research and making my PhD possible.

DEDICATION

To all my family. For the enduring support and lessons in dedication, honesty, and modesty from my parents, Jim and Jill. For the courage my sister and brother-and-law, Molly and Boyd, inspired with their own independence and success, and their constant encouragement to find mine. For the kindness and generosity of my family in Spokane (Dan, Barb, Amy, Steve, Amber, Scott, and the boys) that made my journey to and in Seattle easier. Finally, for the patience and unconditional love of my best friend and amazing wife, Erin, who never doubted me even when I did. I love and thank you all.

Introduction

Forage fish are small pelagic fish (many of which are clupeoids) with highly variable abundance (Fréon, Cury, Shannon, & Roy, 2005). Large and irregular fluctuations in abundance cause naturally-occurring collapses in forage fish which is attributed to a stronger responsiveness to environmental variability (Checkley, Alheit, Oozeki, & Roy, 2009; Myron A Peck et al., 2014). Similarly, forage fish are thought to recover more quickly from population collapse than other fish taxa (Hutchings, 2000; Hutchings & Reynolds, 2004; Mullon, Fréon, & Cury, 2005). Long periods of low abundance, or recovery failure, remain unexplained in forage fish (Essington et al., 2015; Neubauer, Jensen, Hutchings, & Baum, 2013) and further contradict hypotheses of increased resilience in shorter lived, fast growing species (Hutchings, 2000; Mullon et al., 2005). Herring and sardines, important species in forage fish research, can take more than a decade to recover even in the absence of exploitation (Baumgartner, 1992; McClatchie, Hendy, Thompson, & Watson, 2017; Petitgas, Secor, McQuinn, Huse, & Lo, 2010). Prolonged periods of low forage fish abundance can be substantially detrimental to both the ecosystem and economy (Pikitch et al., 2014; Smith et al., 2011).

Herring are a scientifically, sociologically, and ecologically important forage fish. Herring populations are perhaps the most widely assessed and managed forage fish. Of 55 assessed forage fish stocks within the RAM legacy database (Ricard, Minto, Baum, & Jensen, 2011) including anchovy, capelin, mackerel, menhaden, pilchard, sandeel, sardine, and sprat, there are 26 herring stocks, or nearly 50% of the forage fish stock assessments. Numerous herring stocks support both small- and large-scale fisheries throughout the north Atlantic and north Pacific Oceans (Hay et al., 2001), attesting to their enduring economic and cultural importance. Herring remain a source of ancient cultural knowledge and tradition in regions such

as the Pacific Northwest (McKechnie et al., 2014; Thornton & Kitka, 2015). Beyond an anthropocentric focus of the importance of herring, they are a key forage fish in coastal ecosystems throughout the Northern Hemisphere, facilitating nutrient and energy transfers from low to high trophic levels (Schrimpf, Parrish, & Pearson, 2012; Willson & Womble, 2006).

Herring have long been the focal study fish for understanding the dynamics of highly fluctuating forage species (Hjort, 1914). Most researchers have focused on the role of external drivers on biological processes like recruitment, adult mortality, or somatic growth. Herring fluctuations have been linked to a long list of biological or physical factors, including ocean temperatures (Nagasawa, 2001; Toresen & Østvedt, 2000), predation (Overholtz & Link, 2006; Schweigert, Boldt, Flostrand, & Cleary, 2010; Willson & Womble, 2006), river discharge (Gilbert, 1997; Stocker, Haist, & Fournier, 1985), disease (Marty, Ii, Carpenter, Meyers, & Willits, 2003), food availability (Möllmann, Kornilovs, Fetter, & Köster, 2005), and pollution (Carls, Rice, & Hose, 1999). However, these studies are accompanied by broad uncertainty that hinders accurate predictions over time, especially during periods of collapse (Neubauer et al., 2013). While new factors and explanations are continuously being explored for a single stock, revisiting previously published hypotheses and conclusions with new data or different methods serves to improve and validate these findings.

Most herring populations collapsed at some point during the 20th century due to overfishing (Hay et al., 2001). While many of the herring fisheries subsequently recovered after fishing stopped and management approaches reactively improved, preventing collapse remains difficult. Overfishing has been shown to amplify both the frequency and magnitude of collapses of forage fish populations (Essington et al., 2015). Still, overfishing only partially clarifies the onset and nature of collapse, an event that is frequent for herring.

Explaining post-collapse dynamics is at least as important as explaining the collapse itself, while being more poorly understood. Intense overfishing that continues through collapse delays recovery (Neubauer et al., 2013), though still does not appear to explain differences in recovery time among forage fish stocks (Essington et al., 2015). Anecdotally, herring show a wide range of recovery times (Hay et al., 2001; Pearson, Elston, Bienert, Drum, & Antrim, 1999). For Chapter 1, I address the following: 1) How often and for how long should we expect low biomass and recruitment to occur for herring? 2) Which factors are most important in explaining the frequencies of low biomass and high recruitment, and 3) Which factors are most important in explaining the maximum lengths of time at which biomass and recruitment can remain low? I created a database of Pacific and Atlantic herring time series from worldwide. Using these time series, I used generalized linear models and random forest regression to relate recovery times in herring abundance and recruitment to environmental, ecological, and fishing factors (Trochta, Branch, Shelton, & Hay, 2020). This is the first study to examine the recovery potential of herring generally, as opposed to a single stock or across multiple taxa.

The collapse of Pacific herring in Prince William Sound is an enduring, enigmatic example of a forage fish collapse. Herring adult biomass declined from ~120,000 metric tons in 1989 to less than 30,000 tons by 1993. This collapse coincided with the Exxon Valdez Oil Spill of 1989, but it has long been debated whether the oil spill caused the decline (Pearson et al., 1999; Pearson et al., 2012). Except for two years (1997 and 1998), the fishery has been closed since 1993, and has not experienced any strong recruitment classes. The failed recovery of Prince William Sound herring is not well understood.

The herring of Prince William Sound likely experienced compensatory effects at their peak abundance in the late 1980s that combined with reduced food supply to lead to their demise

around 1993 (Pearson et al., 1999). The first signs of a widespread outbreak of the viral hemorrhagic septicemia virus (VHSV) were observed in 1993 (Meyers & Winton, 1995), and VHSV has been suggested to increase adult mortality during the collapse and the subsequent period of failed recovery (Marty et al., 2010; Marty et al., 2003). As recovery failure became more evident, the focus shifted to other potential factors impacting recruitment. Studies have reported a link between hatchery-releases of pink salmon and herring recruitment dynamics (Deriso, Maunder, & Pearson, 2008; Pearson et al., 2012), although mechanistic evidence for predation or competition is lacking at present (Sturdevant, Brenner, Fergusson, Orsi, & Heard, 2013). Overwintering juvenile survival is also assumed to be critical in determining eventual recruitment to adult biomass (Pearson et al., 2012), but the drivers of year-to-year variation in survival are uncertain. Increased humpback whale predation is also likely, given the upward trending presence of summer and overwintering whales in the Gulf of Alaska which feast heavily on juvenile herring schools (Pearson et al., 2012; Straley et al., 2017).

Modeling studies have been a critical part of the effort to synthesize information on various hypotheses of biomass collapse and recruitment drivers. Previous studies of these factors each found different significant factors including air temperature (Williams & Quinn, 2000), zooplankton availability (Brown & Norcross, 2001), hatchery pink salmon (Deriso et al., 2008), freshwater input (Ward et al., 2017), and Gulf of Alaska juvenile walleye pollock (Sewall, Norcross, Mueter, & Heintz, 2017). Most of these analyses used post-hoc recruitment estimates from the age-structured assessment (ASA) model used by the Alaska Department of Fish & Game (ADF&G) with a spawner-recruitment function in a GAM-type framework. These approaches do not directly estimate recruitment, and rely on strong structural assumptions about productivity coming from the stock-assessment-based estimates of recruitment. Only one

analysis also tested hypotheses on adult survival and concluded there was weak support for nutrition and winter ocean temperatures as correlates of survival (Deriso et al., 2008).

Whereas Chapter 1 is broad in the scope of what may explain failed herring recovery following collapse, Chapter 2 specifically addresses this same question in relation to Prince William Sound herring recovery. I re-evaluated covariates supported from these previous studies and evaluated new ones using Bayesian model selection that directly weighs different hypotheses against each other. For model selection, I incorporated covariates into the most recent version of the ASA model which is implemented in a Bayesian framework. The Bayesian age-structured assessment for Prince William Sound herring is specifically referred to as BASA in this dissertation. Covariates were modeled to impact either recruitment and adult natural mortality in regulating Prince William Sound herring population dynamics. I found the most support for at least one covariate (total pink salmon) and ambiguous support for several others.

Disease in particular has long been hypothesized as an critical factor in Prince William Sound herring population dynamics (Marty et al., 1998; Marty et al., 2003). The collapse of Prince William Sound herring in 1993 coincided with a widespread outbreak of external gross lesions in Prince William Sound herring from which a strain of viral hemorrhagic septicemia virus (VHSV) was isolated (Meyers et al., 1994). Increased prevalence of another infection by *Icthyophonus hoferi* in older herring was also observed post-collapse (Marty et al., 1998), especially after the large VHSV outbreaks of the 1990s (Marty et al., 2010). While initially presumed to be a primary mechanism for the 1993 collapse and a significant suppressing factor thereafter, information on the prevalence of these epidemics has been criticized for not providing reliable estimates of adult mortality in Prince William Sound herring (Elston & Meyers, 2009; Hershberger, Garver, & Winton, 2015; Pearson et al., 2012).

The role of disease in herring mortality has motivated its current use in the ASA and Bayesian ASA models used for management (Marty et al., 2010; Muradian, Branch, Moffitt, & Hulson, 2017). The effect of disease is measured with *I. hoferi* and a combined VHSV/ulcer prevalence index (e.g. percent infection). Though separate infections, years with high prevalence of ulcers related to bacterial infections only occurred with high VHSV prevalence, and the index combining both infection types improved ASA model predictions while individual indices for each did not (Marty et al., 2010; Marty et al., 2003). These models incorporate disease prevalence as an additive effect on constant baseline mortality with the effect starting in 1992 onward, implicitly assuming prevalence is proportional to the mortality from infection. However, the current disease prevalence index provide a very narrow interpretation of infection; it does not detect surviving fish that no longer show signs of infection since outbreaks are rapid, nor does it measure actual mortality since dead fish are not sampled (P. Hershberger, USGS, pers. comm.). As a result, the current index fails to accurately represent the overall disease mortality in the manner it is included in the BASA model. Furthermore, combining VSHV and ulcer indices has no basis in the epizootiology, or the ecological principles underlying the dynamics of VHSV.

To overcome these issues, new data from recent field- and laboratory-based studies of VHSV in herring have been developed: measurements of antibody prevalence. Antibodies are produced only after VHSV exposure and accumulate over years within cohorts that experience multiple infections over their lifetime. In other words, seroprevalence offers a synoptic overview of infection history. Chapter 3 uses a simulation analysis to evaluate the utility of antibody prevalence data (seroprevalence), in informing population-wide infection and disease-associated mortality estimates in fisheries stock assessment. We demonstrate that not only are seroprevalence data useful in estimating age- and time-varying effects of disease-associated

mortality, but their use also provides a substantial improvement over current use of a VHSV prevalence index in the Prince William Sound herring assessment model, even when the model is misspecified. After demonstrating through simulations that VHSV seroprevalence data provide useful information for stock assessment, we included the real VHSV seroprevalence data in the BASA model for Prince William Sound herring, to estimate mortality and disease prevalence.

Overall, my dissertation provides insights into and improves models of herring population dynamics, especially regarding processes of population collapse and recovery. These processes largely remain unclear due to a myriad of factors including environmental and ecological variability, population-specific characteristics and considerations, observation error, and model misspecification and uncertainty. This dissertation demonstrates modeling approaches that better account for and reduce uncertainty by better quantifying impacts of these factors. Additionally, much of this work has served to advance the stock assessment of Prince William Sound herring and contribute more robust methods for evaluating population status and trends.

Chapter 1. The highs and lows of herring: A meta-analysis of patterns and factors in herring collapse and recovery

1.1 ABSTRACT

Pacific and Atlantic herring populations (genus Clupea) commonly experience episodic collapse and recovery. Recovery time durations are of great importance for the sustainability of fisheries and ecosystems. We collated information from 64 herring populations to characterize herring fluctuations and determine the time scales at low biomass and at high and low recruitment, and use generalized linear models and Random Survival Forests to identify the most important bottom-up, top-down, and intrinsic factors influencing recovery times. Compared to non-forage fish taxa, herring decline to lower minima, recover to higher maxima, and show larger changes in biomass, implying herring are more prone to booms and busts than non-forage fish species. Large year-classes are more common in herring, but occur infrequently and are uncorrelated among regionally grouped stocks, implying local drivers of high recruitment. Management differs between Pacific and Atlantic herring fisheries, where at similarly low biomass, Pacific fisheries tend to be closed while Atlantic fisheries remain open. This difference had no apparent effect on herring recovery times, which averaged 11 years, although most stocks with longer recovery periods had not yet recovered at the end of the observation period. Biomass recovery is best explained by median recruitment and variability in Sea Surface Height anomalies and Sea Surface Temperatures—higher variability leads to shorter recovery times. In addition, the duration of recruitment failure is closely linked with low biomass. While recovery times rely on the nature of the relationship between spawning biomass and recruitment, they are still largely governed by complex and uncertain processes.

1.2 INTRODUCTION

Population collapses of exploited forage fish may lead to serious socio-ecological repercussions and a significant challenge to sustainable fisheries management. Atlantic herring (Clupea harengus) and Pacific herring (Clupea pallasii) are commercially and culturally important small-bodied pelagic fish that play an integral role in coastal ecosystems in the world's northern oceans (Smith et al., 2011). Fishery closures prompted by low herring abundance deprive the livelihoods of local fishing communities and affect long-standing traditions that define cultural identities focused on herring, especially in the North Pacific (Gauvreau, Lepofsky, Rutherford, & Reid, 2017; Hamada, 2015; Jones, Rigg, & Pinkerton, 2017; Menzies, 2016; Thornton & Kitka, 2015). Prolonged closures or poor landings translate to widespread economic effects, including the collapse of markets based on herring (Dickey-Collas et al., 2010). Herring collapse may also lead to significant changes in key predator-prey interactions since herring are an important prey linking primary production to the highest level consumers including marine mammals, seabirds, and larger fish (Ainsworth, Pitcher, Heymans, & Vasconcellos, 2008; Smith et al., 2011; Surma, Pakhomov, & Pitcher, 2018a; Surma et al., 2018b); although, how much predators are truly impacted is debated (Hilborn et al., 2017). The seasonal migration of Pacific herring in particular to spawn in intertidal and upper subtidal water provides an annual pulse of marine nutrients to marine and terrestrial predators (Fox, Paquet, & Reimchen, 2018; Fox, Paquet, & Reimchen, 2015; Willson & Womble, 2006).

Most exploited herring populations collapsed in the 20th century and overfishing was implicated as the most prevalent cause of collapse (Hay et al., 2001). Many herring fisheries recovered after fishing stopped and while management approaches have improved, they still cannot anticipate and prevent all causes of collapse. Natural variability for herring is similar to

that of other forage fish, which display large and irregular fluctuations attributed to a tighter coupling with environmental variability (i.e. through bottom-up forcing) and where collapses occur naturally (Checkley et al., 2009; M. A. Peck et al., 2014).

Explaining post-collapse population dynamics is important for understanding recovery. Across all fish taxa, intense overfishing that continues through collapse periods delays recovery (Neubauer et al., 2013), but for forage fish, intense overfishing does not explain differences in recovery time (Essington et al., 2015). Furthermore, while some argue that forage fish are more likely to recover than other fish taxa (Hutchings & Reynolds, 2004), others suggest they are more vulnerable to collapse because of schooling behavior effects on catchability (Pitcher, 1995) and show enhanced sensitivity to environmental variability resulting from heavy exploitation (Essington et al., 2015; Pinsky & Byler, 2015; Pinsky, Jensen, Ricard, & Palumbi, 2011).

Several prominent examples of herring collapse (Hay et al., 2001; Pearson et al., 1999) highlight the lack of resilience for at least some populations. Notably, the large Norwegian spring spawning stock of Atlantic herring crashed to record low spawning biomass (<1 million tons) during 1968–1988 compared to averages of 8 million tons before this period and 5 million tons afterwards. Similarly, Hokkaido-Sakhalin herring in the western Pacific experienced a steady decline through the mid-1950s after catches peaked at nearly 1 million metric tons in 1897. Since then, catches have remained below 40,000 metric tons, with no indication that a resumption of large-scale commercial fishing is probable. Finally, Prince William Sound herring (Gulf of Alaska, Northeast Pacific) declined from 130,000 tons of spawning biomass in 1988-89 to less than 30,000 tons by 1994 and has not rebounded since.

These herring collapses have shown long recovery times even though fishing was drastically reduced or halted. Their occurrence begs questions regarding the expected recovery

times for herring and the intrinsic and extrinsic factors that may control differences in these times amongst herring. Intrinsic factors relating to species biology, such as the age-at-maturity, growth, and body size, are related to generation times (Bjørkvoll et al., 2012; Hsieh et al., 2006; Inchausti & Halley, 2003; Pinsky & Byler, 2015). Extrinsic factors include bottom-up influences on recruitment and growth originating from the physical environment (Brunel & Dickey-Collas, 2010; Hay, Rose, Schweigert, & Megrey, 2008; Ito et al., 2015; Williams & Quinn, 2000). Strong top-down influences from predators have also been suggested as important for herring (Moran, Heintz, Straley, & Vollenweider, 2018a; Read & Brownstein, 2003; Schweigert et al., 2010; Surma & Pitcher, 2015; Tjelmeland & Lindstrøm, 2005). Finally, just like questions about whether chicken or the egg came first, there has been a long debate about whether low recruitment is a result of low spawning biomass, or low spawning biomass is a result of periods of low recruitment for herring (Gilbert, 1997; Myers & Barrowman, 1996), with more recent evidence backing recruitment-driven biomass in forage fish (Szuwalski et al., 2019; Szuwalski, Vert-Pre, Punt, Branch, & Hilborn, 2015). The nature, and more specifically direction of this relationship will have significant implications for expected recovery times.

Here we investigate four key questions about herring collapse and recovery by adopting a comparative, cross-population perspective on herring dynamics. Across herring populations we ask: 1) To what extent do herring biomass, recruitment, and catch dynamics fluctuate, and how does this compare to other fish species? 2) How often and for how long should we expect low biomass and recruitment to occur for herring? 3) What factors are most important in explaining the frequencies of low biomass and high recruitment? 4) What factors are most important in explaining the durations of periods of low biomass and recruitment? This study is based on the

largest compilation of information on Atlantic and Pacific herring (n = 64 populations), which enables us to answer these questions.

1.3 METHODS & MATERIALS

We collected historical records of catches and time series of estimated adult biomass and recruitment for herring populations worldwide and defined collapse and recovery based on trends in biomass. We used generalized linear models to predict the expected number of 'collapsed' years from covariates representing stock-specific biology and population dynamics, environmental conditions, predator trends, and fishing histories; and Random Survival Forests (Ishwaran, Kogalur, Blackstone, & Lauer, 2008) to explore whether these same covariates may explain differences in time-to-recovery.

1.3.1 Data sources and types

We obtained spawning biomass, recruitment, and catch time series from government agencies, public databases, and the published literature (Table A1; Trochta & Branch, 2018). The collated data set included 54 spawning biomass, 46 recruitment, and 64 catch time series (Fig. 1.1). Data were not available for all fished herring stocks now and in the past (e.g. Barents and White Sea populations; Hay et al., 2001), although the data in this study still have comprehensive global coverage. Furthermore, not all herring stocks have formal stock assessments, and as a result some biomass time series are raw population survey estimates (n=27) and others are outputs from stock assessments (n=27). Since variability characteristics, specifically the spectral frequencies and autocorrelation structure, of survey time-series differed substantially from those for stock assessment outputs (see Fig. A1), we applied a 3-year moving average to survey data so that survey

and stock assessment information show more similar frequencies and autocorrelation for analysis (see Fig. A2). Eight time series had at least one missing year that we interpolated with the moving average. Nine time series missing more than two consecutive years were not interpolated. Assessment estimates were also derived from different modeling approaches (catch-at-age analyses, state-space models, or virtual population analyses), although there were no noticeable differences in variability of estimates amongst approaches. No catch data are available for Squaxin Pass. For all stocks and types of data we calculated the minimum of each time series and the largest interannual changes across herring stocks.

1.3.2 Identifying and characterizing collapse

Population collapses are generally recognized as substantial declines in abundance from some baseline. Fish population collapses are often defined in reference to the biomass at maximum sustainable yield (B_{MSY}), which provides a defensible theoretical basis and management relevance (Neubauer et al., 2013; Pinsky et al., 2011). However, estimates of absolute abundance are unreliable (Hilborn, 2002), and when combined with the frequency of large-scale fluctuations and regime shifts in forage fish, make MSY and B_{MSY} difficult to estimate (McClatchie et al., 2017). Instead we focus on relative trends from surveys and stock assessments since these can be more accurately estimated and still provide information on the potential extent and duration of low abundance, assuming consistency in survey and assessment methodology within a time series.

To standardize relative biomass trends we divided each time series by the mean values falling within the upper 90th percentile of data in that time series. This is preferable to a time series maximum or mean since those are sensitive to the value and frequency of outliers of both low and high abundance. The threshold for low abundance, or collapse, is defined to be 30% of

the mean of those observations within the upper 90th percentile of each time series (abundances that are not necessarily contiguous):

$$\widehat{\text{biomass}}_{\text{i}} = \frac{\text{biomass}_{\text{i}}}{\text{mean}(\text{biomass}_{\text{90th}})}$$

where $biomass_i$ is the original estimate for year *i*, and $\widehat{biomass}_i$ is the normalized estimate for year *i*. The mean of the highest abundances is hereinafter referred to as Mean High Biomass. This same normalization calculation is applied to the catch and recruitment time series used for analysis, where the observed catches are divided by Mean High Catch to produce relative catch and the observed recruitment values are divided by Mean High Recruitment to produce relative recruitment. These Mean High values were based on the full time series, even for analyses that focused on standardized values for only the most recent 30 years.

We examined both the frequency and duration of collapse periods. Since time series among herring populations vary widely in coverage, we used only the last 30 years of each time series to determine the frequency or number of years below Mean High Biomass (calculated over entire time series). For most herring fisheries, the 30 most recent years follow peak industrial fishing and the implementation of harvest control rules based on stock assessments (Hay et al., 2001). Biomass collapse duration is defined as the number of consecutive years in which biomass is below 30% of Mean High Biomass. Since these durations can be censored if the period of low biomass includes the starting or ending year of the time series, we use a Kaplan-Meier analysis (Kaplan & Meier, 1958) to produce a time-to-recovery curve (or the time to the end of observations, whichever comes first). In this analysis, recovery is when biomass exceeds the collapse threshold. The calculated recovery probability at each observed time interval is the cumulative proportion of stocks collapsed beyond the preceding intervals on the time-to-recovery curve. Only stocks with two or more years of low biomass or low recruitment are included in the analysis to limit the effect of high-frequency

variability (i.e. measurement error). Collapse durations are also identified within the 30 most recent observations from each time series.

Recruitment dynamics may closely couple with population collapse and recovery. Consequently, low biomass provides an incomplete characterization of collapse because recruitment failure can underlie prolonged collapse while strong recruitment promotes recovery; low biomass is thus only a symptom. We use a threshold of 50% of Mean High Recruitment with which to determine the frequency of years above (moderate-to-strong recruitment) and the maximum duration below (recruitment failure) this threshold. A Kaplan-Meier analysis is also applied to low recruitment durations to calculate recruitment times-to-recovery. This Kaplan-Meier analysis is identical to the one outlined in the preceding paragraph.

1.3.3 Statistically modeling collapse predictors

We used negative-binomial linear mixed-effects (NBLME) models, either zero-inflated or not depending on the data (see Appendix A for more details). Predictor variables are listed in the next section. Random intercepts in the NBLMEs that reflect regional groupings based on management definitions (groupings are color coded in Fig. 1.1) did not change the results, so we only present results from the NBLMEs without random effects models for simplicity. We test the significance of these NBLMEs using a parametric bootstrapping procedure that simulates zero-inflated counts to which an intercept-only model and each predictor model are fit (see Appendix A). All models are coded in R 3.3.2 (R Core Team, 2016) using the glmmTMB package (Brooks et al., 2017). These models were applied in the following ways:

(1) To predict the number of years that biomass was below the collapsed threshold (30% of Mean High Biomass in the base case). The model was zero-inflated model.

(2) To predict the number of years of high recruitment (above 50% of Mean High Recruitment). The model was not zero-inflated.

We used Random Survival Forests (Ishwaran et al., 2008) to assess which factors best predict the number of years to recovery from low biomass or low recruitment (see Appendix A). Random Survival Forests ranked the importance of various effects and are an adaptation of random forest analysis for survival or event-time data. Variable importance (VIMP) was assessed by calculating the out-of-bag prediction error, which in this context was measured using Harrell's concordance index (Harrell, Califf, Pryor, Lee, & Rosati, 1982). Partial plots of predictor effects were generated by inputting predictor values into the fitted Random Survival Forest to obtain ensemble estimates (i.e. the average of 20,000 regression trees) of the expected numbers of biomass recoveries or high recruitment events, which is analogous to a cumulative hazard function in survival analysis (Ishwaran et al., 2008). Models were implemented using the *randomSurvivalforest* package in R (Ishwaran & Kogalur, 2016). These models were applied to the 30 most recent years for each population and address the following problems:

(1) Predicting the number of years to recovery from low biomass (falling below 30% of Mean High Biomass).

(2) Predicting the number of years to recovery from low recruitment (falling below 50% of Mean High Recruitment).

1.3.4 Predictor descriptions

We developed and explored a suite of hypotheses for intrinsic and extrinsic factors that might influence the duration of low herring biomass or recruitment as measured by the number of collapsed years and time to recovery. Specifically, we evaluated effects from both bottom-up (oceanographic conditions) and top-down (trends in predator abundance), fishing, and stock-

specific traits related to life history and population dynamics. The following factors (bold) were used to represent these effects (in no particular order).

Latitude (°N) of each herring stock's spawning location (Table A2) is included because it has previously explained differences in key population dynamics processes such as spawn timing, maturity, and growth of eastern Pacific herring (Hay, 1985; Hay et al., 2008). We assume latitude is a proxy for the climatic gradient in the northern hemisphere.

Freshwater inputs characterize the physical processes and habitat quality of estuaries in which herring spawn and their progeny survive and grow to maturity (Fortier & Gagné, 1990; Hay & McCarter, 1997). **Mean freshwater flux (km^3/yr)** come from the Global Runoff Data Centre's (GRDC) global hydrological model *WaterGAP* (Doll, Kaspar, & Lehner, 2003), based on river discharge measurements from the global network of GRDC stations (GRDC, 2014). The provided decadal means from 1961 through 2009 were binned within 5° degree latitudinal zones along coastlines. We identified the coastal zone in which each herring stock's spawning grounds are located and for this predictor used the mean of the decadal means encompassed by each stock's time series.

SST and SSHA are both attributes of physical ocean processes (e.g. basin circulation, eddies, currents, hydrographic discontinuities, ocean heat content, stratification, etc.) working over very different spatial and temporal scales, which leads to various hypotheses on how, when, and where they impact herring dynamics (Somarakis, Tsoukali, Giannoulaki, Schismenou, & Nikolioudakis, 2018). We constrained the list of environmental hypotheses by extracting SST and SSHA metrics at specific locations and times. For low biomass, this was the approximate timing and location of peak spawning activity; for low and high recruits, this was the period following

peak spawning at this same location (i.e. the egg-to-larval stages, or the "critical period"). SST and SSHA for these locations and times are mapped in Figs. S4-S5.

Sea surface height anomalies (SSHA) (cm) came from the JPL Physical Oceanography DAAC (Boulder, 2013; Hamlington, Leben, Strassburg, & Kim, 2014) and are based on satellite altimetry measurements and historical tide gauge data. The resulting data products are weekly imagery composites from June 1950 through 2010 with 0.5° degree spatial resolution. SSHA time series were created by first identifying the 0.5° x 0.5° cell nearest each stock's spawning location (Table A2), then taking the median of a 2° x 2° composite centered on the identified cell. The **Mean SSHA (cm)**, **SD of SSHA** (standard deviation), and **Linear trend in SSHA (cm/year)** over the herring time series were used in the analyses to determine the effect of average oceanographic conditions, in addition to variability and long-term trends in sea level (Tables S3-S4; Figs. S4-S5).

Sea surface temperature (SST) (°C) came from NOAA's Extended Reconstructed Sea Surface Temperature (ERSST v3b) dataset. The ERSST is a smoothed and filtered product derived from the International Comprehensive Ocean-Atmosphere Dataset with monthly averages at 2° spatial resolution (Smith, Reynolds, Peterson, & Lawrimore, 2008). The SST time series for each stock was extracted from the latitude and longitude coordinates of the 2° x 2° cell nearest the spawning location of each stock (Figs. S4-S5). The **Mean SST (°C)**, **SD of SST**, and **Linear trend in SST (°C /year)** over the range of years matching the 30 most recent years from each stock's time series were used to assess the influence of different averages in the thermal environment, interannual thermal variability, and long-term temperature trends respectively (Tables S3-S4).

Predator trends are included to determine their potential association with herring biomass collapse or recruitment failure. Predation information (e.g. consumption rates) is sparse although there are some data for key marine mammal and fish predators of herring. Marine mammal trend

data were obtained from Magera, Flemming, Kaschner, Christensen, and Lotze (2013). We focused on pinnipeds, ignoring cetaceans since cetacean numbers are estimated over large swathes of the ocean (e.g. Northeast Pacific humpbacks) with little information on how populations might feed on individual herring populations. Pinniped populations were paired with herring stocks based on their population area descriptions. Fish predator data were obtained from the RAM Legacy Stock Assessment Database (Ricard et al., 2011) for stocks identified to cohabit areas with herring stocks based on area descriptions. Trends in predator populations (slope and 95% confidence intervals) were estimated using robust linear regression ("lmRob" function) in the *robust* library in R (Wang et al., 2017) with code provided by Magera et al. (2013). The three predictors used were:

(1) **No. of increasing fish** or **pinniped populations**, the number of significantly increasing predator populations.

(2) **No. of decreasing fish** or **pinniped populations**, the number of significantly decreasing predator populations.

(3) **Mean of standardized population trends: Pinnipeds** or **Fishes**, the mean of the linear slope coefficients of predator time series identified as potential predator populations for a herring stock.

Peak fishing pressure is summarized from each stock's catch record in three ways to represent the effects of fishing since fishing mortality rates were not available for most herring stocks:

(1) **No. years relative catch > 0.75** is the number of years in the most recent 30 years that catch remained above 75% of Mean High Catch as a measure of the period of sustained high exploitation.

(2) **Mean of highest Catch/Biomass ratio** is the mean of the three largest ratios of relative catch to relative biomass in each year for the most recent 30 years. A larger mean ratio indicates that catch was high when biomass was low, which is indicative of unsustainable fishing.

(3) **Years Catch increased while Biomass decreased** is the number of years in the most recent 30 years where catches are maintained or increased from the previous year while biomass decreased from the previous year. A greater number of years again indicates a greater tendency to overfish, which may prolong collapse durations.

Zero catch (no. years) is also included for an effect of fishery closures on collapse times, which could associate with more (i.e. low biomass drives decisions to close fishing) or less (i.e. more closures promotes recovery) years of low biomass or recruitment.

First age at maturity is included to reflect differences in regeneration times that is hypothesized to impact the resilience of populations (Table A2).

Log(max catch) (metric tons) or the log of maximum catch is a measure of total population size, that has been shown to be strongly correlated with maximum sustainable yield (e.g. Srinivasan et al. 2010), since large populations are able to produce large catches. This is used in preference to estimates of absolute biomass since such estimates are highly uncertain, and are unavailable for most populations. The expectation is that larger populations may be more resilient than smaller populations (Table A2).

Mean age 5 weight (g) came from various literature sources (Hay et al., 2008; Hay et al., 2001; ICES, 2014; Ito et al., 2015; Naumenko, 2002; Stick, Lindquist, & Lowry, 2014) and is used to evaluate the potential effect of differences in body size on herring vulnerability. Differences reflect somatic growth rate and asymptotic weight. Age 5 is used to standardize mean weights for

comparison amongst stocks. This measure may impact population trends through variability in individual body condition, intrinsic population growth rates, and size selective predation (Table A2).

CV of biomass over the most recent 30 years is a measure of the variability of relative biomass (biomass divided by Mean High Biomass).

Median relative biomass over the most recent 30 years is included in the NBLME models and Random Survival Forests predicting recruitment years to evaluate the importance of the relationship between biomass and recruitment.

CV of recruitment measures the variability of relative recruitment (recruitment divided by Mean High Recruitment) over the most recent 30 years. The irregular pulses characteristic of recruitment time series are likely a measure of environmental variability.

Median relative recruitment is used to evaluate how well average recruitment associates with low biomass year counts and durations.

CV of R/SSB is the time series of recruits per spawning biomass (R/SSB) and is obtained by taking the ratio of recruits to spawners, standardized by dividing by Mean High R/SSB, and reflects changes in each population's ability to produce new recruits. The CV is obtained after standardization since each stock's units on spawning biomass and recruitment differ from one another. The interpretation of the resulting relative R/SSB metrics is only useful when compared among stocks.

Median R/SSB is defined as the median R/SSB (after standardization as described above) is used to find whether average productivity is well below the peaks or close to the peaks for each population.

1.4 RESULTS

1.4.1 To what extent do herring biomass, recruitment, and catch dynamics fluctuate?

Our database reveals highly varied biomass dynamics with dramatic changes in abundance for most stocks (Fig. 1.1). The average CV of relative biomass for 53 herring stocks with time series longer than 10 years, was 0.58 (95% confidence interval, CI, 0.30-0.93), compared to a median CV of 0.44 (95% CI 0.11-1.10) for the 307 non-forage fish species in the RAM Legacy database (Table 1.1; Fig. A3). The average minimum relative biomass was 0.097 of Mean High Biomass (95% CI 0.001-0.315) compared to a median of 22.2% (95% CI 1.1-67.6%). Out of the 53 herring populations, 28% (15) fell below 5% of Mean High Biomass, and 96% (51) fell below our collapse threshold of 30% of Mean High Biomass, in at least one year. Of the 51 collapsed stocks, 7 were at their lowest level in the final year, and 39 had recovered to above the collapse threshold after their low point. Recovery time among these 39 stocks averaged 13 years but was highly variable (95% CI 4-38 years). Far fewer non-forage fish (n = 307) fell below 5% of Mean High Biomass (10% vs. 28%), or below the collapse threshold (66% vs. 96%), but the time to recovery was similar, with a median of 12 years (95% CI 1-48 years).

Extreme interannual fluctuations (i.e. first differences of time series) are larger in herring than non-forage fish species. The magnitudes of the largest declines (median -28%, 95% CI -13 to -60%) are similar to the largest increases (28%, 95% CI 6-61%) in 48 herring stocks with continuous time series of biomass (excluding Goodnews Bay, Humboldt Bay, Nelson Island, Nunivak Island, and Security Cove). These fluctuations are much greater than those in non-forage fish species, with biggest declines of -14% (95% CI -3 to -58%) and biggest increases of 12% (95% CI 0-64%) (Table 1.1). These extreme interannual changes typify the “boom and bust” nature of herring population dynamics in which seldom, short periods of extreme changes are followed

by longer periods of much smaller changes. On the whole, interannual changes for herring stocks averaged -0.1% (95% CI -3.2 to 4.1%) which was similar to those for non-forage fish species (-0.7%, 95% CI -3.5% to 3.4%).

More herring stocks display pulses of strong cohorts compared to other fish, although such pulses are uncommon amongst stocks. Assuming log-normally distributed recruitment, herring stocks have much higher recruitment CV (median 0.16, n = 45, 95% CI 0.04 to 0.64) than non-forage fish (median 0.03, n = 263, 95% CI 0.005 to 0.14). For the herring stocks, 39% (17 of 45) had single-year recruitment increases greater than Mean High Recruitment, compared to only 21% of non-forage fish species (55 of 263).

Directed fisheries on herring stocks are also more likely to experience closures (i.e. zero catches ignoring bycatch from other fisheries) than non-forage fish species: 48% (31 of 64) were closed in at least one year compared to 14% (67 of 490) non-forage fish stocks. Similar results are obtained when examining stocks with catches less than 5% of Mean High Catch (73% of herring stocks vs. 45% of non-forage fish stocks). Herring stocks are more than twice as likely to be closed in recent years (after 2005, 28%; 1995-2005, 13%), largely because of recent depressed levels of many Pacific herring stocks—indeed Pacific herring account for 24 of 25 of all herring stocks with catches less than 5% of Mean High Catch.

1.4.2 How often should we expect low biomass and high recruitment for herring?

For this analysis, we focused on the 30 stocks that have at least 30 years of biomass estimates, and used the last 30 years of each time series (summarized in Fig. 1.2). The median biomass of each stock is correlated with median recruitment (Spearman's ρ=0.69, p<0.001), but not with median relative productivity (R/SSB) (ρ=0.19). Instead, CV of recruitment is correlated

with relative productivity (ρ=0.61, p<0.001). Median catch is also correlated with median biomass, but less so (ρ=0.40, p=0.02), and the correlation is almost significant for Atlantic herring (ρ=0.45, p=0.06), but is significant for Pacific herring (ρ=0.65, p=0.02).

To address the frequency of collapse in biomass, we focused on the lower tails of this distribution (Fig. 1.2). Collapse frequency is defined as the total number of years that biomass is below 30% of Mean High Biomass for each stock. Comparing among all stocks, biomass is collapsed for eight years on average. Atlantic herring are collapsed in fewer years (median 6.5 years) than Pacific herring (median 8.5 years), and one quarter of stocks are collapsed for half the time of the most recent 30 years (Fig. 1.3). The herring stocks with the most collapsed years (more than 15 years) are those in the Irish Sea, Northwest Ireland, West of Scotland, Western Newfoundland fall and spring spawners, Haida Gwaii, Kamishak, and Prince William Sound.

For recruitment, the analysis focused on the frequency of moderate to strong recruitment events: those above 50% of Mean High Recruitment (Fig. 1.3). Recruitment was above this threshold for a median of 8 years out of the most recent 30 years and differed little between Atlantic and Pacific herring. Stocks with few strong cohorts (3 or fewer years out of 30) are from West of Scotland, St. Mary's & Placentia Bays, West Newfoundland spring spawners, the Scotian Shelf and Bay of Fundy, Southern Gulf of St. Lawrence, Togiak, Kamishak, and Prince William Sound. Counter-intuitively, only four stocks with few strong cohorts also had many years of low biomass, namely West of Scotland, Kamishak, and Prince William Sound. In these stocks, strong year-classes occurred for a relatively short period before entering a protracted period of low recruitment with low variability.

1.4.3 For how long should we expect low biomass and recruitment to persist among herring?

Biomass collapse duration, or the number of years to recovery, is the consecutive number of years (minimum two years) below 30% of Mean High Biomass. Such collapses occurred in 23 of the 30 stocks (Fig. 1.3). The median collapse duration is 11 years and does not differ between Pacific and Atlantic herring. Only six stocks were collapsed for more than 15 years.

Prolonged recruitment failures are defined as the consecutive number of years below 50% of Mean High Recruitment (n=30 stocks). The median duration is 10 years, with shorter periods for Pacific herring (7.5 years) than Atlantic herring (11.5 years). For 14 of the stocks, long periods of recruitment failure (some up to 30 years) had not ended by the most recent year of data. This highlights large uncertainty in the outlook of herring recruitment dynamics because the absence of strong cohorts can span time scales of decades.

1.4.4 What factors are most important in explaining the frequencies of low biomass and high recruitment?

The variability seen in biomass collapse frequencies is the result of a complex suite of environmental, fishing, and biological factors (Figs. 4-5). Since many factors are tested which increases the risk of Type I error, we applied the Holm–Bonferroni method (Holm, 1979) to adjust the significance associated with our bootstrapped *p*-values. Greater frequencies of low biomass were most associated (based on NBLMEs) with lower median recruitment (Holm–Bonferroni $p<0.01$), and secondarily with lower SST standard deviation (Holm–Bonferroni $p<0.1$). We found no other significant associations between low biomass and all other factors (Table 1.2). Stock groupings (as color coded in Fig. 1.1) as a random effect did not impact estimates.

The frequency of high recruitment years was also analyzed using NBLMEs (Table 1.3), which found that across stocks, the number of years with high recruitment is positively associated with median biomass (Holm–Bonferroni $p<0.01$), and negatively associated with biomass CV (Holm–Bonferroni $p<0.01$), and the mean of the highest catch-to-biomass ratio (Holm–Bonferroni $p<0.05$).

These results were further checked with a suite of sensitivity tests to determine if any outlier values had a significant impact on the results (Appendix A), which marginally changed significance of some of the minor predictors, but had no major influence on the key predictors.

1.4.5 What factors are most important in explaining the durations of low biomass and recruitment?

For 23 stocks with low biomass (the other 7 of 30 stocks did not have low biomass persisting longer than two years), the most important predictors of the duration of low biomass were identified using variable importance (VIMP) in Random Survival Forests analysis. These predictors were, from most to least important, SD of SSHA, trend in SSHA, CV of recruitment, median relative recruitment, and trend in SST (Fig. 1.8). Other variables were unimportant (within or near the shaded region in Fig. 1.8). The out-of-bag error rate using Harrell's concordance index (Fig. A6) was 0.32, where values less than 0.5 indicate better predictive accuracy.

The generated Random Survival Forests were then used to determine the partial dependencies, i.e. how each predictor affected the probability that biomass collapse would last a given number of years. In Fig. A8, the probabilities that a stock would be collapsed for five or ten years are shown as a function of each predictor, demonstrating that collapse probability declines

with increasing SD of SSHA, increases with CV of recruitment, decreases with median relative recruitment, and has nonlinear dependencies with the other two predictors.

For 30 durations of low recruitment (i.e. time to between high recruitment events), the most important predictors from the Random Survival Forests were median relative biomass, SD of SSHA, the highest relative catch-to-relative biomass ratio, and mean predatory fish trend (Fig. 1.9). The error rate was 0.23. The results were little changed when two outliers were removed.

Partial dependency plots revealed that higher proportions of stocks with low recruitment is associated with lower median relative biomass, lower SD of SSHA, higher catch-to-biomass ratio, and lower rates of decline in predatory fish abundance (Fig. A9). However, low recruitment only strongly depends on these predictors at the lower end of predictor ranges, suggesting that predictor importance relies on a small subset of observations (Fig. A9).

Sensitivities in the Random Survival Forest analysis were also checked, including removing collinear predictors and re-running on times derived from different collapse thresholds (Appendix A). While predictive accuracy of the analyses changed, ranking of the key predictors did not which largely upholds the key findings presented here.

1.5 DISCUSSION

Our study characterizes population collapse and subsequent recovery of herring stocks, finding that herring have more extreme interannual swings in biomass and recruitment than non-forage fish species. For herring stocks, the average duration at low biomass is 11 years, moderate to high recruitment occurs approximately one-fourth of the time (8 of 30 years) on average, and consistently low recruitment spans a decade on average. Our investigation of explanatory factors

highlight how spawning biomass and recruitment relate to physical environmental conditions, fishing, region, and predators (both for Atlantic herring and Pacific herring).

1.5.1 Herring experience more population variability than non-forage taxa

Compared to non-forage fish species in the RAM Legacy database, herring biomass drops to lower minima, recovers to higher maxima, and exhibits larger maximum interannual increases and decreases. These observations are consistent with previous meta-analyses examining how clupeids decline and subsequently increase, which attributed their resilience to their short-lived, fast-growing life histories (Hutchings, 2001; Hutchings & Reynolds, 2004). However, these life history traits were also shown to increase clupeid vulnerability to collapse (Pinsky et al., 2011), and to exacerbate their vulnerability to collapse when overfishing occurs (Essington et al., 2015; Pinsky & Byler, 2015).

Herring are more likely to have very large recruitment events than non-forage fish and overall to display greater variability in recruitment. Strong cohorts (> one unit of relative recruitment) occur in 3.1% of years within Pacific and Atlantic herring, and in survey and stock-assessment-derived estimates. Of the 17 herring stocks with the largest cohorts, three pairs within the same region share cohorts in the same years (Prince William Sound and Sitka in 1987 and 1991, Central Coast and Haida Gwaii in 1979, and West Newfoundland fall and spring spawners in both 1982 and 1983). The 1987 and 1991 year classes in Prince William Sound and Sitka also match somewhat smaller recruitment pulses in other central Gulf of Alaska stocks (Kamishak and Kodiak, Fig. 1.1), but it is possible these pairs may just be coincidences since correlations reported between these stocks in the 1980s ceased to exist thereafter (Rice & Carls, 2007). More commonly,

nearby stocks had unmatched large year-classes, implying that stock-specific local conditions are important in determining the magnitude of cohorts.

1.5.2 Fisheries closed for Pacific herring, but not Atlantic herring

Nearly half of herring stocks experienced very low to no catches during the time frame of analysis, which is a greater proportion than observed in non-forage fish stocks; and half of these herring fisheries remained closed for more than ten years. Many herring fisheries were closed recently (from the mid-2000s on), and closures were all in the Pacific. Since both Pacific and Atlantic herring stocks showed similar patterns of declines in catches and biomass, that more fishery closures occur among Pacific herring likely reflect a number of differences in fishing dynamics and their management. For example, nearly all Pacific herring stocks in our analyses were fished by a single country (Canada or U.S.), while all Atlantic herring stocks in the Northeast Atlantic are targeted by multiple countries. A previous study has shown that when more countries jointly fish a stock, this increases the risk of overexploiting the stock (McWhinnie, 2009). However, Northwest Atlantic herring too have shown no fishery closures (along with high frequencies and durations of low biomass) even though they are not shared between countries (fished and managed by either Canada or U.S.).

Differences also exist in the types of fishers targeting Pacific and Atlantic herring stocks. For example, Northeast Atlantic stocks are dominated by commercial fisheries, while many Pacific herring stocks (mostly in the Northeast) are managed by dividing a total catch among a variety of fishers including large industrial fleets, commercial fishers, recreational fishers, subsistence fishers, and indigenous fishers such as the coastal First Nations of British Columbia. Conflicting goals among user groups resulted in stark disagreement between industry and First Nations on the

Department of Fisheries and Oceans (DFO) management policy for Haida Gwaii fishing (Lam et al., 2019), following a court injunction to reverse DFO's decision to open commercial fishing on Haida Gwaii and West Vancouver Island herring (Jones et al., 2017). In other words, cultural considerations (whether preemptive or by intervention) may promote more fishery closures amongst Pacific herring stocks, perhaps in addition to other differences in fishing and/or management (e.g. different gear types, management definitions and units of a stock, management areas, seasons, etc.).

1.5.3 Half of herring stocks collapsed for a decade or more

Biomass in all herring stocks declined below 30% of Mean High Biomass at some point, with an average collapse duration of 11 years. These time scales are similar to those discussed in other studies on the recovery of fished populations after collapse (Neubauer et al., 2013; Petitgas et al., 2010). Paleo records of anchovy and sardine abundance off the California coast indicate an average time of 1-2 decades to return to "fishable biomass" (33% of mean peak biomass) (McClatchie et al. (2017). Recovery is not inevitable: only 4 of 12 stocks that had remained collapsed for 11 or more years, had recovered above the 30% threshold by the end of the time series. Thus recovery may involve far longer durations than could be recorded in our data set. This uncertainty is worrisome given that prolonged collapse times also may include severely contracted geographical ranges and the loss of spawning components (Hay et al., 2001; Melvin & Stephenson, 2006). These biological effects can have severe long-term consequences for fishing fleets (Dickey-Collas et al., 2010), and cultures whose long-standing traditions rely on herring such as roe-on-kelp harvests (Gauvreau et al., 2017; Jones et al., 2017; Menzies, 2016; Thornton et al., 2010).

Herring stocks with long periods of recruitment failure and few strong year-classes are notable for a "flatline" in their recruitment time series. We found that protracted periods of herring recruitment failure correlate with higher autocorrelation (ρ=0.60, p<0.001). High autocorrelation in recruitment may be driven by spawner abundance (Somarakis et al., 2018), low frequency ecological drivers (Pepin, 2015; Pyper & Peterman, 1998; Ricard, Zimmermann, & Heino, 2016), or the combined effects of both (Punt, Szuwalski, & Stockhausen, 2014; Szuwalski et al., 2015; Vert-pre, Amoroso, Jensen, & Hilborn, 2013).

1.5.4 Recovery hinges on the link between biomass and recruitment

Distributions of spawning biomass, recruitment, catch, and productivity (defined as the relative recruits/spawning biomass, or R/SSB) across stocks reveals no relationship between median spawning biomass and median productivity. Furthermore, the number of low biomass years is best predicted by lower median recruitment and higher recruitment variability, and the duration of low recruitment by low median biomass, while the frequency of high recruitment is best predicted by high median biomass and high biomass CV.

Both high and low stock-recruitment associations have been previously shown for herring (Myers & Barrowman, 1996) and there is recent evidence for statistically significant stock-recruitment relationships in forage fish (Somarakis et al., 2018). Still, other recent studies have shown that cross correlations are strongest when recruitment leads spawning biomass across stocks (Szuwalski et al., 2019; Szuwalski et al., 2015) with arguments positing the overwhelming effects of environment and/or life history compared to spawner abundance (sensu Pepin, 2015). Our results do not robustly test one claim over the other; instead, our results suggest a strong association

between recruitment and biomass at high and low levels as the primary determinant of collapse times and eventual recovery across herring stocks.

1.5.5 Catch patterns and recovery

While catches generally track spawning biomass, catches alone are less useful than time series of fishing mortality. However, fishing mortality values are not available for many stocks. Therefore, we developed three proxies for the duration and magnitude of unsustainable exploitation (the number of years in which relative catch exceeded 0.75; years in which catch increased while biomass decreased; and mean of the highest catch/biomass). The linear mixed models and Random Survival Forest both found little relationship between times at low biomass and our exploitation proxies.

In contrast, results the linear mixed models and Random Survival Forests found that stocks with larger maximum catch-to-biomass ratios in their record were likely to have fewer years of strong recruitment and longer recruitment failure. This connection could reflect that: 1) recruitment failure precludes biomass recovery, which is exacerbated by increased exploitation, or that 2) this is the result of recruitment overfishing, as has been previously noted for herring stocks (e.g. Cushing, 1971; Dickey-Collas et al., 2010; Hay et al., 2001). Fishing also mainly continues on Atlantic herring, but usually ceases on Pacific herring at the low relative levels we defined. These continuing catches of Atlantic herring imply a risk of increasing fishing mortality and overfishing especially given the density-dependent catchability that is characteristic of forage fish (Pitcher, 1995). Fishery closures that are more common among Pacific herring may negate this risk; however, closing does not seem to guarantee a speedy recovery since long durations of low biomass still exist for several stocks despite being closed to fishing.

1.5.6 The importance of oceanographic variability

Variability both in sea surface temperature (SST) and sea surface height anomaly (SSHA) were key predictors of low herring biomass. Greater environmental variability in SSHA and SST were associated with fewer years of low biomass. This makes sense given that higher spawning biomass is driven by occasional large recruitment events, which is in turn driven by recruitment variability. Variability in SSHA is correlated with periods of failed recruitment, and likely also adult mortality and individual growth, given the strong negative correlation between weight-at-age 5 and SSHA variability. While the link between environmental and recruitment variability is implicit in hypotheses postulated by other authors (e.g. the "optimal stability window" hypothesis from Gargett, 1997), the correlation between body condition and variability has not been made before in the literature and warrants further investigation into the mechanisms controlling somatic growth and asymptotic size.

Unlike SSHA variability, we found no relationship between SST variability and recruitment. Interannual temperature variability may indirectly force greater variability in growth and survival through changes in prey availability and distribution (e.g. Corten, 2001; Cushing, 1990; Southward, Hawkins, & Burrows, 1995), or directly force frequent or larger changes in growth and survival (Boltaña et al., 2017) via cumulative changes in individual metabolism and physiology (Pörtner & Knust, 2007; Pörtner & Peck, 2010). In other words, more variable SST may amplify variability in spawning biomass separately from the effect of recruitment variability. However, Pinsky and Byler (2015) found evidence that fast growing forage fish in more variable thermal environments are more vulnerable to collapse as a result of overfishing. Higher variability in SSHA was associated generally with shortened periods of failed recruitment in our analysis.

Recent work has implied strong links between recruitment success of various fish stocks and ocean currents and gyre-related circulation patterns, which directly influence SSHA, in the North-east Atlantic (Zimmermann, Claireaux, & Enberg, 2019). Other studies have more specifically linked SSHA to recruitment of other species in the Gulf of Alaska and California Current (Stachura et al., 2014), the northern Benguela (Hardman-Mountford, Richardson, Boyer, Kreiner, & Boyer, 2003), and in the Leeuwin Current (Caputi, Fletcher, Pearce, & Chubb, 1996; Pearce & Phillips, 1988). Coastal upwelling, which also influences SSHA variability, is also strongly correlated with herring recruitment variability (Reum, Essington, Greene, Rice, & Fresh, 2011).

1.5.7 Other factors

We found no significant effects of predator trends on the number of years of low biomass or high recruitment, but larger declines in Pacific fish predators were associated with shorter periods of low recruitment. Inference on this potential relationship is complicated by the spatial context: juvenile herring are generally found within coastally enclosed areas (e.g. sounds, bays, and straits), while our data on fish predators spans larger regions of open ocean over the eastern Pacific shelf. This spatial (and likely temporal) mismatch motivated our exclusion of other predators such cetaceans and seabirds from analysis, since these groups migrate over even vaster distances. Seasonal and local estimates of their numbers may be available (e.g. Bishop, Watson, Kuletz, & Morgan, 2015; Surma & Pitcher, 2015; Teerlink et al., 2015), but estimates that are relevant to analysis were not generally available. Nevertheless, predator-prey relationships of cetaceans and seabirds on herring have been implicated as top-down controls on herring (Moran et al., 2018a; Read & Brownstein, 2003; Straley et al., 2017) and as recipients of bottom-up effects from herring (Pikitch et al., 2012; Pikitch et al., 2014; Smith et al., 2011). Herring have also been

identified as a key prey item for pinniped populations throughout the Pacific and Atlantic, including Southeast Alaska (Gende & Sigler, 2006; Sigler, Gende, & Csepp, 2017; Womble & Sigler, 2006), British Columbia (Olesiuk, 1999, 2008), Puget Sound (Lance, Chang, Jeffries, Pearson, & Acevedo-Gutiérrez, 2012), the Southern Gulf of St. Lawrence (Hammill, Stenson, Proust, Carter, & McKinnon, 2007), the Gulf of Maine (Overholtz & Link, 2006; Read & Brownstein, 2003), and the North Sea (Sveegaard et al., 2012). A more detailed and thorough investigation of these relationships involving various herring predators is needed, as has been done by Surma et al. (2018b).

Factors that were not found to have any effects were latitude, freshwater influx, first age-at-maturity, age-5 weight, and maximum catch. Given the importance of fishing and oceanographic predictors from our analyses, extrinsic factors may determine herring recovery times to a greater degree than the intrinsic stock-specific characteristics considered here.

1.5.8 Challenges of approach

A variety of assumptions in our approach were explored in more detail in sensitivity analyses (Appendix A). Chief among these was the choice of a threshold to define low or high levels in biomass or recruitment. Naturally, levels below 30% of the Mean High Biomass may or may not equate to true population collapse in some stocks. Other fisheries meta-analyses have taken a similar approach of defining collapse with a reference point determined from the time series themselves to standardize comparisons (e.g. Essington et al., 2015; McClatchie et al., 2017; Mullon et al., 2005; Pinsky et al., 2011; Worm et al., 2009). Applying alternative thresholds did not substantially change our conclusions except for the Random Survival Forest model of low biomass durations. Due to the loss of predictive accuracy, thresholds that were lower (20%) or

higher (40%) modified the duration data in a way that made them less informative for our specific objective. For example, changing the threshold to 20% reduced the sample size from 23 to 16 since many stocks did not experience prolonged times below this threshold. Different thresholds in recruitment also did not alter our conclusions.

Testing of the various factors that may explain recovery times is also caveated by the scale with which we conducted this analysis. As with herring predators, ecosystem factors are particularly nuanced because of the time and spatial scales with which herring biological processes and oceanographic variables interact. For example, more localized oceanographic processes may influence herring biomass and/or recruitment than captured in our broad-scale factors so that even if such factors do impact some herring stocks, the cross-stock effect is absent or undetectable. This may explain why we do not find more significant predictors (for more discussion, see Appendix A). However, small stocks for which only survey estimates are available comprise over half of the biomass data, and they lack stock-specific ecological information needed to develop more appropriate time series for analysis. As a result, our use of broader ecological factors is the best possible given available data, with the caveat that we can only detect effects from the variables for which data are available.

1.6 CONCLUSIONS

The data presented brings together a wealth of herring knowledge in the most comprehensive compilation of herring population dynamic datasets to date, extending far beyond the data contained in the RAM Legacy database (Ricard et al., 2011). Stock assessment models do not exist for many herring populations in the Pacific, and Pacific herring have much greater representation in our data set. Few other fish taxa have as many records across stocks and years

and thus this analysis provides the largest available meta-analysis on a single marine fish group. This "treasure trove" of herring data was first glimpsed in Hay et al. (2001), which served as one of the foundations for this paper and its questions.

We found a wide variety in the extent and duration of collapse and time to recovery in herring, implying that any management strategies must be robust to this broad range of possibility to avoid the risk of losing sub-stocks and eroding the long-term resilience of stock complexes. Timely reductions in fishing effort may counteract initial rapid and drastic declines, allowing some leverage through robust fisheries management (Bakun & Broad, 2003; Essington et al., 2015). Achieving this timeliness remains an obstacle and requires an ongoing investigation of biology and ecosystem interactions for each stock, since these factors likely change over time. One approach is to more specifically identify smaller scale environmental indicators of impending productivity changes that would allow prompt reductions in harvest rates (Lindegren, Checkley, Rouyer, MacCall, & Stenseth, 2013). However, the success of this approach relies on the accuracy of these indicators in predicting changes (Punt et al., 2013). Novel methods have demonstrated the promise of accurate predictions from environmental indicators (Deyle et al., 2013), and applying such methods to datasets such as ours along with more context specific environmental variables would be useful.

Our analyses are correlative, highlighting links between factors and recovery times without offering further evidence for specific mechanisms (Pepin, 2015; Williams & Quinn, 2000). Distinguishing recovery probability based on biologically plausible factors is a useful assessment of the stocks most at risk for prolonged recovery when collapsed. Knowing which factors are important for distinguishing recoveries (e.g. "risk factors") amongst stocks can inform comparisons of management procedures amongst management areas. Identifying these risk factors

also directs research to diagnose their causal pathways to longer or shorter collapses and recruitment failures amongst stocks. Our analyses demonstrate the potential of these data and similar datasets to address important questions on fish population dynamics. The outcome of such analyses have the potential to inform fisheries management on the types of policies that are best suited to promote faster potential recovery times after population collapse.

1.7 ACKNOWLEDGEMENTS

Funding for this work was provided by the Exxon Valdez Oil Trustee Council (EVOSTC). T.A.B. was additionally supported in part by the Richard C. and Lois M. Worthington Endowed Professor in Fisheries Management. Various scientists and agencies shared herring data and/or stock-specific information (in no particular order): S. Dressel with the Alaska Department of Fish and Game; J. Cleary with the Pacific Biological Station, Fisheries and Oceans Canada; J. McDermid, C. Bourne, and T. Doniol-Valcroze with Fisheries and Oceans Canada in the Northwest Atlantic; J. Deroba with NOAA Northeast Fisheries Science Center; K. Stick with Washington Department of Fish and Wildlife; T. Greiner with California Department of Fish and Wildlife; V. Radchenko with the North Pacific Anadromous Fish Commission; G. Óskarsson with the Marine Research Institute in Iceland; H. Okamura with the National Research Institute of Fisheries Science in Japan; the International Council for the Exploration of the Sea (ICES) stock assessment database; and the RAM Legacy database. We thank A. Punt, T. Essington, and S. Dressel and three anonymous reviewers for invaluable comments on the manuscript.

1.8 DATA AVAILABILITY STATEMENT

Most of the data underlying this study are available at https://doi.org/10.24431/rw1k1i. Restrictions apply to the availability of the remaining data, which were used under privacy agreements for this study, and those data may be requested from the data providers identified in Table A1 of the Appendix A.

1.9 TABLES

Table 1.1. Characteristics of time series of herring stocks collected for this study and non-forage fish stocks from the RAM Legacy database. For values derived from distributions, the median is shown first followed by the 95% confidence intervals in parentheses. For entries indicating the number of stocks meeting a certain criterion (e.g. "No. stocks…"), this applies across all years available for a stock so that a single instance when a criterion is met qualifies that stock.

	Herring	**Non-forage fish**
Biomass		
No. series available	53	307
CV of relative biomass	0.58 (0.30-0.93)	0.44 (0.11-1.10)
Minimum relative biomass	0.097 (0.001-0.315)	0.222 (0.011-0.676)
No. stocks below 5% Mean High Biomass	15	30
No. stocks below 30% Mean High Biomass	51	202
No. stocks with max. biomass greater than 30% following the historical minimum	39	120
Median years to reach max biomass greater than 30% following the minimum	13 (4-38)	12 (1-48)
Median value of this maximum biomass following the minimum	1.00 (0.11-1.25)	0.500 (0.056-1.22)
Largest 1-year decrease in biomass	0.28 (0.13-0.60)	0.14 (0.03-0.58)
Largest 1-year increase in biomass	0.28 (0.06-0.61)	0.12 (0.004-0.64)
Median interannual change in biomass	-0.001 (-0.032-0.041)	-0.0065 (-0.035-0.034)
Recruitment		
No. series available	45	263
Median CV of log relative recruitment	0.16 (0.06-0.64)	0.03 (0.005-0.14)
No. stocks with at least one exceptionally large cohort (> Mean High Recruitment)	17	55
Proportion of years with exceptionally large cohorts	0.031 (0.017-0.057)	0.009 (0.026-0.05)
Median relative recruitment of stocks with exceptionally large cohorts	0.22 (0.05-0.43)	0.46 (0.08-0.90)
Catch		
No. series available	64	490
No. series where catch=0 at least once	31	67
No. series where catch <5% of Mean High Catch at least once	47	222
Median time when catch=0	9 (1-25)	--

Table 1.2. Estimated parameters for single-variable zero-inflated negative binomial GLMMs predicting **the number of low biomass years** among herring populations (*n*=30). The estimated intercepts (Int.), effect coefficients (Eff.), and the lower (L 95% CI) and upper 95% confidence interval bounds (U 95% CI) on the effect are provided for each model. All variables are scaled by their mean and standard deviation. The *p*-values are empirically derived, based on the proportion of parametrically bootstrapped likelihood ratios between the full and null models (only the means in the conditional and zero-inflated models with a random-intercept on herring locale) that are as extreme as the observed likelihood ratio. These bootstrapped likelihood ratios derive from converged model fits from the null model (# sims). Given the large number of predictors, we also apply the Holm-Bonferroni Method on the empirical p-values to correct for Type I errors (0.01 = ***, 0.05 = **, 0.1 = *).

		Conditional model			**Zero-inflated model**				
Explanatory factor	***n***	**Eff.**	**L 95% CI**	**U 95% CI**	**Eff.**	**L 95% CI**	**U 95% CI**	**# sims**	***p***
Median relative recruitment	30	-0.57	-0.87	-0.26	1.56	0.23	2.88	991	*0.000***
SD of SST	30	-0.45	-0.71	-0.18	0.83	-0.21	1.87	991	*0.004**
CV of recruitment	30	0.25	-0.01	0.51	-1.92	-3.98	0.14	997	0.012
CV of R/SSB	30	0.06	-0.22	0.34	-1.95	-3.83	-0.06	993	0.021
SD of SSHA	30	-0.48	-0.84	-0.12	0.13	-1.07	1.33	988	0.039
Peak fishing: no. years relative catch > 0.75	30	0.04	-0.33	0.41	1.73	-1.23	4.69	986	0.046
Zero catch (no. years)	30	0.19	-0.04	0.42	NA	NA	NA	895	0.051
Latitude (˚N)	30	0.17	-0.16	0.49	0.80	-0.41	2.01	990	0.265
Mean freshwater influx (cu.km/year)	30	0.24	-0.05	0.54	-0.16	-1.13	0.81	994	0.287

		Conditional model			Zero-inflated model				
Explanatory factor	***n***	**Eff.**	**L 95% CI**	**U 95% CI**	**Eff.**	**L 95% CI**	**U 95% CI**	**# sims**	***p***
First age at maturity (year)	30	0.20	-0.09	0.49	0.44	-0.45	1.33	1000	0.299
Mean age 5 weight (g)	30	0.32	-0.10	0.74	0.12	-1.10	1.35	992	0.362
Peak fishing: mean of highest Catch/Biomass ratio	30	0.15	-0.09	0.40	-0.91	-3.87	2.05	998	0.374
Mean SSHA (cm)	30	-0.11	-0.43	0.21	0.60	-0.35	1.55	989	0.392
Log(max catch) (metric tons)	30	0.07	-0.26	0.40	0.66	-0.57	1.89	991	0.480
No. of decreasing fish populations	30	0.16	-0.12	0.44	-0.24	-1.24	0.76	999	0.514
Mean of standardized population trends: Pinnipeds	30	-0.16	-0.45	0.14	0.28	-0.79	1.36	995	0.524
Linear trend in SSHA (cm/year)	30	-0.03	-0.39	0.33	0.56	-0.40	1.53	988	0.531
Linear trend in SST (˚C/year)	30	0.03	-0.32	0.39	0.68	-0.88	2.23	999	0.573
Median R/SSB	30	-0.11	-0.41	0.19	0.33	-0.65	1.32	997	0.618
No. of increasing fish populations	30	0.15	-0.17	0.48	0.19	-0.69	1.07	999	0.640
Peak fishing: years Catch increased while Biomass decreased	30	0.05	-0.19	0.30	-0.41	-1.42	0.61	992	0.671
Mean of standardized populations trends: Fishes	30	-0.03	-0.34	0.28	0.37	-0.69	1.42	986	0.784

		Conditional model			Zero-inflated model				
Explanatory factor	***n***	**Eff.**	**L 95% CI**	**U 95% CI**	**Eff.**	**L 95% CI**	**U 95% CI**	**# sims**	***p***
No. of increasing pinniped populations	30	0.09	-0.20	0.37	-0.06	-1.04	0.92	1000	0.860
Mean SST (˚C)	30	0.00	-0.27	0.28	-0.26	-1.27	0.74	996	0.894

Table 1.3. Estimated parameters for single-variable negative binomial GLMMs predicting the **number of high recruitment years** (*n*=30). The estimated intercepts (Int.), effect coefficients (Eff.), and the lower (L 95% CI) and upper 95% confidence interval bounds (U 95% CI) on the effect are provided for each model. All variables are scaled by their mean and standard deviation. Adjusted significance of the p-values using the Holm-Bonferroni Method are also indicated (0.01 = ***, 0.05 = **, 0.1 = *)

		Conditional model				
Explanatory factor	**n**	**Eff.**	**L 95% CI**	**U 95% CI**	**# sims**	**p**
Median relative biomass	30	0.48	0.29	0.67	1000	*0.000***
CV of biomass	30	-0.49	-0.73	-0.26	1000	*0.000***
Peak fishing: mean of highest Catch/Biomass ratio	30	-0.91	-1.56	-0.25	1000	*0.002***
Peak fishing: years Catch increased while Biomass decreased	30	-0.40	-0.66	-0.15	1000	0.005
Median R/SSB	30	0.35	0.11	0.58	1000	0.017
Mean of standardized population trends: Pinnipeds	30	0.31	0.05	0.57	1000	0.050
No. of decreasing pinniped populations	30	-0.32	-0.62	-0.03	1000	0.091
CV of R/SSB	30	-0.24	-0.48	-0.01	1000	0.109
SD of SSHA	30	0.21	-0.02	0.45	1000	0.186
Mean freshwater influx (cu.km/year)	30	-0.20	-0.45	0.06	1000	0.340
No. of increasing fish populations	30	-0.21	-0.48	0.06	1000	0.358
Mean of SSHA (cm)	30	0.15	-0.10	0.41	1000	0.456
Mean of standardized populations trends: Fishes	30	-0.13	-0.36	0.11	1000	0.492
No. of decreasing fish populations	30	-0.13	-0.40	0.14	1000	0.518
Log(max catch) (metric tons)	30	0.14	-0.15	0.43	1000	0.521
Zero catch (no. years)	30	-0.08	-0.35	0.19	1000	0.558

		Conditional model				
Explanatory factor	**n**	**Eff.**	**L 95% CI**	**U 95% CI**	**# sims**	**p**
Linear trend SSHA (cm/year)	30	0.07	-0.16	0.30	1000	0.565
Peak fishing: no. years relative catch > 0.75	30	0.06	-0.18	0.30	1000	0.579
SD of SST	30	0.05	-0.25	0.34	1000	0.586
Latitude (˚N)	30	-0.04	-0.28	0.20	1000	0.590
Age at recruitment (year)	30	-0.03	-0.30	0.24	1000	0.594
Mean SST (˚C)	30	0.04	-0.19	0.27	1000	0.595
Linear trend in SST (˚C/year)	30	-0.02	-0.24	0.21	1000	0.604
No. of increasing pinniped populations	30	-0.01	-0.30	0.27	1000	0.605

1.10 FIGURES

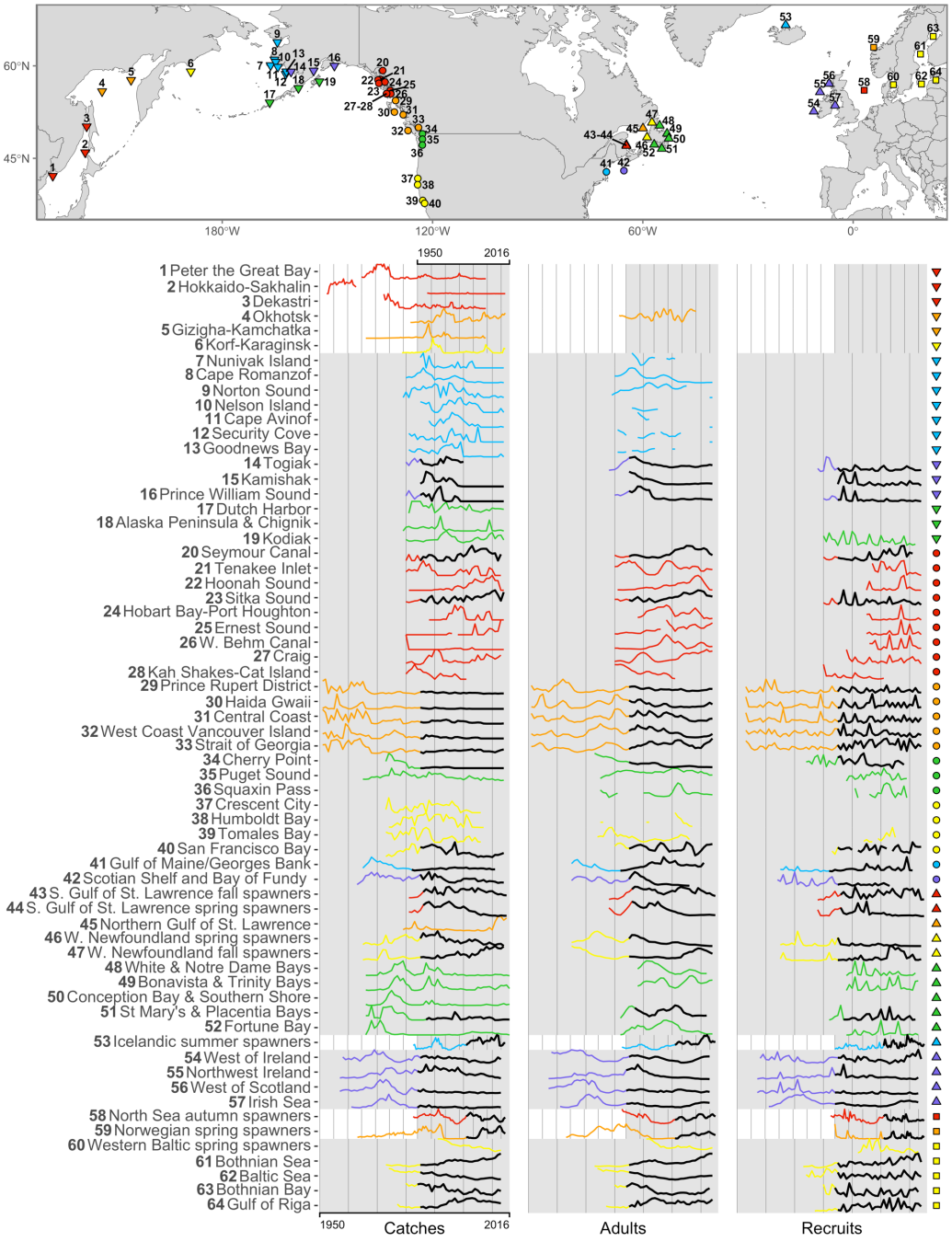

Figure 1.1. Herring time series and their locations from west (point 1) to east (point 64). Observations of each time series are divided by their respective historical maxima so they are visualized on the same scale. The filtered versions of the spawning biomass (Adults) series are shown for those stocks whose estimates came from surveys. The unique color-shape combinations indicate the management area or stock complex to which each time series belongs in both the map and time series. The shaded regions represent a reference time frame (1950-2016), since some records start earlier and are "zoomed out" above to show their entirety. The select herring stocks and years used in the analyses of this paper are emboldened in black.

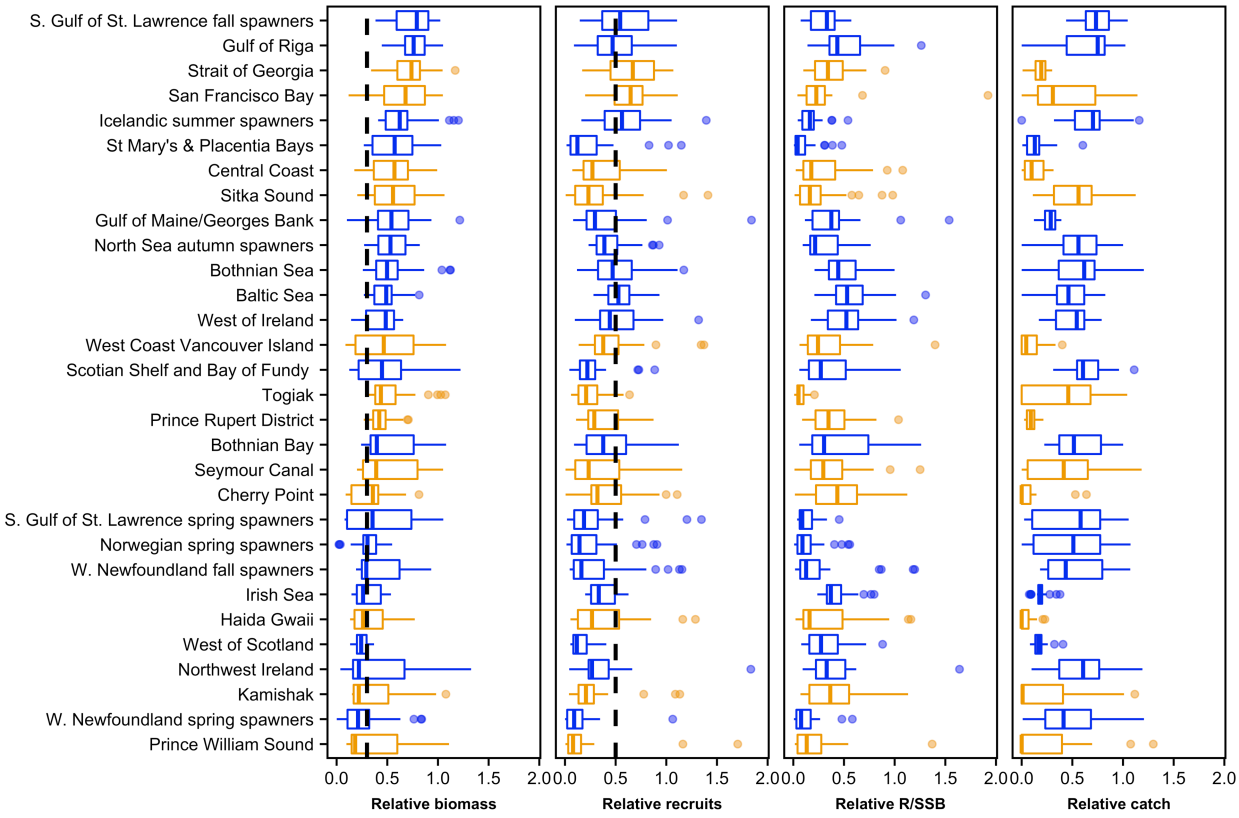

Figure 1.2. Box plots of time series distributions for stocks used in our analysis (at least 30 years of observations). Each time series is divided by its respective mean high value (i.e. mean of the values in the 90th percentile). Stocks are ordered by their median relative biomass from the largest median at the top to lowest median at the bottom and color coded by species (Atlantic herring are blue and Pacific herring are orange). The 30% of Mean High Biomass and 50% of Mean High Recruit thresholds used to identify collapse levels are also shown by dashed lines in the Relative biomass and Relative recruits plots.

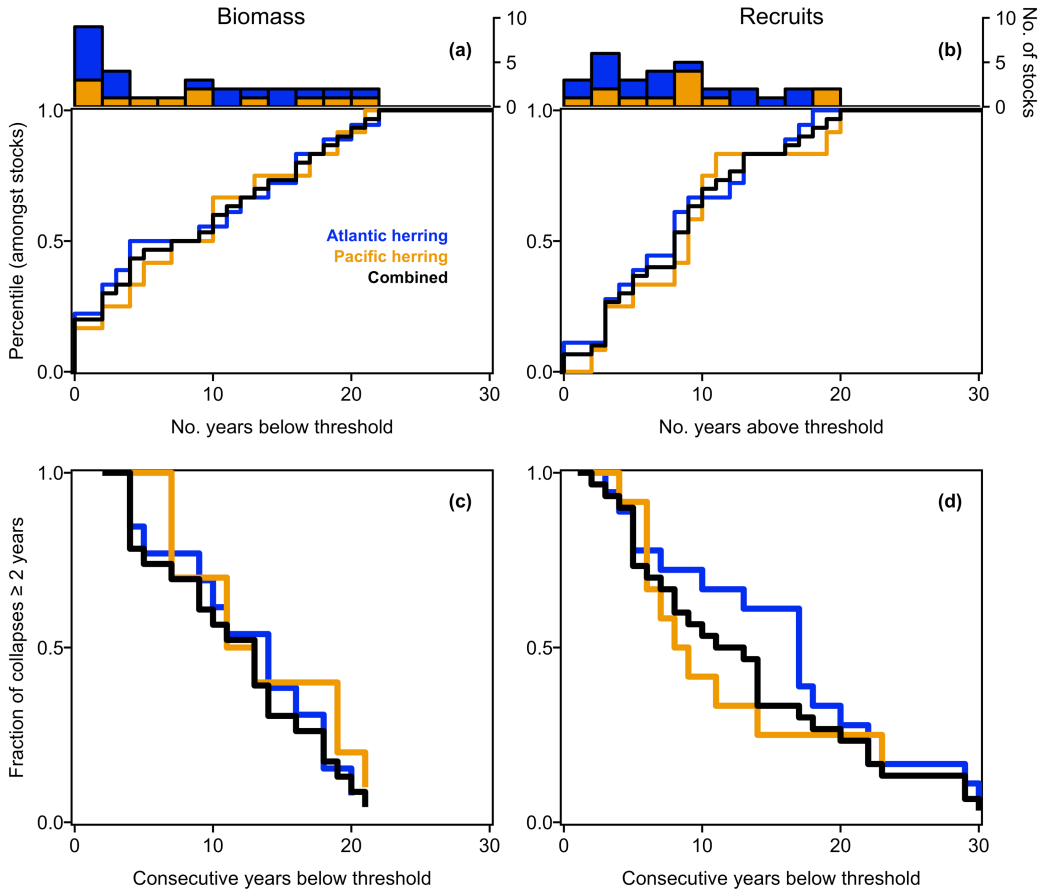

Figure 1.3. Distributions of the number of years below or above a threshold amongst herring stocks (N=30). Both the histogram (top) and empirical cumulative distribution (bottom) are shown for (a) the total number of years in which biomass falls below 30% of Mean High Biomass and (b) the total number of years recruits exceed 50% of Mean High Recruitment. Also shown are the frequencies of events with more than X consecutive years of (c) biomass below 30% of Mean High Biomass (N=23) and (d) recruits below 50% of Mean High Recruitment (N=30). The minimum period considered was more than two years. Since stocks may show multiple periods longer than two years below the stated thresholds, each stock's maximum period is considered. For all stocks, only the 30 most recent years of each stock's time series are used to determine the number of years on all plots.

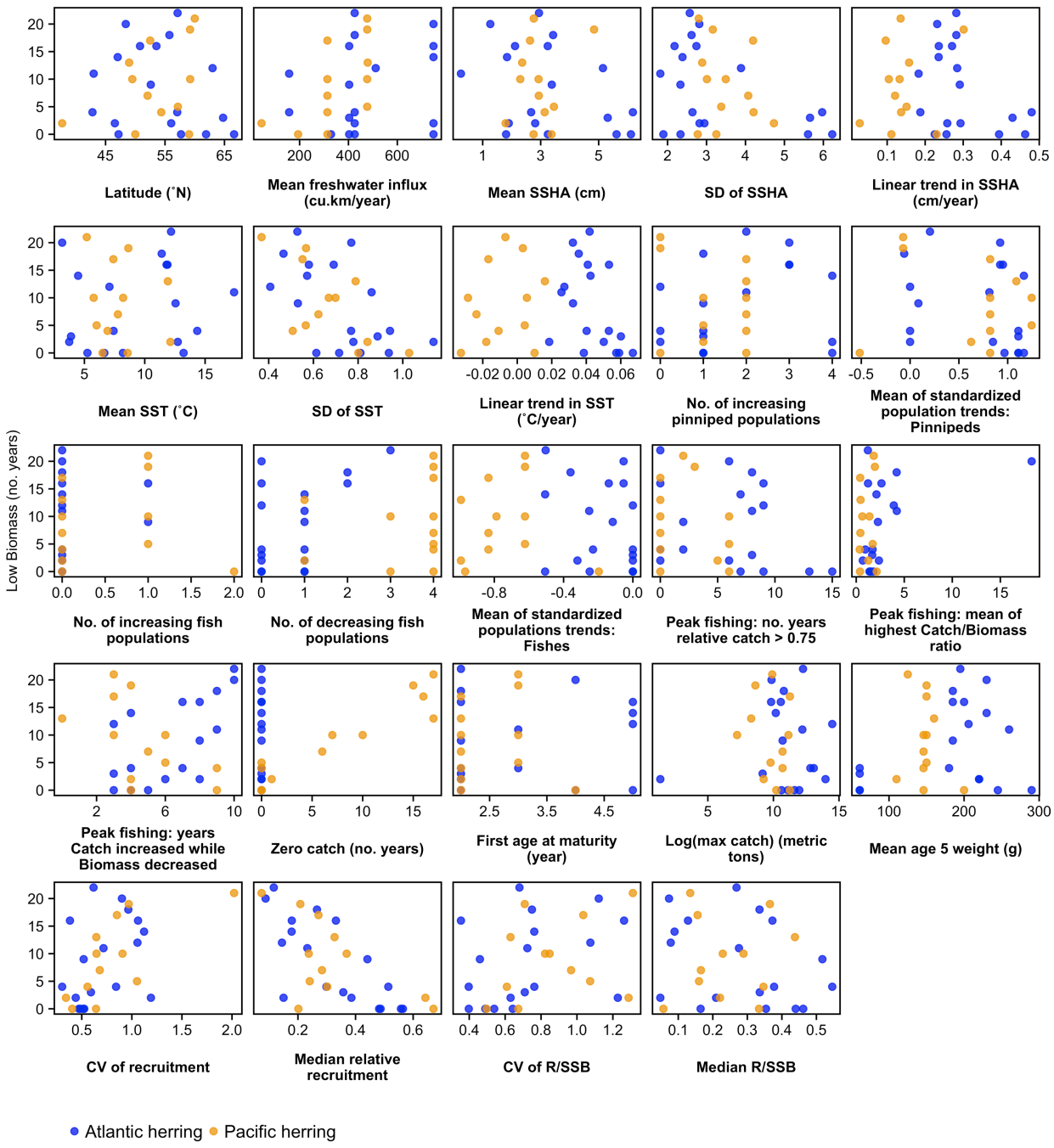

Figure 1.4. Observed low biomass year counts < 30% of Mean High Biomass with their respective covariate values for each of the variables explored with the zero-inflated generalized linear mixed effects models. Observations are colored by species (Atlantic herring=blue, Pacific herring=orange).

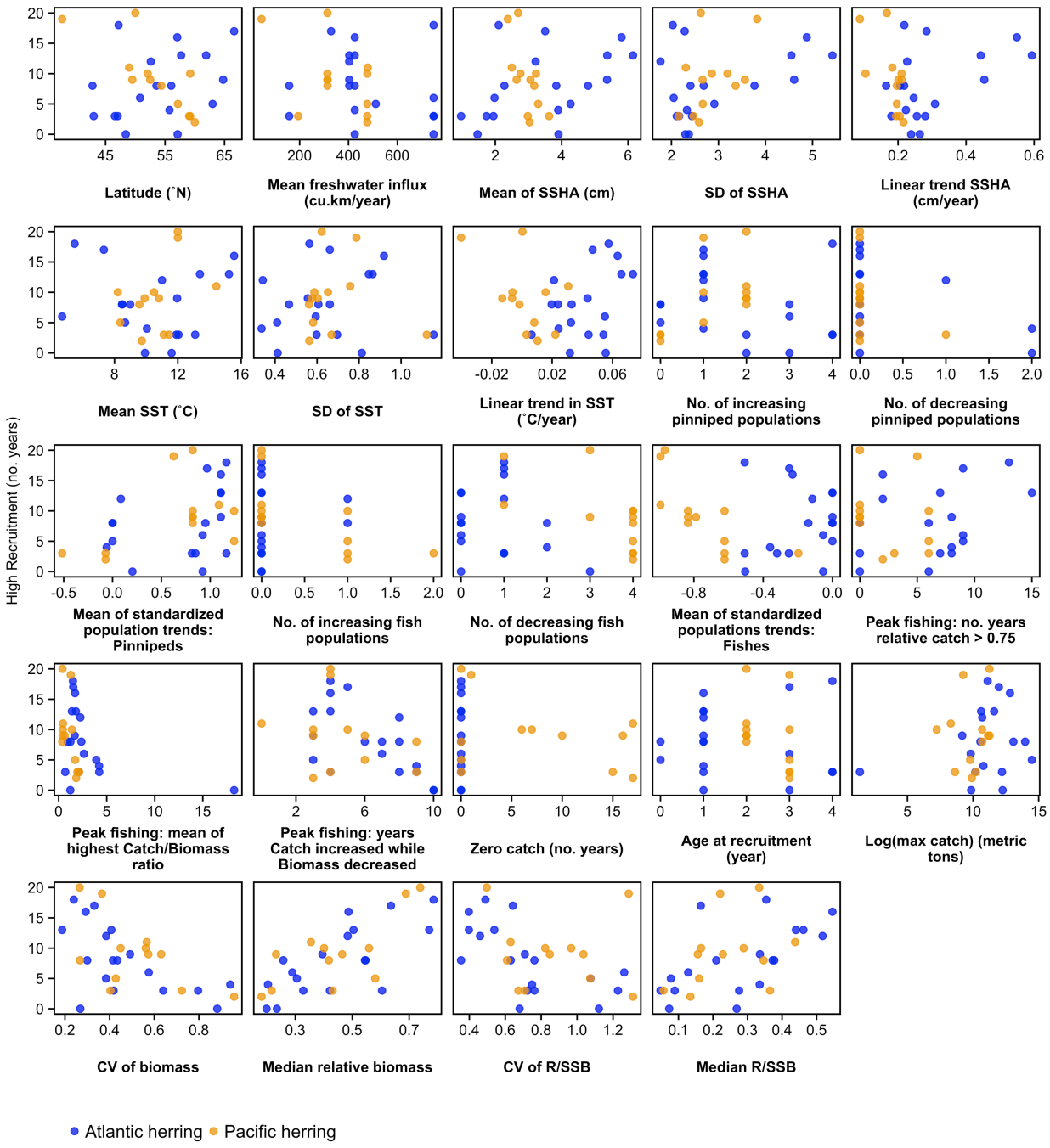

Figure 1.5. Observed high recruit year counts > 50% of Mean High Recruitment with their respective covariate values for each of the variables explored with the zero-inflated generalized linear mixed effects models. Observations are colored by species (Atlantic herring=blue, Pacific herring=orange).

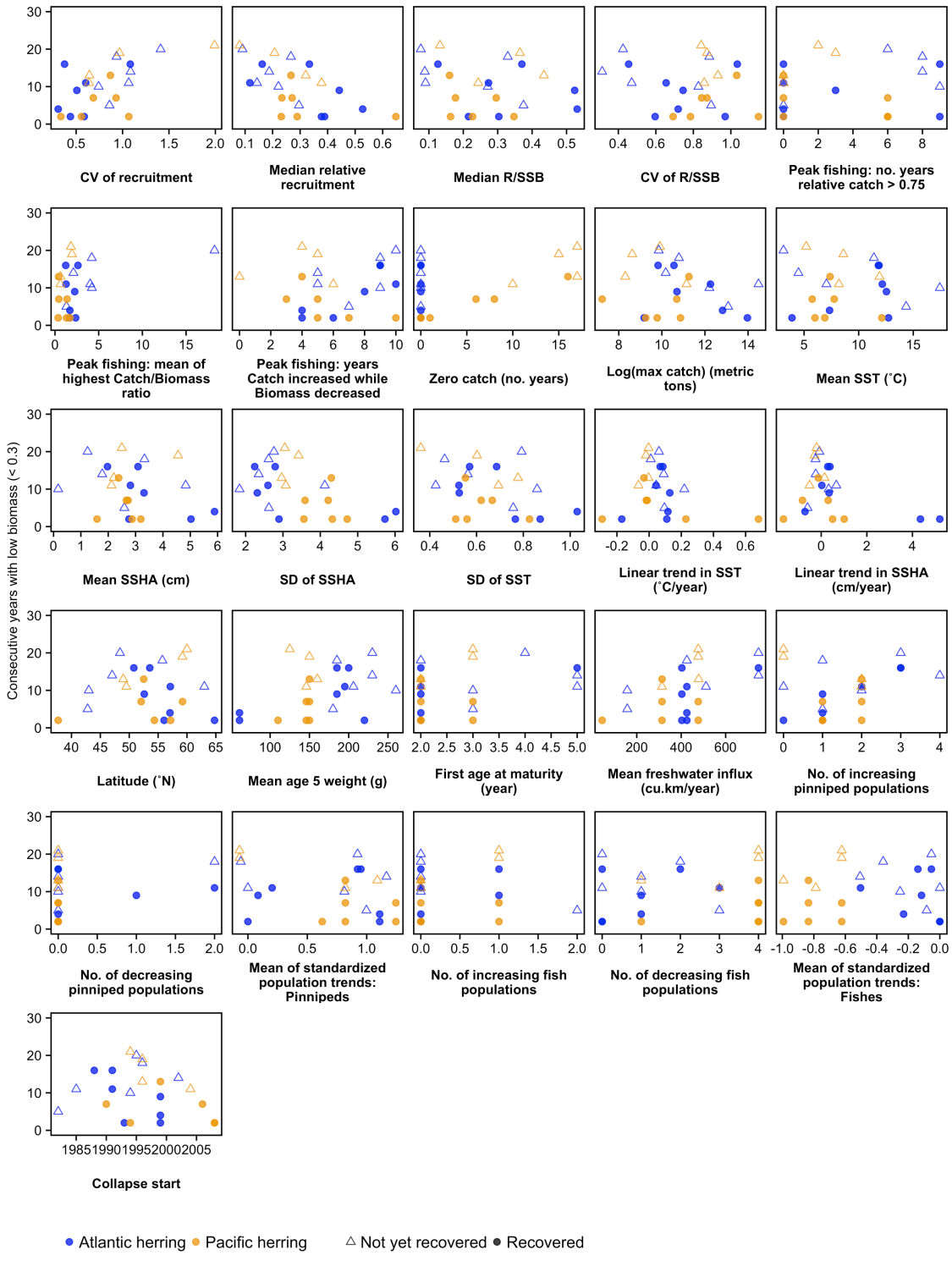

Figure 1.6. Observed low biomass durations < 30% of Mean High Biomass (N=23) with their respective covariate values for each of the variables explored with the Random Survival Forest. Observations are identified by species (Pacific herring=orange, Atlantic herring=blue) and also by whether the end of the collapse time was observed within the time series (Recovered=solid circle) or not (Not recovered=empty triangle).

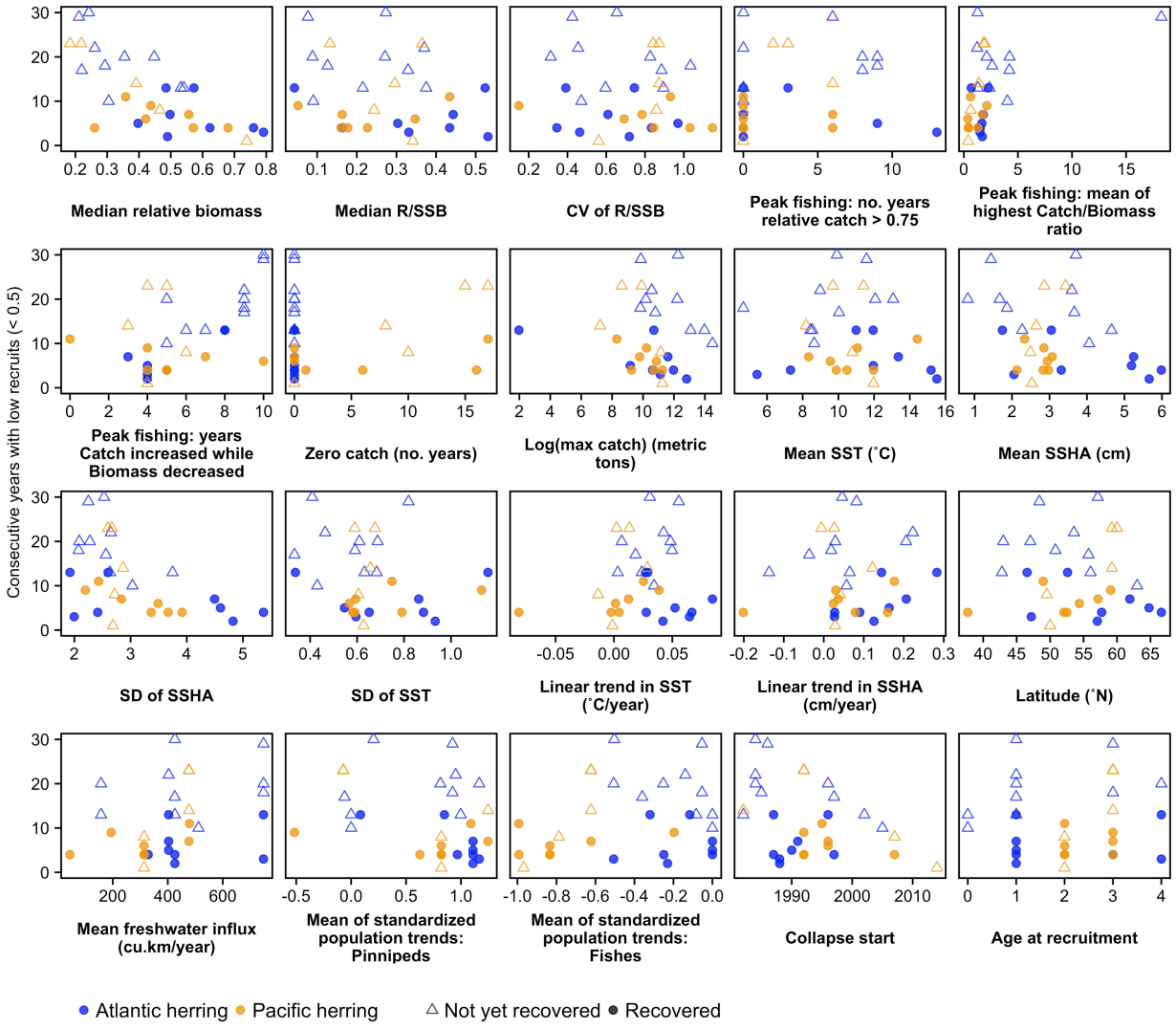

Figure 1.7. Observed low recruitment durations < 50% of Mean High Recruitment (N=30) with their respective covariate values for each of the variables explored with the Random Survival Forest. Observations are identified by species (Pacific herring=orange, Atlantic herring=blue) and also by whether the end of the collapse time was observed within the time series (Recovered=solid circle) or not (Not recovered=empty triangle).

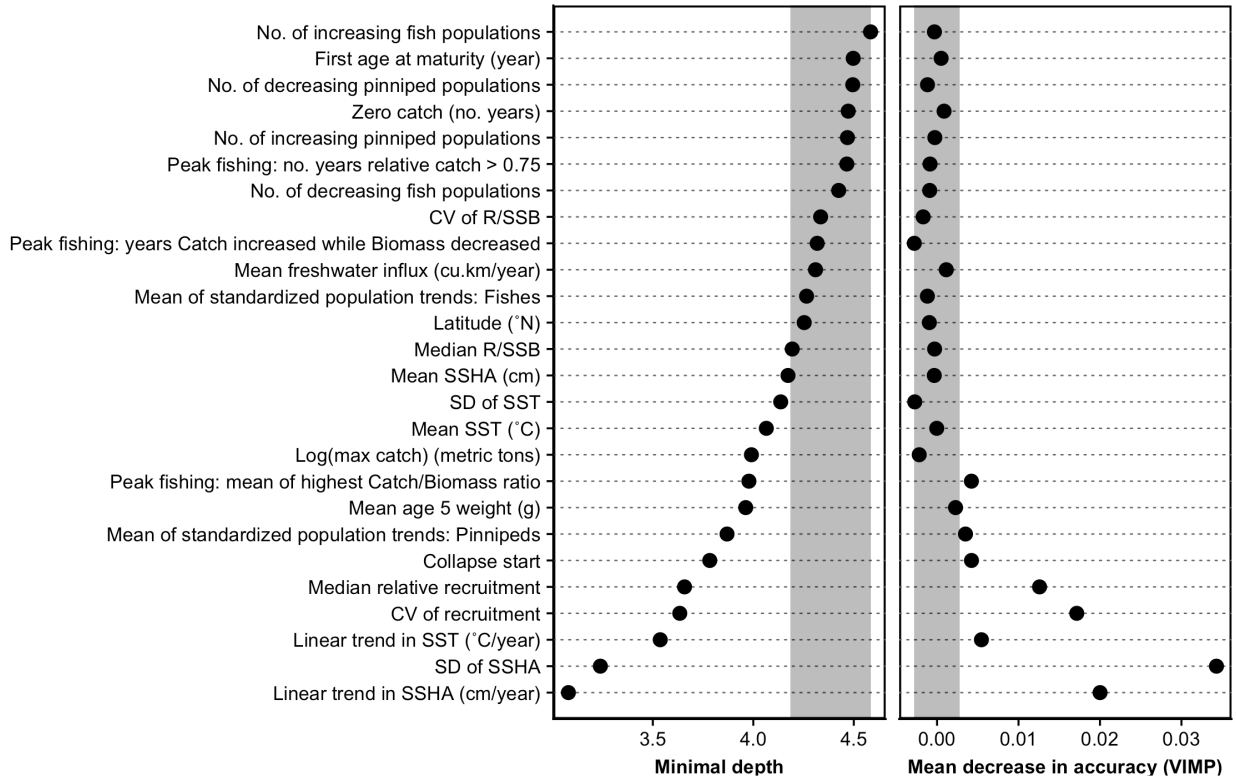

Figure 1.8. Metrics for selecting the most important variables from the Random Survival Forest regression on 23 collapse durations (collapse as biomass < 30% Mean High Biomass). Variables are ordered by their importance in terms of minimal depth (top variables at the bottom), or the average node on which the variable is selected to split across all trees in the Random Survival. The lower the node, the more frequently the variable best splits all the data. Alternatively, variable importance or VIMP is the difference in OOB Harrell's concordance index before and after random permutation, with larger values indicating an increase in error when randomly permuted. The larger the error, the larger the decrease in predictive accuracy and greater chance there is a true effect. Variables near and within the shaded regions are considered unimportant, which is above the median of all minimal depth values, and below the absolute value of the minimum VIMP score.

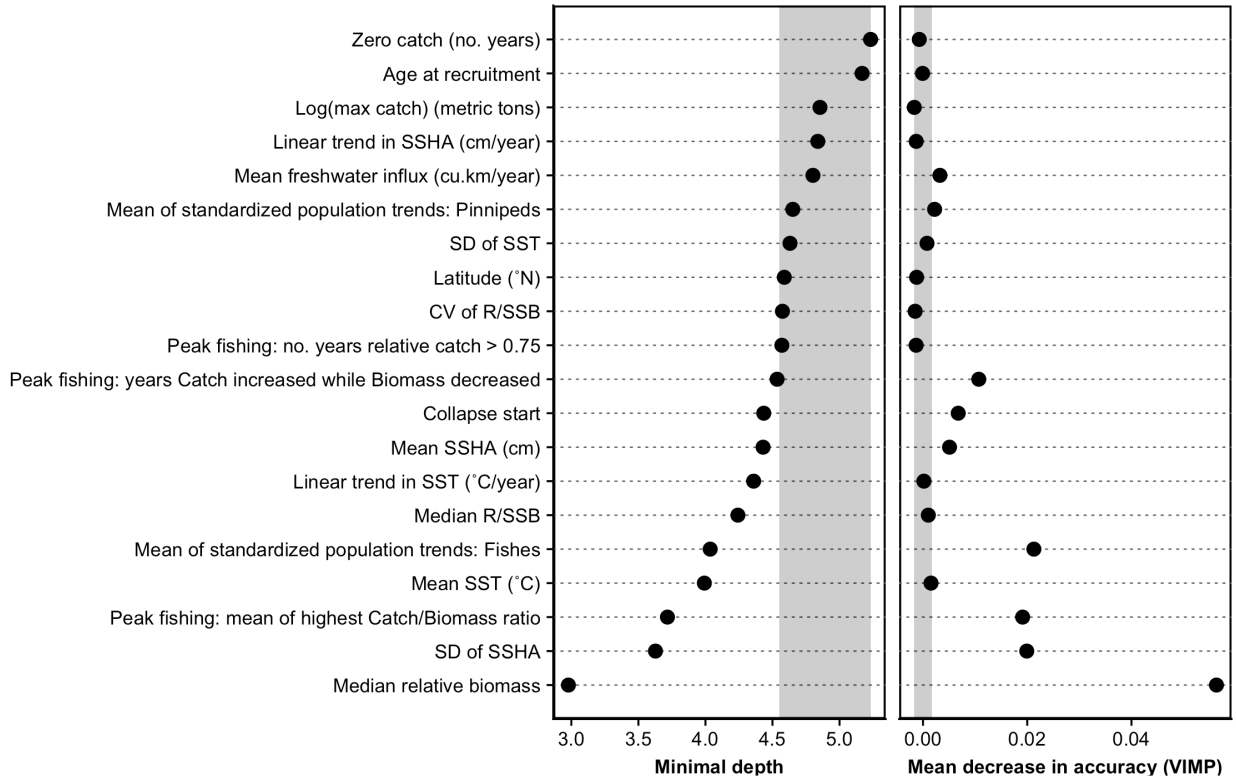

Figure 1.9. Metrics for selecting the most important variables from the Random Survival Forest regression on 30 low recruitment periods (low as recruits <50% of Mean High Recruitment). Variables are ordered by their importance in terms of minimal depth (top variables at the bottom) and displayed alongside their respective variable importance (VIMP). Variables near and within the shaded regions are considered unimportant, which is above the median of all minimal depth values, and below the absolute value of the minimum VIMP score.

Chapter 2. This clears up nothing! Applying multiple Bayesian model selection criteria to ecological covariates in stock assessment

2.1 ABSTRACT

Incorporating ecological covariates into fishery stock assessments may improve predictions, but some covariates are estimated with error. Model selection criteria are often used to identify support for covariates, have some limitations and rely on assumptions that are often violated. For a more rigorous evaluation of ecological covariates, we developed an approach using four popular selection criteria to identify covariates influencing natural mortality or recruitment in a Bayesian stock assessment of Pacific herring (*Clupea pallasii*) in Prince William Sound, Alaska. Within this framework, covariates were incorporated either as fixed effects or as latent variables (i.e., covariates have associated error). We found most support for pink salmon increasing natural mortality, which was selected by three of four criteria. There was ambiguous support for other fixed effects on natural mortality (walleye pollock and the North Pacific Gyre Oscillation) and recruitment (hatchery-released juvenile pink salmon and a 1989 regime shift). Generally, similar criteria values among covariates suggest no clear evidence for a consistent effect of any covariate. Models with covariates as latent variables performed worse, implying the assumed amount of covariate error is a key consideration. We recommend using multiple criteria and exploring different statistical assumptions about covariates for their use in stock assessment.

2.2 INTRODUCTION

Population dynamics models, such as those used in fisheries management, are governed by biological parameters including growth, recruitment, and natural mortality (Hilborn & Walters, 1992). Estimating recruitment and natural mortality is perhaps the most challenging obstacle to conducting accurate fisheries stock assessments. Recruitment predictions that rely on a relationship with estimated parameters are a key source of uncertainty in stock assessment (e.g. Maunder & Deriso, 2003; Maunder & Watters, 2003; Needle, 2001), in large part because of the high variance around estimated stock-recruitment relationships for many fish stocks (Gilbert, 1997; Lee, Maunder, Piner, & Methot, 2012; Szuwalski et al., 2015). Natural mortality of young and old fish is also a key uncertainty (Clark, 1999; Vetter, 1988), proving difficult to estimate and causing biased estimates of stock status when mis-specified, especially when ignoring time-varying mortality (Deroba & Schueller, 2013; Johnson et al., 2014). In fisheries research, one of the major driving questions is which ecological factors are most responsible for variation in recruitment and natural mortality. Little progress has been made in addressing this question (Pepin, 2015; Rice & Browman, 2014), but efforts continue because improving the accuracy and precision of stock assessments would result in more sustainable fish stocks and fisheries.

Reliably modeling ecological effects on recruitment or natural mortality can involve a variety of functions and analyses, but often starts (and sometimes stops) with linear models. In other words, ecological covariates, or the observable variable, are often used as predictors of recruitment or productivity in a linear or log-linear manner, and their effects are additive (Deriso et al., 2008; Maunder & Watters, 2003). This form treats covariates as fixed effects and provides a convenient link between ecological and population dynamics and accommodates hypotheses regarding the specific biological processes that are impacted.

Despite the ease and convenience of this approach, inappropriate assumptions about the covariates often undermine the robustness of these models. One of these inappropriate assumptions occurs because the observations used for covariates have statistical error (i.e. the "errors-in-variables" problem; Walters & Ludwig, 1981). Many ecological covariates are estimates from external analyses that are themselves uncertain such as time series of abundance for predator species that come from population dynamics models. Overlooking this uncertainty when incorporating as covariates into stock assessment may lead to erroneous conclusions (Brooks & Deroba, 2015). Additionally, the covariate itself may imperfectly represent the true forcing ecological factor and act in concert or interact with other factors, and this unexplained variability should be treated as random effects (Deriso et al., 2008; Maunder & Watters, 2003). Methods to address these issues include state-space formulations (Maunder, Deriso, & Hanson, 2015; Miller, Hare, & Alade, 2016) or modeling covariates “as data” (Crone, Maunder, Lee, & Piner, 2019; Schirripa, Goodyear, & Methot, 2009); both approaches more generally treat covariates as latent variables. Such models more appropriately represent covariate uncertainty, although their performance results in little improvement compared to models with fixed covariate effects and an appropriate bias correction (Crone et al., 2019).

Another covariate-specific issue is how to deal with missing covariate data. Various approaches may address this issue including estimating random effects in years of missing covariate data, substituting the mean of the available covariate data for missing years (“imputation”), or ignoring all fitted data in the missing year (Maunder & Deriso, 2010). Maunder and Deriso (2010) show through simulations that random effects, or state-space framework are good approaches (Maunder & Thorson, 2019), although substituting the mean of the covariate data may also perform well under some circumstances.

As with other species, including ecological covariates in herring (genus *Clupea*) stock assessments has long been a focus (Deriso et al., 2008; Deroba et al., 2018; Hulson, Miller, Dressel, Quinn II, & Van Kirk, 2018; Okamoto et al., 2020). This is in part because herring displays large fluctuations in abundance (Hjort, 1914) as well as prolonged periods of low adult abundance and recruitment (Trochta et al., 2020). Some stocks can remain at low levels for decades. For example, Pacific herring in Prince William Sound, Alaska, failed to recover following population collapse despite fisheries being closed for more than two decades. This failure to recover from low levels is unusual for fish stocks (e.g. Hilborn, Hively, Jensen, & Branch, 2014). Various studies have investigated biological and ecological factors that may inhibit the recovery of Prince William Sound herring, each providing different answers (Brown & Norcross, 2001; Deriso et al., 2008; Pearson et al., 2012; Sewall et al., 2017; Ward et al., 2017; Williams & Quinn, 2000). The reliance of various marine species (Moran, O'Dell, Arimitsu, Straley, & Dickson, 2018b; Stocking, Bishop, & Arab, 2018; Straley et al., 2017; Thomas & Thorne, 2003) and fisheries (Beaudreau et al., 2019) on herring motivates the continued need for reliable predictions of herring stock status from the current single-species Bayesian age-structured stock assessment model (Muradian et al., 2017).

This Bayesian assessment allows for the incorporation of priors for parameters and characterizes uncertainty using posterior distributions, which offers robust probabilistic decision analysis for fisheries management (Muradian et al., 2017; Punt & Hilborn, 1997). The Bayesian assessment also provides a valuable tool for evaluating different hypotheses about which covariates influence stock status. A variety of model-selection methods is available for Bayesian inference on evaluating support for individual covariates, each with its benefits and limitations (Hooten & Hobbs, 2015).

Commonly used Bayesian model selection criteria include the Deviance Information Criterion (DIC; Spiegelhalter et al., 2002), Watanabe-Akaike's Information Criterion (WAIC; Watanabe, 2013), and posterior predictive loss (herein PPL; Gelfand & Ghosh, 1998), which maintain popularity largely because of their easy computation. Existing posterior samples from draws of a Markov Chain Monte Carlo sampler are used to calculate DIC and WAIC, while posterior predictive draws are used in PPL. Another criterion, the Pareto-smoothed Importance Sampling Leave-one-out Criterion (PSIS-LOO), was more recently developed and shown to be more robust than these other criteria (Piironen & Vehtari, 2017; Vehtari, Gelman, & Gabry, 2017). Statistically, all these criteria are approximations of a "true" utility function that measures the predictive performance of a model (i.e. the Kullback-Leibler divergence between the true data generating distribution and the predictive distribution of a candidate model; Piironen & Vehtari, 2017). However, any one criterion is vulnerable to selecting the incorrect model, especially when models are overfitted or misspecified (Hooten & Hobbs, 2015; Piironen & Vehtari, 2017).

Here we evaluated the predictive ability of ecological covariates in the stock assessment of Prince William Sound herring using multiple Bayesian model selection criteria. The essence of the approach was to incorporate ecological covariates directly into the mortality and recruitment functions within the Bayesian assessment. We investigated the implications on how covariates are incorporated by running individual Bayesian assessment models with covariates incorporated as fixed effects and as latent variables. Since several covariates are systematically missing observations (e.g. data started or ended part way through the modeling time period), we created sets of models covering shorter or longer time periods, each of which had more complete data for all covariates. The models with longer time periods incorporated only those covariates

with long time series, while the models with shorter time periods were able to include more covariates. This approach allowed temporally consistent information for comparing models using Bayesian model selection. Finally, we applied DIC, WAIC, PPL and PSIS-LOO model selection criteria to check for inconsistencies in support between criteria. Altogether, our methods provide a framework for accounting for major technical issues involved in incorporating and selecting covariates for fisheries stock assessments: covariate data errors, missing covariate data, and model selection fallibility.

2.3 METHODS & MATERIALS

We reviewed the literature on hypotheses related to ecological factors driving Prince William Sound herring recruitment and natural mortality (hereafter "mortality") and collected corresponding covariate time series for inclusion in the Bayesian assessment model for Prince William Sound herring. We also describe the model fitting procedure and modifications made to the Bayesian assessment to incorporate covariates; how we dealt with missing covariate data; how we evaluated covariates using Bayesian model selection; and the alternative modeling approach incorporating covariates as latent variables.

2.3.1 Covariates of ecological factors impacting Prince William Sound herring

Various ecological factors have been proposed to impact Prince William Sound herring recruitment and adult (i.e. 3+ years and older) mortality rates. Modeling studies suggest that recruitment and mortality drive current population dynamics in Prince William Sound and that food quality and quantity, predation, oceanographic processes, and broad-scale climate drivers may explain their variability over time (Brown & Norcross, 2001; Deriso et al., 2008; Pearson et

al., 2012; Sewall et al., 2017; Ward et al., 2017; Williams & Quinn, 2000). Here we describe the covariates examined in this study (notated by I_y in equations below), with an overall summary and references given in Table 2.1.

Viral hemorrhagic septicemia virus (VHSV) and *Ichthyophonus hoferi*. Disease, specifically epizootics of VHSV and ulcers, and continuous infections of the protozoan parasite *I. hoferi*, have been hypothesized to be major determinants of Prince William Sound herring mortality (Marty et al., 1998; Marty et al., 2010; Marty et al., 2003; Quinn, Marty, Wilcock, & Willette, 2001). Three sets of disease data are currently incorporated into the Bayesian assessment model for Prince William Sound herring as an additive effect on adult natural mortality (Muradian et al., 2017): a combined prevalence index of VHSV and ulcers assumed to affect the mortality rate of ages 3-4, *I. hoferi* prevalence from field collections during 1994-2006 assumed to affect ages 5+, and *I. hoferi* prevalence from a new survey during 2007-present assumed to affect ages 5+. Since previously supported models incorporate all disease data (Marty et al., 2010; Muradian et al., 2017), we either include or exclude all disease data in the model.

Summer upwelling. Upwelling drives coastal primary productivity which may influence bottom-up control on herring productivity. The summer upwelling index describes the magnitude and direction of water transport and is calculated as the average of monthly Bakun (1973, 1975) upwelling indices ($m^3/s/100$ m) over May-September at a 3-degree cell centered on 60°N 146°W (https://oceanwatch.pfeg.noaa.gov/products/PFELData/upwell/monthly/upindex.mon).

North Pacific Gyre Oscillation (NPGO). NPGO reflects patterns in the variability of sea level, westerlies, winter temperatures, and precipitation (Di Lorenzo et al., 2008), which may also influence primary productivity dynamics in the Gulf of Alaska. NPGO is the second Principal Component from the empirical orthogonal function of sea-surface temperature (SST) and sea-

surface height anomalies (SSHA) over the Northeast Pacific (http://www.o3d.org/npgo/). Here, summer NPGO is the average over May-September (i.e. when herring primarily feed and generate lipid storage for future energy expenditure), and winter NPGO the average over November-March (i.e. when overwintering herring may need to rely on energy stores if prey availability is low) the following year.

Pacific Decadal Oscillation (PDO). PDO (the first Principal Component of SST and SSHA variability) is a pattern of climate variability in the mid- to north-Pacific that is expressed as phases of warmer or cool SST in the northeast Pacific and correlates with many marine populations (Mantua & Hare, 2002; Mantua, Hare, Zhang, Wallace, & Francis, 1997; Polovina, Mitchum, & Evans, 1996). Values were downloaded from http://research.jisao.washington.edu/pdo/. Here, summer PDO is the average over May-September, and winter PDO the average over November-March the following year.

Total pink salmon run and hatchery pink salmon releases. Pink salmon in Prince William Sound prey on herring and other species (Kaeriyama et al., 2000; Sturdevant et al., 2013), and may also compete with them for food. Total numbers of wild pink salmon (escapement + harvest) returning to Prince William Sound each year were obtained from ADF&G estimates (R. Brenner, pers. comm.). Releases of juvenile pink salmon from Prince William Sound hatcheries predicted long term shifts in Prince William Sound herring recruitment, implying that pink salmon either competed with or preyed on herring (Deriso et al., 2008; Pearson et al., 2012). These releases drastically increased in the late 1980s and have remained stable since the early 1990s. The number of hatchery-released pink salmon fry in Prince William Sound were obtained from ADF&G (pers. comm. R. Brenner, unpublished data).

Gulf of Alaska arrowtooth flounder total spawning biomass. Herring are eaten by Gulf of Alaska adult arrowtooth flounder (>20 cm), which has increased in abundance substantially since the 1980s (Spies, Aydin, Ianelli, & Palsson, 2017). We used stock assessment estimates of arrowtooth flounder total biomass (ages 1+) from the Gulf of Alaska (Spies et al., 2017).

Gulf of Alaska walleye pollock spawning biomass (SB) and age-1 numbers (lagged 1-yr). While walleye pollock eat herring within Prince William Sound (Gray, Bishop, & Powers, 2019; Thorne, 2008), a stronger effect may be reflected by the relative availability of walleye pollock and herring to dominant predators in the Gulf of Alaska such as arrowtooth flounder (Barnes, Beaudreau, Dorn, Holsman, & Mueter, 2020; Dorn et al., 2017; Oken, Essington, & Fu, 2018) and Steller sea lions (Trites & Donnelly, 2003). Local estimates of walleye pollock in Prince William Sound are unavailable, but spawning biomass estimates from Gulf of Alaska walleye pollock are available and used here (Dorn et al., 2017). Thus, the hypothesis we specifically evaluate is that Gulf of Alaska walleye pollock abundance decrease mortality due to prey switching by shared predators. Age-1 Gulf of Alaska walleye pollock were strongly and positively correlated with Prince William Sound herring productivity up to 2012, suggesting shared bottom-up effects of zooplankton prey or prey switching by shared predators (Sewall et al., 2017). Numbers of age-1 walleye pollock were obtained from the Gulf of Alaska stock assessment (Dorn et al., 2017), and lagged by 1 year to match the brood year of Prince William Sound herring.

Humpback whales. Humpback whales are major predators of herring throughout the northeast Pacific and in Prince William Sound (Moran et al., 2018a; Straley et al., 2017). Two separate time series of humpback whale abundance are used in this analysis: model estimates of summer Prince William Sound humpback whale abundance through 2009 (Teerlink et al., 2015) and humpback

whale counts from sighting surveys within Prince William Sound during the fall and winter (https://gulfwatchalaska.org/monitoring/pelagic-ecosystem/humpback-whales/).

Freshwater discharge. Freshwater discharge into Prince William Sound impacts quality of nearshore nursery habitats for juvenile herring, changing zooplankton prey timing and quantity (Ware & Thomson, 2005) and altering salinity, which in turn cues changes in larval and juvenile fish behavior (Boehlert & Mundy, 1988). We used annual indices of freshwater discharge near Seward, AK (Royer, 1982), which is positively associated with productivity of Prince William Sound herring (Ward et al., 2017).

First-year scale growth increment. First-year scale increments in Prince William Sound herring measures growth rates in the first year of life, and is strongly correlated with planktonic prey abundance and warmer summer temperatures (Batten, Moffitt, Pegau, & Campbell, 2016). We included a time series of scale increments from archived herring scale images collected from Prince William Sound (Haught & Moffitt, 2018).

1989 regime shift. The year 1989 marked two ecologically significant events in Prince William Sound: the Exxon Valdez oil spill and a climate regime shift (Hare & Mantua, 2000). These two events confound analyses on the cause of dramatic decreases of herring and salmon populations in Prince William Sound that occurred during or shortly after this same time (Ward et al., 2017). To account for these factors, we included a time-block effect with a shift in estimated mean recruitment.

Null model. The null model includes no covariates on natural mortality and recruitment.

2.3.2 Bayesian age-structured assessment model

Each ecological covariate was incorporated into the Bayesian assessment (BASA for shorthand; Muradian et al., 2017), an updated version of the ADF&G assessment model used in previous modeling studies (Deriso et al., 2008; Deriso, Maunder, & Skalski, 2007; Hulson et al., 2007). Six key datasets were fit by the model: relative abundance indices from two hydroacoustic surveys conducted respectively by the Prince William Sound Science Center (PWSSC) and ADF&G; a relative abundance index from an aerial survey of milt coverage standardized by length of shoreline and days surveyed; an absolute abundance index from an egg deposition survey; fishery-dependent age compositions from the purse-seine fishery; and fishery-independent age compositions from seine and cast net surveys on pre-spawning aggregations of herring (Muradian et al., 2017). Since the Bayesian assessment has been thoroughly documented in earlier literature, we only describe the equations incorporating covariates. We also made minor changes in how Muradian et al. (2017) calculated mature biomass, to improve estimation, and altered the model to start at age 0 instead of age 3 to allow for covariates to affect younger ages. These changes are described in Appendix B.

Recruitment (R_y) was modeled as spawner independent where process error varies around constant mean recruitment. Ecological effects contribute to the process error in proportion to an estimated β (the effect size of covariate I_y), where ε_y is the estimated unexplained error in recruitment variation with log-normal bias-correction and $\bar{R}$ is mean stationary recruitment across time:

$$R_y = \bar{R} e^{\beta I_y + \varepsilon_y - 0.5\sigma^2}$$

$$\varepsilon_y \sim N(0, \sigma^2)$$

$$\sigma \sim U(0.0001,2)$$

There is a uniform prior that constrains σ (recruitment standard deviation) to a positive variance, and differs from BASA (Muradian et al., 2017), which freely estimated annual recruitment.

Survival is a function of mortality that was modeled for two periods within each year to account for the seasonal fisheries that once operated in Prince William Sound. Survival ($S_{a,y,b}$) of adult herring of age *a*, in year *y*, and half-year *b* (1 or 2) is:

$$S_{a,y,b} = e^{-0.5\bar{M}+\beta I_{a,y,b}} \quad 0 \le a \le 8$$

in which $\bar{M}$ is the assumed average annual instantaneous mortality rate multiplied by 0.5 for a bi-annual mortality rate, and an estimated β measures the influence of covariate $I_{a,y,b}$ in half-year b. The value of $\bar{M}$ is a constant and fixed at 0.25 yr^{-1} (Muradian et al., 2017). Half-year survival in the age 9+ group is:

$$S_{9+,y,b} = \begin{cases} e^{-0.5\bar{M}_{9+}+\beta I_{9+,y,b}} & y = 1980 \\ S_{9+,y-1,b}\left(\dfrac{S_{8,y,b}}{S_{8,y-1,b}}\right) & y > 1980 \end{cases}$$

in which $\bar{M}_{9+}$ is the instantaneous mortality rate of the plus group in the first year, and is estimated. In all other years, whatever changes were made to age 8 survival are also made to age 9+ survival; therefore, any covariate applied to age 8 is also referred to as having affected age 9+.

Each covariate was normalized to have a mean of 0 and standard deviation of 1 over the time series, and only one covariate was included in BASA at a time to provide a suite of independent models (Table 2.1). Each covariate was assumed to affect one or more age groups: the affected age groups were all affected in the same way, while the unaffected age groups had β = 0. Covariate effects on ages 9+ survival were implicit since they were related to age 8 survival. This linear age-structured formulation for mortality is identical to the current formulation in BASA

that incorporates an index of disease prevalence rate (Muradian et al., 2017), except that the disease indices were not normalized and were assumed to influence mortality over the entire year.

BASA includes two additional, freely estimated mortality parameters that were added to $m_{a,y,b}$ in 1992-1993 to account for the sudden and significant loss of biomass observed in the milt and acoustic surveys in those years (Hulson et al., 2007; Marty et al., 2010). One mortality parameter acted on ages 3-4 and the other on ages 5-8 (Muradian et al., 2017). Excluding these two parameters made no difference in the top models our analysis selected and resulted in worse fits to the data and poorer convergence. Here, we report values of these two parameters for each model as a check on whether covariates may partially explain increased mortality in 1992-1993.

2.3.3 Other modifications to BASA

We modified two other components of BASA to correct for potential model misspecification in Muradian et al. (2017). The first was changing hydroacoustic survey biomass to represent pre-fishery mature biomass instead of age 3+ biomass, as follows:

$$\widehat{H}_y = e^q \sum_{a \in A} \rho_{M,a} N_{y,a} w_{y,a}$$

in which $\widehat{H}_y$ is the acoustic estimate in year y, q is a scalar for the acoustic estimate (log-link), $\rho_{M,a}$ proportions mature-at-age, $N_{y,a}$ numbers-at-age, and $w_{y,a}$ weight-at-age. Muradian et al. (2017) previously assumed hydroacoustic surveys captured total adult biomass (all age 3+), although recent data from sampling of the aggregations targeted by acoustic surveys revealed they are mostly mature fish (unpublished data W. Pegau). The second modification was to estimate proportions mature at ages three and four over the entire time period (1980-2017) instead of estimating two sets of proportions for two time periods split at 1997 (Muradian et al., 2017).

Sensitivity analysis shown negligible difference in biomass and recruitment estimates, while estimating a maturity for a single time period improved model convergence.

2.3.4 Addressing missing covariates

Multiple covariates have observations that start or finish during the modeled time frame and are missing values especially in early years. To make model comparisons and selection consistent so that the same time periods are affected across all covariates, we re-ran the model on four time periods with different numbers of years removed from the early or later part of covariate time series with cut-off years matching the first of last year of observations for incomplete time series (see Table 2.1). The time periods are 1980-2009, 1980-2017, 1994-2017, and 2007-2017. The complete records (1980-2017) of the six fitted data sets are used in all models across all time periods.

We then compared model results within each time frame. This approach is similar to that used by Sewall et al. (2017). Some covariates are missing values in individual years or for several years at the end (see Table 2.1). We did not systemically omit these years in other covariates because these instances are too few to substantially impact results and would require running many more Bayesian models. Additionally, since all covariates are normalized to have a zero mean, missing years are analogous to an effect of the mean covariate value within the model (i.e. substituting missing years with the mean value), which was previously demonstrated as a possible alternative for addressing missing covariate values (Maunder & Deriso, 2010).

2.3.5 Bayesian model-fitting

BASA was implemented in AD Model Builder (ADMB; Fournier et al., 2012). Parameter estimation was done using the no-U-turn sampler (NUTS), a more efficient Markov chain Monte Carlo (MCMC) algorithm for sampling from the posterior distribution (Monnahan, Thorson, & Branch, 2017). We used the R package "adnuts" (Monnahan & Kristensen, 2018) to run ADMB with NUTS inside R (R Core Team, 2020). Three chains of 3000 samples were generated using a diagonal mass matrix (the default in adnuts) adapted with a warm-up phase of 500 samples and a target acceptance rate of 0.925. The results from all chains were combined. To assess convergence in each model, we checked for sufficient potential scale reduction $\hat{R}$ values (<1.1; Gelman, Carlin, Stern, & Rubin, 2014a) of each parameter across chains and zero divergences. Models typically converged in 30-90 minutes.

2.3.6 Hypothesis evaluation

We used two general approaches to evaluating support for each ecological covariate: 1) computing posterior probabilities of estimated effects and 2) Bayesian model selection.

The posterior probabilities of the estimated effects (β) of each covariate was calculated as the proportion of posterior draws greater than or less than zero depending on the sign of the effect implied by the hypotheses (Table 2.1). We directly compared effect probabilities among models because all covariates were normalized and thus estimates for β are on the same scale.

Various Bayesian model selection criteria have been developed, but we focused on those that have either been widely used, especially for stock assessment, or could be readily computed from existing posterior draws, namely Deviance Information Criterion (DIC), Watanabe Akaike Information Criterion (WAIC), Posterior Predictive Loss (PPL), and Pareto-smoothed Importance

sampling Leave-one-out cross validation (PSIS-LOO). Calculating criteria values involves multiple computational steps (Fig. 2.1). Details on how they are applied to the multiple data sets within BASA are provided in Appendix B.

There are similarities in how these criteria are computed, such as the use of posterior densities for model estimates of the data (WAIC and PSIS-LOO), but also key differences and caveats to each. DIC (Spiegelhalter, Best, Carlin, & Van Der Linde, 2002) has been widely used with stock assessment (Brooks et al., 2019; Punt, Hurtado-Ferro, & Whitten, 2014; Wilberg & Bence, 2008), but poorly characterizes and favors model complexity, is biased when the posterior distribution is not multivariate normal, and disregards uncertainty, a key benefit of Bayesian inference (Hooten & Hobbs, 2015). WAIC is preferred to DIC because it integrates over the posterior densities (Fig. 2.1) and asymptotically approximates conventional leave-one-out cross-validation, but its reliance on reusing the data to estimate out-of-sample prediction error can lead to high variance and result in choosing the nonoptimal model (Piironen & Vehtari, 2017; Vehtari et al., 2017). PPL (Gelfand & Ghosh, 1998; Ibrahim & Laud, 1994; Laud & Ibrahim, 1995) considers simulated measurements from the posterior estimates of the data (i.e. posterior predictive distributions) and has been shown to penalize more complex models, but may be biased especially with non-normal posterior predictive distributions (Piironen & Vehtari, 2017).

PSIS-LOO has been shown to be a more reliable approximation of leave-one-out cross-validation and more robust to weak priors and influential observations compared to WAIC (Vehtari et al., 2017). However, PSIS-LOO is still subject to incorrectly estimating prediction accuracy under these conditions or when data are sparse. An added benefit to using this criterion is accompanying output that provides diagnostics on the reliability of PSIS-LOO values.

Specifically, calculating PSIS-LOO involves estimating tail shape parameters of the generalized Pareto distribution ($\hat{k}$) for each fitted observation; values of $\hat{k}$ should not exceed 0.7 for most predictions (Vehtari et al., 2017). Many problematic $\hat{k}$ indicate the PSIS-LOO value may be unreliable and in these cases, full *K*-fold cross-validation or model changes are recommended. We did not run *K*-fold cross-validation for models with many problematic $\hat{k}$ because there is no clear way to do this with a model that integrated catch-at-age model such as BASA. Further details on PSIS-LOO diagnostics are provided in Appendix B.

We used DIC, WAIC, PPL, and PSIS-LOO to rank the most supported hypotheses, including the null model. The null model provides a benchmark for comparison in which alternative models need to have lower criteria values than the null model to be considered better (i.e. better than using no covariate information). The best model should produce the lowest values under each criterion. Model selection with the full model (all covariates) was beyond the scope of this study which is to evaluate individual hypotheses and identify the single most important factors.

Another important aspect of our criteria computations is that random effects (i.e. estimated recruitment deviations) and latent variables were sampled along with other parameters, resulting in conditional likelihoods that enter the calculated posterior densities. Using conditional likelihoods resulted in suboptimal model selection with DIC and WAIC using marginal likelihoods (where random effects are marginalized out at in each posterior draw during sampling) in a state-space surplus-production stock-assessment model (Kai & Yokoi, 2019). However, the computation of marginal likelihoods in MCMC sampling is computationally infeasible for the much more complex BASA model.

2.3.7 Impact of covariates on population estimates

While it is important to examine estimated effect probabilities and support from model selection criteria, of even greater importance is the impact of selected covariates on key management quantities. Fisheries management relies on estimates of spawning biomass to decide on catch levels and rebuilding strategies, and on estimates of recruitment to predict future trends in spawning biomass. Therefore, we compared the posterior distributions of spawning biomass and recruitment estimated from each model with those from the null model to determine the implications of covariates on spawning biomass and recruitment.

2.3.8 Treating covariates as latent variables

The models assumed that covariates are fixed effects without error, but many covariates are model estimates themselves with accompanying estimates of uncertainty. This is true of stock assessment estimates we used as covariates (e.g. arrowtooth flounder, walleye pollock). To address this issue, we ran model variants where we incorporated ecological time series as latent variables of estimable process error. For the survival model, this alternative formulation is:

$$S_{a,y,b} = e^{-0.5\bar{M}+\varepsilon_{a,y}} \quad 0 \leq a \leq 8$$

where $\varepsilon_{a,y}$is a parameter estimated for each year that is available from the ecological time series and across ages impacted by the changes in mortality (e.g. if ages 3+ are impacted, then a single parameter is estimated for that age group in each year). A normal error distribution was specified for $\alpha\varepsilon_{a,y}$, where α is an estimated nuisance parameter that scales $\varepsilon_{a,y}$ to the normalized ecological time series, $I_{a,y}$:

$$\alpha\varepsilon_{a,y} \sim N(I_{a,y}, {\sigma_y}^2)$$

We fixed year-specific variance parameters, $\sigma_y{}^2$, to estimates of annual standard error or deviation values that are available for some time series. Time series without accompanying standard errors were assumed to have a constant standard deviation of σ_y=0.3 in all years. While this is arbitrary, it is a reasonable value similar to error magnitudes provided or estimated for other data included in BASA (i.e. fitting a covariate is roughly equally weighted compared to fitting other data). The prior is:

$$\sum_y \left[\ln(\sigma_y) + \frac{(\alpha \varepsilon_{a,y} - I_{a,y})^2}{2\sigma_y{}^2} \right]$$

The equations for the recruitment model and contribution to the objective function follow similar forms, but with lognormally distributed deviates and an unstandardized ecological time series:

$$R_y = \bar{R} e^{\varepsilon_y - 0.5\sigma^2}$$

$$\sum_y \left[\ln(\sigma_y) + \frac{(\ln(\alpha e^{\varepsilon_y}) - \ln(e^{I_y}))^2}{2\sigma_y{}^2} \right]$$

For the latent-variable model variants, we calculated DIC, WAIC, PPL, and PSIS-LOO to select the best models and compare their estimates of spawning biomass and recruitment with the null model.

2.4 RESULTS

2.4.1 Posterior probabilities of effects

For the model fitted to the longest time series of data (1980-2017), multiple covariates have high probabilities of an effect on natural mortality (>95%), which increased with higher winter and summer NPGO, higher total pink salmon returns, lower summer PDO, lower GOA walleye pollock

SSB, and higher GOA arrowtooth founder total biomass (Fig. 2.2). These estimated effects were mostly consistent in 1994-2017 (with the addition of an increasing effect with higher *I. hoferi* before 2007) and 1980-2009, except the probability for a total pink salmon effect substantially decreased for the 1980-2009 data. Over 1980-2009, a negative effect of winter PDO and positive effect of summer upwelling had high probabilities, as did a positive effect of summer humpback whales. For the shortest time period data (2007-2017), most covariates have low probabilities, except for summer upwelling and total pink salmon.

High probabilities (>95%) of increasing recruitment with lower hatchery-released juvenile pink salmon, higher GOA walleye pollock age 1, and an upward regime shift in 1989 are shown from 1980-2017 (Fig. 2.3). The median proportions of variance explained in $\log(R_y)$ from 1980-2017 is substantial for hatchery-released juvenile pink salmon and the 1989 regime shift, both at 0.37 (95% uncertainty from 0.10-0.79), while other covariates explained 0.16 or less.

Identical posterior probabilities for the effects of hatchery-release juvenile pink salmon and a 1989 regime shift are seen in all four time periods, but only in 1980-2017 and 1980-2009 for GOA walleye pollock. In 2007-2017, recruitment also likely correlated with higher summer PDO, lower summer NPGO, and high age-0 scale growth, which explained 0.43-0.67 of $\log(R_y)$ variance in those years.

2.4.2 Bayesian model selection

For natural mortality effects, model selection most consistently supported the model with total pink salmon returns (Fig. 2.4). The total pink salmon returns model is best in three of four criteria (PSIS-LOO, WAIC, and DIC) in 1980-2017, 1994-2017, and 2007-2017, but not in 1980-2009. In 1980-2009, total pink salmon returns were the worst model under all four criteria (Fig. 2.4) and

had the most number of $\hat{k}$ values from PSIS-LOO that were problematic (9 values) compared to the other covariate models (5-7 values for each model). Altogether, this suggests that total pink salmon returns from 2007-2017 are highly influential in model predictions for this same time period.

An effect of GOA walleye pollock SSB on natural mortality is the best model under one criterion (PPL) in 1980-2017, while winter NPGO was selected under this same criterion in 1994-2017 and 1980-2009 (Fig. 2.4). Multiple recruitment covariates were selected as well, including a tie between age 0 scale growth, summer PDO, and summer NPGO in 2007-2017 (under PPL), and hatchery-released pink salmon in 1980-2009 (under PSIS-LOO and WAIC). However, differences in criteria values between recruitment covariates and the null model are negligibly small, suggesting these models did not improve predictions. Furthermore, most models resulted in a number of problematic $\hat{k}$ values from PSIS-LOO (4-10), suggesting that PSIS-LOO values (and the other criteria) may be inaccurate or the models misspecified.

When incorporating covariates "as latent variables" into BASA, model selection differed substantially (Fig. 2.5), vaulting disease indices to the top position in three of four criteria (PSIS-LOO, WAIC, and PPL). Additionally, humpback whale counts rank similarly to disease indices under PPL. The model with total pink salmon returns is still the best under DIC.

2.4.3 Explaining the 1992-1993 decline

Model performance was evaluated with respect to their ability to explain the decline in spawning biomass in the early 1990s. If any of these covariates were able to at least partially explain this mass herring mortality, or a substantial decline in biomass in general (e.g. through persistent low recruitment), we would expect lower estimates of the two 1993 additive mortality parameters

compared to the null model. However, none of the mortality covariates reduced these parameter estimates, and some even increased the estimate of 1993 mortality (Fig. 2.6). Additional analyses running BASA with each covariate and without these two additional mortality parameters all resulted in worse performance amongst model selection criteria compared to the present results.

2.4.4 Consequences to population estimates

We examined the impacts of including the top covariates on resulting estimates of spawning biomass and recruitment—key outputs from BASA that are used by management (Fig. 2.7). Top natural mortality covariates (as fixed effects and as latent variables) tended to produce lower estimates of spawning biomass and more pronounced differences in trends in recent years, although the most consistently supported covariate, total pink salmon returns, estimated spawning biomass and recruitment that differed little from the null model. Hatchery-released juvenile pink salmon, one of the top recruitment models, also had no impact on spawning biomass and recruitment estimates; in fact, all recruitment covariates, when implemented as fixed effects, had little impact on recruitment estimates (not shown). However, when recruitment covariates were incorporated as latent variables (none of these models were unsupported by selection criteria), drastically different estimates of spawning biomass and recruitment resulted (not shown).

2.5 DISCUSSION

An effect of total pink salmon returns (including catch and escapement) on adult natural mortality had the most consistent support amongst criteria and in different time periods, but not in earlier years (before 2009). Still, the effect of pink salmon did not impact population estimates from BASA. Evidence for other covariates was more ambiguous: many covariates had a high probability

of an effect, fewer had support from model selection in general, and none had support for more than one criterion or time period. Altogether, no single covariate was a good predictor for the entire time period of collapse and failed recovery of Prince William Sound herring biomass and recruitment, but at least one covariate may partially inform variability in herring population dynamics.

2.5.1 Supported covariates of natural mortality

Our results support an antagonistic interaction between adult herring mortality and Prince William Sound pink salmon. The specific mechanism for pink salmon causing higher herring mortality is uncertain. Initially, predation of herring by pink salmon within Prince William Sound was thought to be virtually negligible (Okey & Pauly, 1999; Pearson et al., 2012), but there has been recent evidence for irregular localized predation impacts on Prince William Sound herring (Sturdevant et al., 2013). Adult pink salmon migrate inside and outside of Prince William Sound into the Gulf of Alaska, and exhibit a diverse diet that includes adult herring and herring prey items (Sturdevant et al., 2013). Thus there could also be competition between adult herring and pink salmon, as has been shown in Puget Sound, Washington state (Kemp, Beauchamp, Sweeting, & Cooper, 2013). The strengths of interactions with pink salmon through diet may also change with climate, migration, and the degree of overlap between the two species (Kaeriyama et al., 2000; Sturdevant et al., 2013). Interactions between Prince William Sound herring and pink salmon are also likely influenced by highly variable herring movement to and from the Gulf of Alaska (Bishop & Eiler, 2018), as concluded by a previous study that found a significant impact of pink salmon returns on Prince William Sound sockeye salmon productivity (Ward et al., 2017). Our results suggest the

value in further investigating interactions between Prince William Sound pink salmon and herring and characterizing their overlap in space and time.

There is weaker support for higher abundance of Gulf of Alaska walleye pollock being linked to lower age-3+ mortality (i.e. pollock abundance and herring survival are positively correlated). Direct overlap between these two populations is not evident, so the most likely cause is a third factor that impacts both populations. Some predators target both herring and walleye pollock in the Gulf of Alaska, including Steller sea lions (Trites & Donnelly, 2003; Womble & Sigler, 2006) and arrowtooth flounder; our analysis did show a positive correlation between arrowtooth flounder and herring mortality, and other evidence shows herring to be a small component of their diet (Spies et al., 2017; Yang, 1993). Prey switching by predators could occur depending on the relative availability of their prey, as has been implied for Steller sea lions (Trites & Donnelly, 2003). Another reason for their covariation is bottom-up forcing. Adult Pacific herring feed on lipid-rich crustaceans, other zooplankton, and small fish (e.g. Andrews, Strasburger, Farley Jr, Murphy, & Coyle, 2016), which are also eaten by walleye pollock (Dorn et al., 2017). Changes in prey availability and quality for both herring and walleye pollock may then have an identical effect on each species, such as influenced by climate conditions (e.g. Andrews et al., 2016).

Our analysis also suggested climate factors may have an effect on age 3+ mortality as well as recruitment. Posterior probabilities and model selection implicated effects of NPGO and PDO indices from summer and winter. It is difficult to hypothesize and interpret the signs of these effects because NPGO and PDO are not physical processes, but statistical summaries of emergent patterns across space and time, and associated with measurable physical and climate variables (e.g. SST and Sea Level Pressure field; Litzow et al., 2019; Litzow et al., 2020; Puerta, Ciannelli,

Rykaczewski, Opiekun, & Litzow, 2019). PDO had been the dominant climate pattern in the Gulf of Alaska (Di Lorenzo et al., 2008) and correlated with the productivity and abundance of various Gulf of Alaska fish populations; however, this correlation has changed over time and disappeared in recent years (Litzow et al., 2018, 2019; Litzow et al., 2020; Puerta et al., 2019). Following 1988/1989, NPGO explained more climate variance (Di Lorenzo et al., 2010; Yeh, Kang, Noh, & Miller, 2011) and also associated fish population dynamics such as salmon survival in the North Pacific (Kilduff, Di Lorenzo, Botsford, & Teo, 2015). NPGO also more recently lost its association with physical-ecological variables while having strengthened its anticorrelation with PDO (Litzow et al., 2020), which may explain why PDO and NPGO shown more likely, but opposite effects on recruitment in 2007-2017 compared to other time periods (Fig. 2.3). Given the evidence for non-stationarity in PDO and NPGO relationships, a superior approach would be to explicitly model time-varying relationships (e.g. Litzow et al., 2018, 2019; Litzow et al., 2020; Puerta et al., 2019) or identify time blocks that correspond with regime shifts, as has been done in relating PDO to natural mortality in another Gulf of Alaska herring stock (Hulson et al., 2018). This should be the next step for considering these climate indices in BASA and other stock assessment models.

Among the remaining covariates, when included as latent variables, disease was selected as the top model, with secondary support for total pink salmon returns and humpback whales. Previous lab, field, and modeling studies provided evidence that disease, specifically VHSV and *I. hoferi*, increased juvenile and adult herring mortality (Marty et al., 1998; Marty et al., 2010; Marty et al., 2003). However, a synthesis of the available evidence suggests that neither pathogen had a primary role in the collapse nor failed recovery of herring (Pearson et al., 2012). More importantly, the disease prevalence indices do not reflect the proportion that died, but the proportion that were infected and still alive at the time of sampling. In particular, *I. hoferi* can

cause acute mortality or persistent infections with selective mortality (e.g. selective vulnerability to predation) in subsequent years, although this is irregular (Hershberger et al., 2016). This may help to explain the substantial change in probability of *I. hoferi* increasing mortality before and after 2007 (Fig. 2.2). Data on the exposure history to these pathogens, such as from neutralizing antibody tests (Hart, MacKenzie, Purcell, Powers, & Hershberger, 2017), may better allow for a more accurate assessment of the impact of past infections on herring.

Humpback whales (summer estimates and overwinter counts) are also likely to increase mortality (Fig. 2.2). Humpback whales are frequently recorded targeting herring aggregations (Moran et al., 2018a; Pearson et al., 2012; Straley et al., 2017). Importantly, humpback whale consumption within Prince William Sound in the late 2000s was estimated at 21-77% of herring spawning biomass (Moran et al., 2018a). The summer abundance estimates and raw overwinter counts of humpback whales we used likely does not characterize the true extent of humpback predation on herring in Prince William Sound. Ancillary information, such as humpback prey selection and herring energy content as used by Moran et al. (2018a), is necessary to better account for the predation impact of whales within herring models.

The remaining covariates with negligible support, upwelling and arrowtooth flounder biomass, are not likely covariates of herring mortality on their own. Evidence for the influence of upwelling indices on Gulf of Alaska fish populations is not particularly strong (e.g. weaker than coastal SST effects on salmon survival; Mueter, Ware, & Peterman, 2002) despite being linked to herring recruitment elsewhere in the northeast Pacific (Reum et al., 2011; Williams & Quinn, 2000). Arrowtooth flounder in the Gulf of Alaska have a diverse diet where herring are a minor prey item compared to other species (Spies et al., 2017; Yang, 1993) and predation on herring is mitigated by the abundance of other prey species as noted earlier.

2.5.2 Ambiguous support for recruitment covariates

The ultimate aspiration in fisheries science is finding covariates that can explain and predict recruitment, but in our study case we did not find any convincing covariates that consistently predicted Prince William Sound herring recruitment. While effects of hatchery-released juvenile pink salmon and a 1989 regime shift had high probabilities across time periods and explained a moderate amount of variance, their time series showed long-term shifts with very little or no interannual variability that cannot explain the large pulses of individual cohorts that predominate recruitment variability. Furthermore, these two effects were identical in magnitude and proportion variance explained, which suggests a likely shift in average recruitment, but whose specific cause cannot be discerned from our analysis. The little support for scale growth and summer NPGO and PDO in 2007-2017 (by posterior probabilities and PPL) may also be plausible, especially for scale growth because it strongly correlated with the availability of appropriately-sized high-quality prey for young-of-the-year herring that may also affect herring survival (Batten et al., 2016). However, evidence for these three covariates is suspect since few years (11 years) are modeled and evaluated for support.

Our results further contrast with other recent modeling studies that identified predictors of Prince William Sound herring recruitment. Previously, Gulf of Alaska juvenile walleye pollock (Sewall et al., 2017) and freshwater discharge (Ward et al., 2017) were selected as top predictors for herring recruitment. However, these studies evaluated covariates with a Ricker stock-recruitment relationship, an assumption we avoided here because stock-recruitment relationships are poorly estimated for Prince William Sound herring (Muradian et al., 2017). Additionally, these studies used model estimates from the herring stock assessment as input data, which may produce

unreliable results (Brooks & Deroba, 2015). We avoid these problems while also including more years of data (at least for juvenile walleye pollock), which may also erase previously detected environment-recruitment correlations (Myers, 1998).

2.5.3 As fixed effect versus as latent variable models

For our analysis, most covariates exhibited much worse model selection values as latent variables versus as fixed effects. Incorporating covariates as latent variables as opposed to fixed effects follows recommendations for evaluating environmental covariates in a more statistically rigorous manner to account for covariate uncertainty (CAPAM, 2017). Simulation analyses comparing these two approaches in a maximum likelihood framework indicated similar performance in the quality of results (nominally, our 'as fixed effect' is 'as structure,' and 'as latent variable' is 'as data' in Crone et al., 2019; Schirripa et al., 2009). Our results suggest estimated recruitment or mortality deviations 'fit' to latent variables are generally poor and lead to worse model predictions (i.e. large criteria values) than covariates as fixed effects, especially compared to the null model. We suspect the latent variables act as informative priors because of relatively small standard deviations (σ=0.3 or less), that alter the posterior geometry sampled in MCMC and consequently overweight their influence in estimation. That latent variables should perform worse in our analysis may imply these covariates provide poor information or, that there is greater uncertainty in the latent variables than we assumed. Additional error in latent variables may be estimated, as done for the survey indices in BASA (Muradian et al., 2017), to account for the total uncertainty so that it better represents the measurement, estimation, and process errors (e.g. error in estimating predator abundance as well as uncertainty in spatial overlap between predators and herring, and their dietary preference for herring).

Our comparison of both approaches for incorporating covariates also reveal consequences for stock assessment. For modeling covariates as fixed effects on recruitment in particular, where random effects are included, a substantial amount of recruitment variability could not be explained by any covariate. Since random effects captured a majority of recruitment variation, covariates as fixed effects did not produce different estimates of total recruitment and spawning biomass compared with the other recruitment covariates and the null model. This implies a key benefit to modeling covariates as fixed effects with additional random effects; if the wrong covariate(s) is used, there is no consequence to model estimates. Of course, this issue would be consequential for forecasting, where there is no clear approach on how to best use covariates as fixed effects (and additional random error) or as latent variables if substantial uncertainty remains. Scientists should consider this flexibility in using covariates as either fixed effects or latent variables within stock assessments and explore the consequences of both to model predictions.

2.5.4 Mixed messaging in Bayesian model selection

Model selection criteria may fail under certain conditions and with certain models (Gelman, Hwang, & Vehtari, 2014b; Piironen & Vehtari, 2017; Vehtari & Ojanen, 2012). However, results from simulation analyses using these criteria with more complex population-dynamics models are promising; for example, performance of criteria improved when latent variables/random effects were marginalized out of the likelihood compared to criteria based on conditional likelihoods in one study (Kai & Yokoi, 2019), while in another study, criteria were able to favor models that produce reliable estimates, despite failing to select the true model (Dey, Delampady, & Gopalaswamy, 2019). The additional diagnostics available for PSIS-LOO ($\hat{k}$) in

particular provides insight to the reliability of both model selection criteria and the models themselves. That all models in our analysis had multiple problematic $\hat{k}$ may suggest yet unaddressed misspecification within BASA. Alternatively, considering that many covariates showed likely effects (Figs. 2.2-3), differences in criteria were small amongst various models (including the top and null models; Figs. 2.4-5), and population estimates differed little between the best models (Fig. 2.7), the underlying reason might lie in BASA being a very flexible model. Future Bayesian model selection with BASA, or any fisheries stock assessment, should involve simulation testing the performance of criteria in relation to the various issues raised here (e.g. influential observations in covariate time series, amount of error in covariates as latent variables, etc.).

That we found some consensus amongst criteria for one covariate suggests our approach would be useful for evaluating alternative stock assessment models. Various model configurations are typically explored and presented as a part of a stock assessment for fisheries management. Our results emphasize that it is prudent to use multiple established criteria when comparing models to confirm conclusions of support for any one model. Other specific measures such as retrospective bias (Mohn, 1999) and forecasting error afford different perspectives on prediction that would further inform model selection, though this becomes computationally expensive when considering many Bayesian models.

2.6 CONCLUSION

Our study provided a rigorous evaluation of ecological covariates in Bayesian stock assessment. Our approach checks common issues of similar covariate and model selection analyses including missing covariate observations, covariate errors, and the potential failure of any one criterion when

used on its own. The advent of more computationally efficient software and algorithms, such as NUTS (Monnahan, Branch, Thorson, Stewart, & Szuwalski, 2019), makes the evaluation of multiple Bayesian stock assessment models more tractable. This invites the application of our approach for stock assessments while bringing more intensive Bayesian tools within reach (Hooten & Hobbs, 2015; Piironen & Vehtari, 2017). This is a boon to Bayesian stock assessment given the plausibility of multiple models and the fact that multiple ecological factors exert pressure on fish population dynamics with effects of any one factor becoming apparent only under certain conditions. In such cases, Bayesian model averaging should better represent stock assessment uncertainty (Ianelli, Holsman, Punt, & Aydin, 2016) and can, but not always, further reduce prediction errors (Dormann et al., 2018). Super-ensemble models have also been demonstrated to improve predictions of fish population status (Anderson et al., 2017). Still, more simplistic single covariate models with model selection approach remains nearly ubiquitous in fisheries science. Our extensions of this established approach offer the next step for stock assessment researchers to take when moving to Bayesian multi-model inference.

2.7 ACKNOWLEDGEMENTS

Funding for this work was provided by the Exxon Valdez Oil Trustee Council (EVOSTC). T.A.B. was additionally supported in part by the Richard C. and Lois M. Worthington Endowed Professor in Fisheries Management. We also thank Rich Brenner, Stormy Haught, John Moran, and Jan Straley for sharing pink salmon and whale data for use in this analysis; Scott Pegau and Cole Monnahan for valuable discussions and feedback throughout the development of this study; and Andre Punt, Tim Essington, and Ole Shelton for comments that improved the quality and clarity of the manuscript.

2.8 TABLES

Table 2.1. Summary of covariates individually tested with BASA. Covariates are used to model effects on recruitment, natural mortality, or both. For natural mortality, covariates can be modeled bi-annually with the modeled periods indicated (in first, *b*=1, or second, *b*=2, period). In examining alternative timeframes to check for non-stationarity, some covariates have missing years which are ignored in the model and not filled or interpolated. Most hypotheses (beside NPGO and scale growth) have been previously posited and/or supported and those references are provided. Data sources by agency, reference, or url are also provided.

Hypothesis	Indicator	Used for recruitment?	Used for mortality and for which ages?	Index for half-year mortality (*b*)	Years of available data	Timeframes modeled	References for hypothesis	Source
Cause of epizootics in herring that positively associates with mortality in younger fish	VHSV	No	Yes, 3-4	1-2	1994-2017	1994-2017, 2007-2017	Marty et al. (2003); Hulson et al. (2007)	Alaska Department of Fish & Game
Cause of endemic disease in herring that positively associates with mortality in older fish	*Ichthyophonus hofer*i	No	Yes, 3-8	1-2	1994-2006, 2007-2017	1994-2017, 2007-2017	Marty et al. (2003); Hulson et al. (2007)	Alaska Department of Fish & Game
Oceanic conditions associate (positively or negatively) with adult mortality	Summer upwelling index	Yes	Yes, 3-9+	1	1980-2017	1980-2017, 1994-2017, 2007-2017, 1980-2009	Williams and Quinn (2000); Ward et al. (2017)	NOAA Pacific Fisheries Environmental Laboratory (https://oceanwatch.pfeg.noaa.gov/products/PFELData/upwell/monthly/upindex.mon)
Broad-scale summer climate associates (positively or negatively)	Summer NPGO	Yes	Yes, 3-9+	1	1980-2017	1980-2017, 1994-2017, 2007-2017,	NA	http://www.o3d.org/npgo/; Di Lorenzo et al. (2008)

with adult mortality						1980-2009		
Broad-scale winter climate associates (positively or negatively) with adult mortality OR recruitment	Winter NPGO	No	Yes, 3-9+	2	1980-2017	1980-2017, 1994-2017, 2007-2017, 1980-2009	NA	http://www.o3d.org/npgo/; Di Lorenzo et al. (2008)
Broad-scale summer climate associates (positively or negatively) with adult mortality	Summer PDO	No	Yes, 3-9+	1	1980-2017	1980-2017, 1994-2017, 2007-2017, 1980-2009	Deriso et al. (2008)	http://research.jisao.washington.edu/pdo/
Broad-scale winter climate associates (positively or negatively) with adult mortality OR recruitment	Winter PDO	No	Yes, 3-9+	2	1980-2017	1980-2017, 1994-2017, 2007-2017, 1980-2009	Deriso et al. (2008)	http://research.jisao.washington.edu/pdo/
Salmon prey on adult herring and positively associate with mortality	Total pink salmon run	No	Yes, 3-9+	1	1980-2017	1980-2017, 1994-2017, 2007-2017, 1980-2009	Deriso et al. (2008); Pearson et al. (2012); Sewall et al. (2017)	Rich Brenner (ADF&G)
Flounder prey on adult herring and positively associate with mortality	Gulf of Alaska arrowtooth flounder female spawning biomass	No	Yes, 3-9+	1-2	1980-2017	1980-2017, 1994-2017, 2007-2017, 1980-2009	Spies et al. (2017)	Spies et al. (2017)
Pollock are alternative prey source for herring predators (Stellar sea lion and arrowtooth flounder) and negatively	Gulf of Alaska walleye pollock spawning biomass	No	Yes, 3-9+	1-2	1980-2017	1980-2017, 1994-2017, 2007-2017, 1980-2009	Pearson et al. (2012)	Dorn et al. (2017)

associate with mortality								
Humpbacks prey on herring and positively associate with mortality	Humpback whale abundance	No	Yes, 3-9+	1	1980, 1983-2009	1980-2009	Moran et al. (2018a); Pearson et al. (2012)	Teerlink et al. (2015)
Humpbacks prey on herring and positively associate with mortality	Humpback whale counts	No	Yes, 3-9+	2	2006-2008, 2011-2017	2007-2017	Moran et al. (2018a); Pearson et al. (2012)	Moran & Straley (unpublished data)
Bottom-up forcing on near-shore zooplankton timing and quantity associates with juvenile survival	Freshwater discharge	Yes	No	NA	1980-2013	1980-2017, 1994-2017, 2007-2017, 1980-2009	Ward et al. (2017); Ware and Thomson (2005)	Royer (1982)
Prey quantity positively associates with juvenile survival	Prince William Sound Avg. Zooplankton density	Yes	No	NA	1981-1999	1981-1999	Brown and Norcross (2001)	Brown and Norcross (2001)
Growth during first-year positively correlates with first-year survival	Age 0 scale growth	Yes	No	NA	1980-2013	1980-2013, 1994-2013, 2007-2013, 1980-2009	NA	Haught and Moffitt (2018)
Pollock recruitment success correlates with herring recruitment	Gulf of Alaska walleye pollock age-1 (lagged 1-yr)	Yes	No	NA	1980-2015	1980-2015, 1994-2015, 2007-2015, 1980-2009	Sewall et al. (2017)	Dorn et al. (2017)
Juvenile salmon have antagonistic interaction with herring and negatively	Prince William Sound hatchery juvenile pink salmon	Yes	No	NA	1980-2017	1980-2017, 1994-2017, 2007-2017	Deriso et al. (2008); Pearson et al. (2012); Sewall et al. (2017)	Rich Brenner (ADF&G)

associate with recruitment								
Negative shift in mean recruitment, regardless of cause	1989 regime shift	Yes	No	NA	1980-2017	1980-2017, 1994-2017, 2007-2017	Ward et al. (2017)	NA

2.9 FIGURES

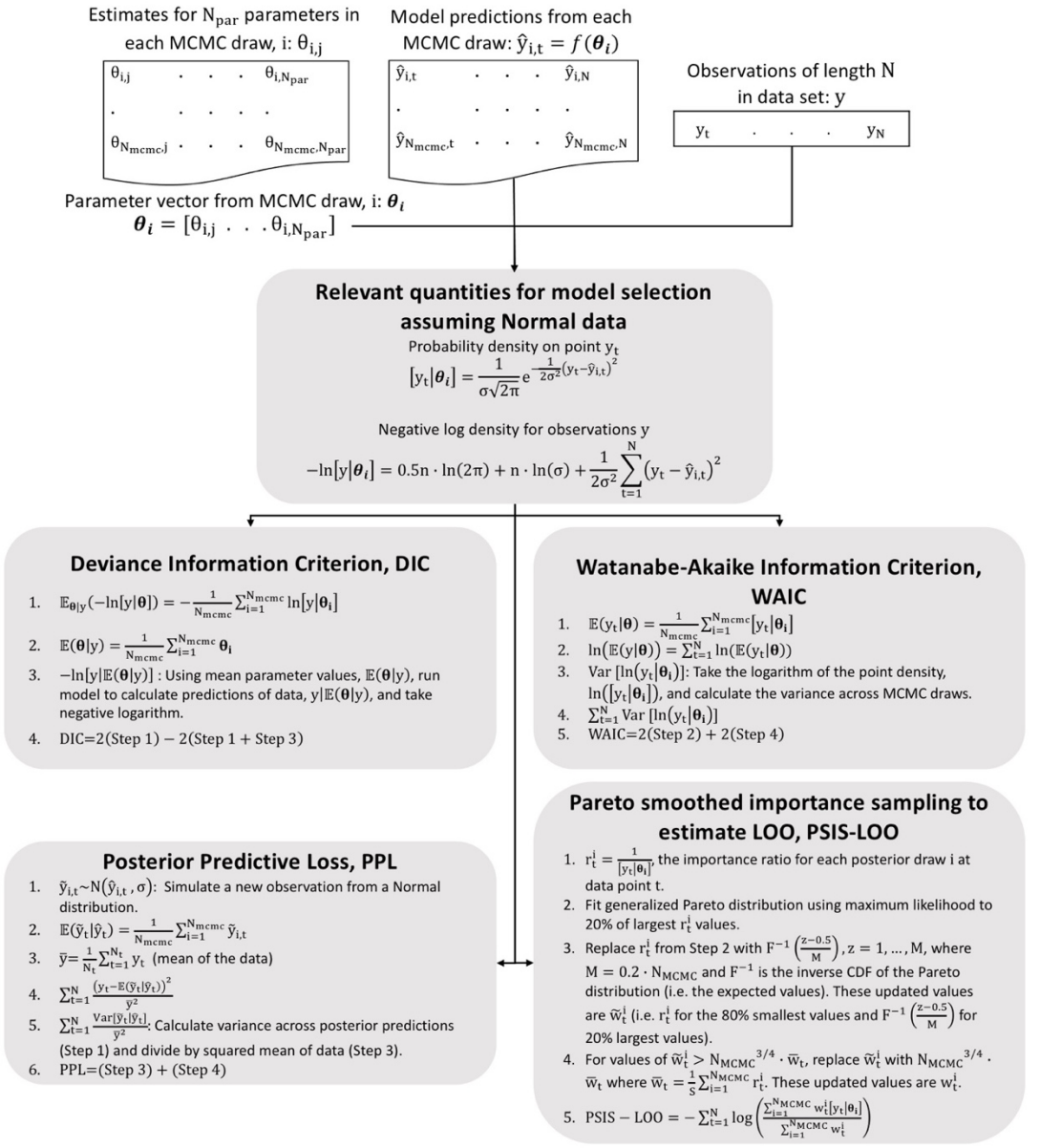

Figure 2.1. Schematic for how Bayesian model selection criteria were calculated in this analysis using a single example data set y with normally-distributed errors. This example data set has *N* total observations as indexed by *t*. Model estimates of the data $\hat{y}_{i,t}$ conditioned on parameter vector $\boldsymbol{\theta}_i$ the ith iteration of a total N_{mcmc} iterations sampled using Markov chain Monte Carlo. Steps for calculating Deviance Information Criterion (DIC), Watanabe Akaike Information Criterion (WAIC), Posterior Predictive Loss (PPL), and Pareto Smoothed Importance Sampled Leave-one-out Cross-validation (PSIS-LOO) are provided as equations that use the log-likelihood or posterior density of the data y.

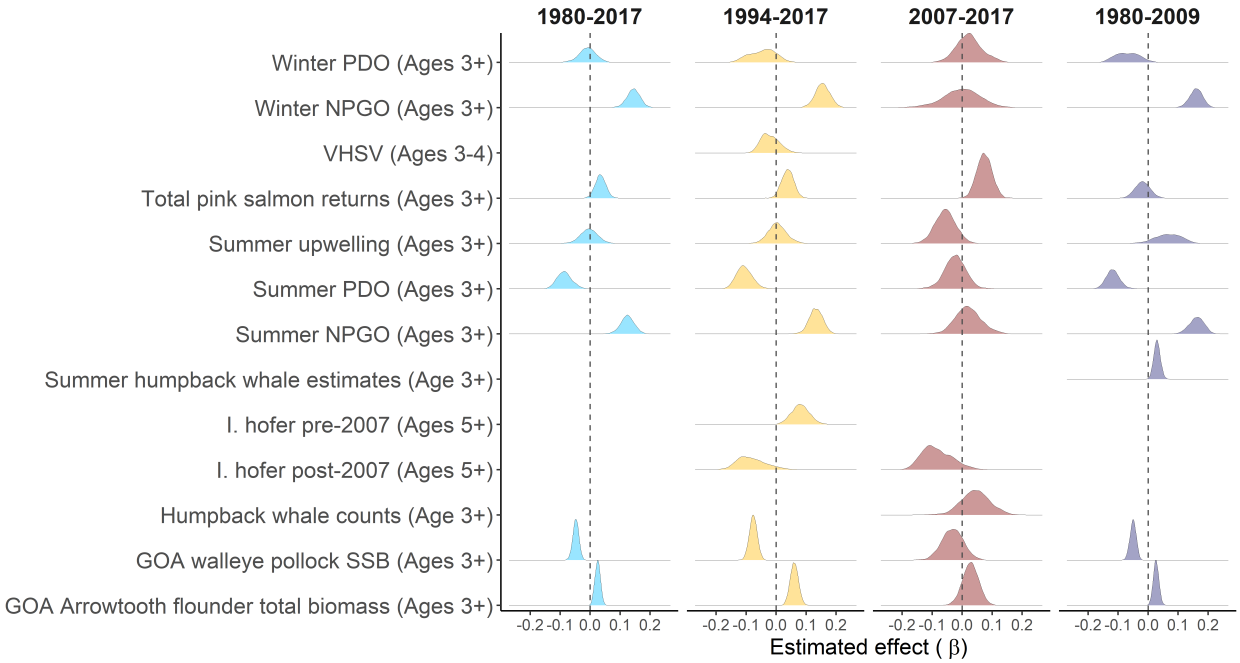

Figure 2.2. Posterior distributions of the estimated effects on natural mortality and for each time frame. A zero effect is denoted by a dashed vertical line. No posterior is shown for VHSV from 2007-2017 because indices were zero in most years except one year (0.0003).

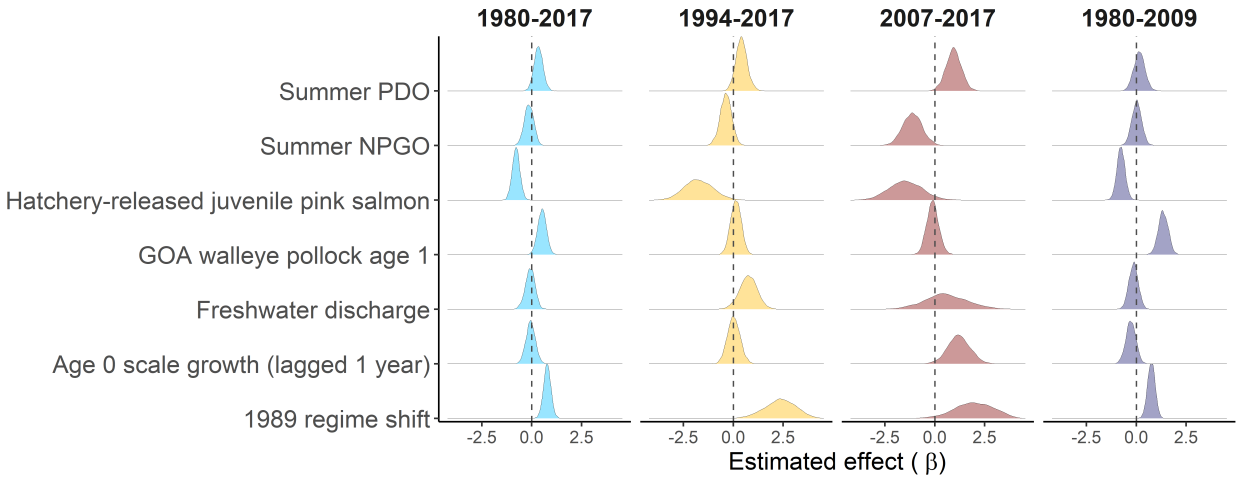

Figure 2.3. Posterior distributions of the estimated effects on recruitment and for each time frame. A zero effect is denoted by a dashed vertical line.

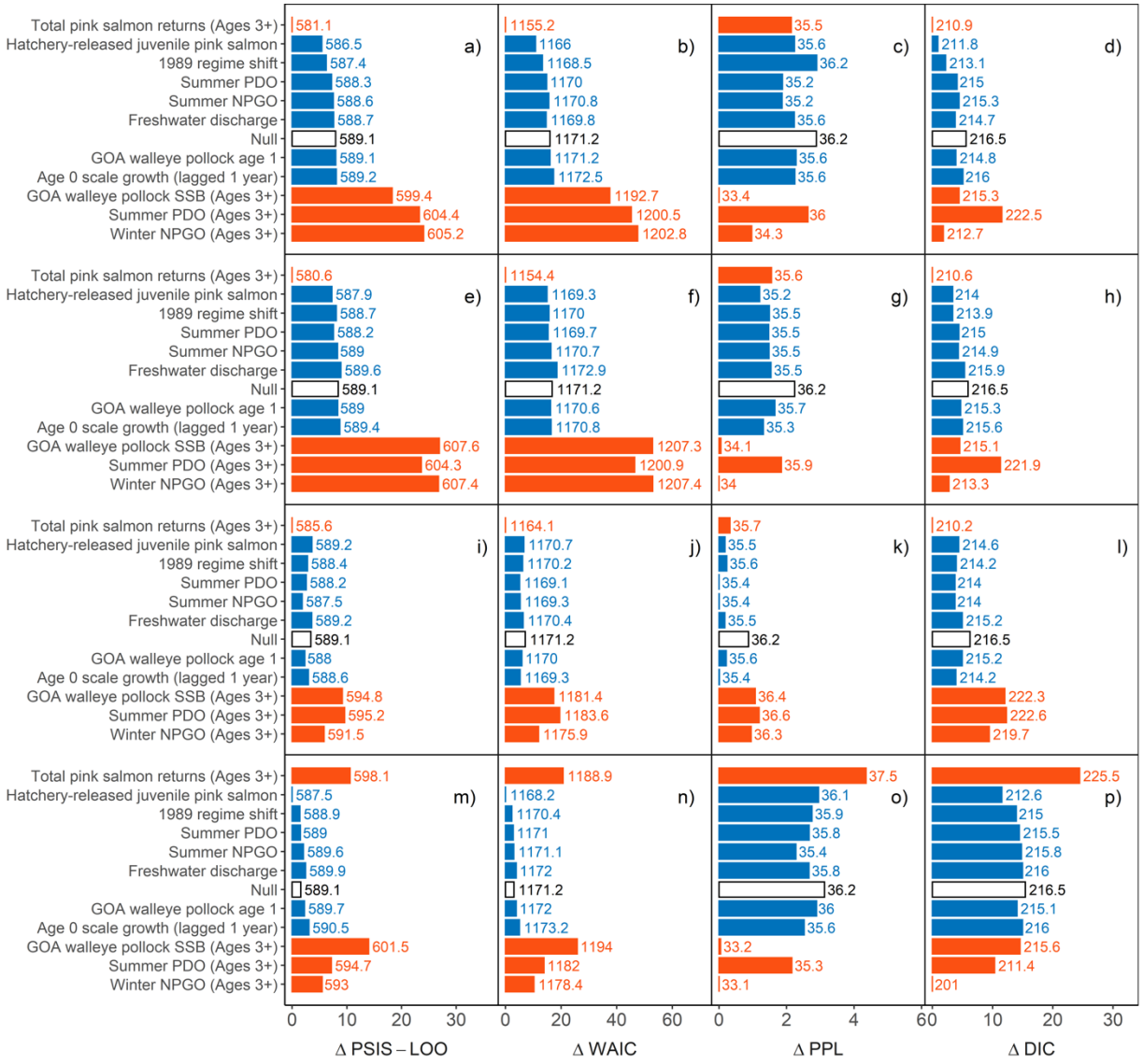

Figure 2.4. Bar charts of model selection values across select covariates as fixed effects model variants of BASA with at least two criteria better than the value of the Null model (empty black bar). Colors indicate the process affected, either natural mortality (red) or recruitment (blue). Each row represents the different time periods modeled: 1980-2017 (a-d), 1994-2017 (e-h), 2007-2017 (i-l), and 1980-2009 (m-p). Each column represents one of the four model selection criteria used (PSIS-LOO, WAIC, PPL, and DIC). Bar lengths measure the difference in the criteria values from the best model (the minimum) in each box. The raw criteria values are labeled next to the bars. The same 12 covariates are shown for all rows and are ordered from the smallest to largest values of ΔPSIS − LOO in plot a).

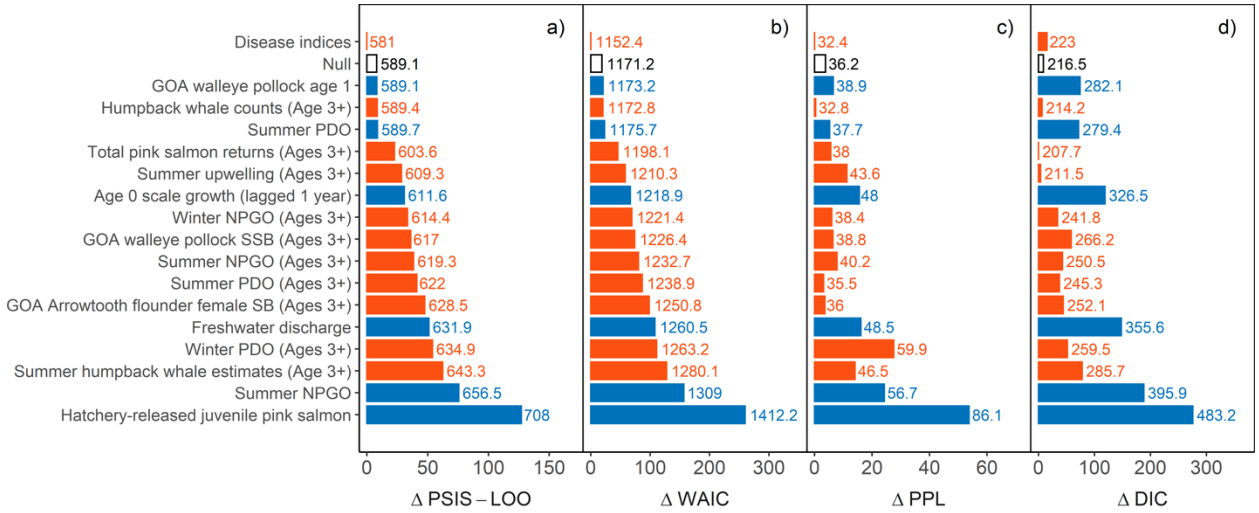

Figure 2.5. Bar charts of model selection values across all covariates as latent variable model variants of BASA. The format is identical to Figure 2.4 (red = natural mortality effect, blue = recruitment effect) and is only shown for one time frame (1980-2017). The ecological covariates are ordered from the smallest to largest values of ΔPSIS − LOO in plot a).

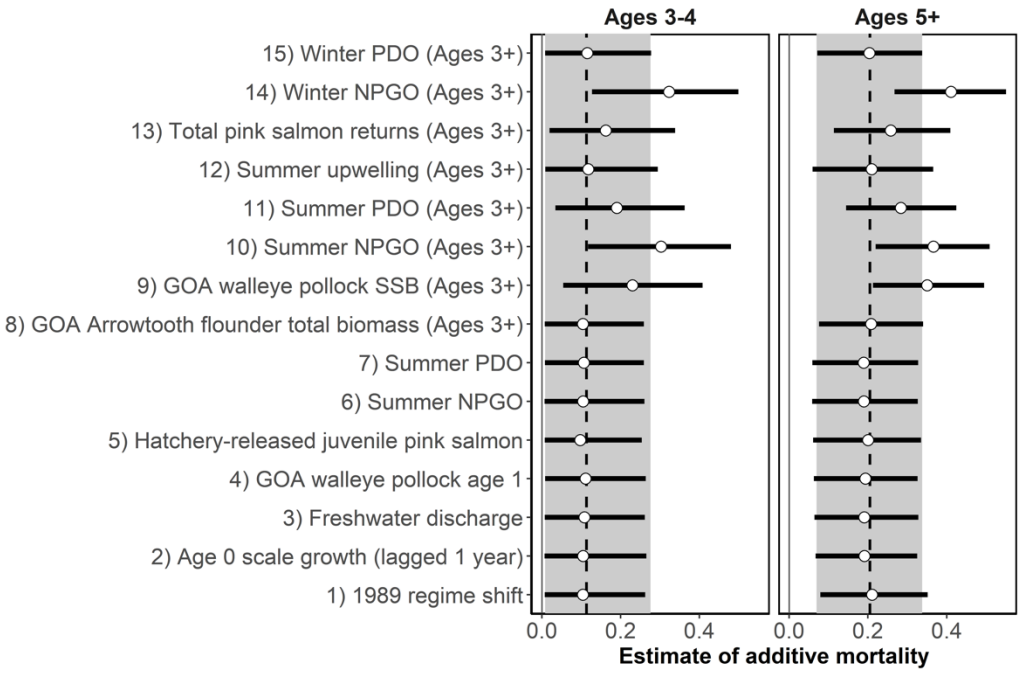

Figure 2.6. Median (empty circle) and 95% credibility intervals (blue lines) of additional mortality in 1993 for two different age groups (Ages 3-4 and Ages 5+). Recruitment (1-7) and natural mortality (8-15) specific effects are shown together with estimates of the null model denoted by the shaded regions (95% interval) and vertical dashed lines (median). If covariates partially explain the decline in biomass in 1993, then we would expect the additional mortality estimates for these covariates to be lower than those of the null model. Estimates are from models using the full covariate time series (1980-2017).

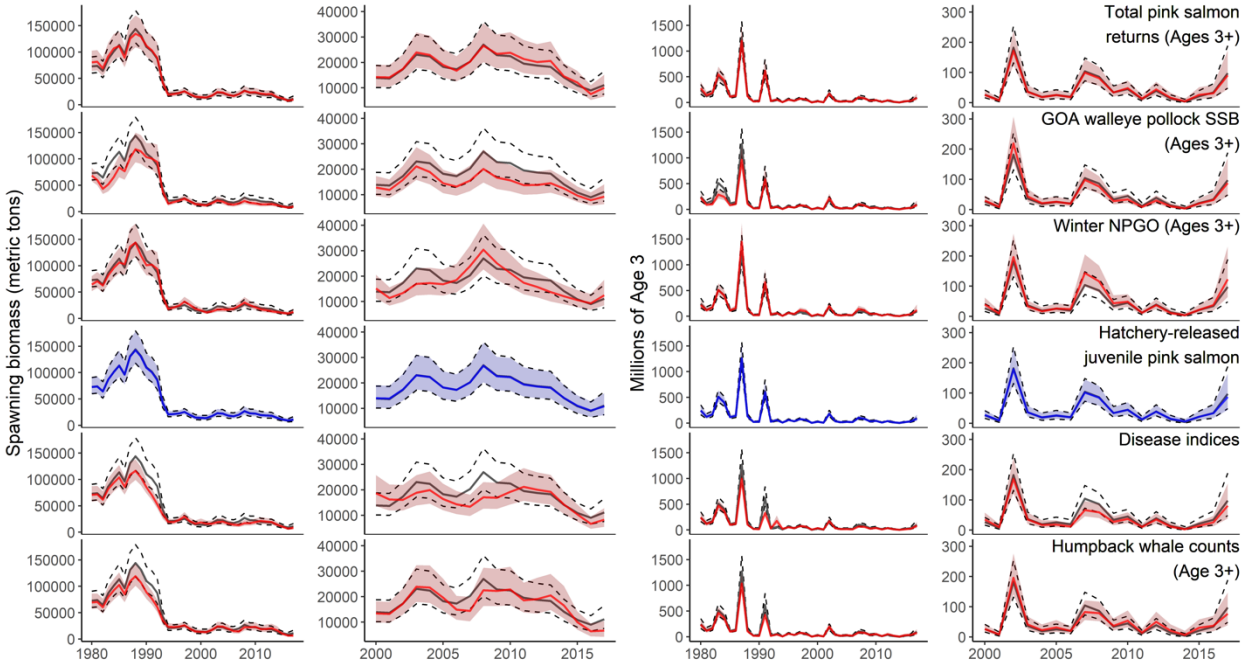

Figure 2.7. Estimates of spawning biomass and recruitment (in millions of age 3 fish) from select models that were the best model in at least one criterion compared to the Null model (dark grey lines). These include top covariates modeled as fixed effects from 1980-2017 (Total pink salmon returns, GOA walleye pollock SSB, and Winter NPGO, and Hatchery-released juvenile pink salmon) and latent variables (Disease indices and Humpback whale counts). Color coding indicates the process affected (red = recruitment, blue = natural mortality). The lines and shaded regions reflect the posterior median and 95% credibility intervals, respectively. The Null model median and uncertainty estimates are shown by the solid and dashed grey lines, respectively. For the hatchery-released juvenile pink salmon model, estimates are virtually an exact match with the Null model because additional random effects are estimated to capture the variability not explained by the covariate. Estimates are shown over the complete time frame (1980-2017) and after 2000 because of the substantial difference in scale of biomass and recruitment dynamics before and after collapse.

Chapter 3. A better way to account for disease in fisheries stock assessment: the potential of seroprevalence data

3.1 ABSTRACT

When attempting to assign disease mortality estimates to fisheries population models, infection and disease surveillance data are often of little value, especially with fast-moving diseases that may go unobserved in populations that are difficult to observe and sample. In the case of viral hemorrhagic septicemia (VHS) in Pacific herring, a more valuable indication of annual disease impacts and future disease potential would involve quantitative assessments of the annual VHS virus exposure history, combined with annual demographic data including abundance, biomass, and age structure assessments. We examine the utility of this approach using a simulation analysis with a novel epidemiological model to simulate population dynamics and seroprevalence data. These data were fitted in a typical integrated catch-at-age model with equations that formally quantify age- and time-varying mortality from seroprevalence. We found that seroprevalence data can indeed inform accurate estimates of infection history and disease-associated mortality. Importantly, even mis-specified models that included seroprevalence data provided accurate estimates of disease-related parameters and the population abundance. In applying this framework to real seroprevalence data of VHSV from Pacific herring in Prince William Sound, we found generally low VHSV infection and disease mortality in recent years, although current seroprevalence data may underrepresent the true immunity. Still, seroprevalence data has the potential to provide valuable information that can improve stock assessments and fisheries management of populations where disease is a major concern and accurate quantification of infection prevalence is not possible.

3.2 INTRODUCTION

Disease can be a major cause of reported mass mortality events in animals that are increasing and frequency, and viral infections cause the largest-magnitude mortality events (Fey et al., 2015). For fish, disease-associated mass mortality events have not only increased in magnitude, but also in frequency over time (Fey et al., 2015). Diseases are caused by variety of pathogens found in many different species, including microparasites in salmon species worldwide (Miller et al., 2014), herpesvirus in Australian pilchard (*Sardinops sagax*) (Alexander, Michael, & Brian, 2003), and *Ichthyophonus hoferi* in multiple pelagic fish species across the Northern Hemisphere (e.g. Pacific and Atlantic herring, Yellowtail flounder, and American shad; reviewed in Burge et al., 2014; McVicar, 2011). Disease outbreaks can unexpectedly reduce fish harvests and threaten the livelihood of fishers.

Very few stock assessments used by fisheries management account for disease even when there are chronic infections or epizootics in exploited stocks (Hoenig et al., 2017). One example is the use of an infection prevalence index (the observed proportion that exhibit visible signs or test positive for an active infection of a pathogen) of mycobacteriosis with tagging data in the stock assessment for Chesapeake Bay striped bass (*Morone saxatilis*) (Vogelbein et al., 2012). An in-depth analysis of these data with a multi-state mark recapture model revealed survival of diseased fish was 25-75% depending on water temperature and infection severity, implying substantial impacts to the exploitable population (Groner et al., 2018). Otherwise, simple logistic regression models have been used outside stock assessments to estimate disease-associated mortality as done for snow crab (*Chionoecetes opilio*) in Newfoundland, Canada and American lobster (*Homarus americanus*) in New England (Hoenig et al., 2017). These examples motivate the development of

alternative approaches to data-intensive tagging studies, but just as rigorous in accounting for disease dynamics directly within stock assessment models.

The Prince William Sound Pacific herring (*Clupea pallasii*) stock assessment also incorporates disease information (Marty et al., 2010; Muradian et al., 2017). Disease mortality from viral hemorrhagic septicemia virus (VHSV) and other pathogens remains a leading hypothesis for the failed recovery of the population after its decline in the early 1990s (Marty et al., 2010; Marty et al., 2003). VHS afflicts many herring populations and other species in the northeast Pacific Ocean (e.g. hake and walleye pollock; Meyers & Winton, 1995; Skall, Olesen, & Mellergaard, 2005), causing epizootics and fish kills (Garver et al., 2013). Annual VHSV prevalence in the population is assessed by cell culture from subsamples of individuals collected during annual survey cruises (pers. comm. P. Hershberger). Infection prevalence indices from these samples are directly input into the stock assessment as a linear fixed effect on natural mortality from 1994 (the first year prevalence samples were collected) to the present (Muradian et al., 2017). However, VHSV prevalence only measures active infection by VHSV and does not count VHSV-associated mortality or survivors of the disease. Therefore, VHSV infection prevalence data have limited utility for estimating annual mortality rates in wild populations that cannot be monitored continuously throughout the year (Elston & Meyers, 2009).

While past tools could estimate current prevalence of infections, surveillance tools have been developed, optimized, and validated to evaluate evidence of prior exposures to VHSV. Serological assays are one of the most promising of these tools and involve testing for antibodies specific to VHSV (Hart et al., 2017; Hershberger et al., 2015; Wilson et al., 2014). Nonlethal age-specific serological tests, such as the indirect enzyme-linked immunosorbent assay (ELISA) and plaque neutralization test (PNT) have been successfully demonstrated for VHSV in Pacific herring

and other freshwater species, although the PNT has been recently optimized to provide more sensitive and repeatable results in Pacific herring (Hart et al., 2017; Wilson et al., 2014). Serological tests can provide a synoptic view of prior exposure to a virus by measuring seroprevalence—the proportion of fish demonstrating detectable levels of antibodies to a specific pathogen. In the case of VHS virus, seroprevalence data may be used to estimate past mortality and forecast future outbreaks because survivors of prior exposure demonstrate long term immunity against future outbreaks of the disease (Hershberger et al., 2015). When included in age-structured fisheries stock assessments, seroprevalence data may enable estimation of annual mortality resulting from unobservable VHSV outbreaks if age-specific seroprevalence data can be obtained from one of the methods listed above.

Here, we create the first application of using age-specific seroprevalence data in fisheries stock assessment models. We (1) develop an epidemiological model with annual population dynamics to simulate 'realistic' outbreaks, (2) develop an age-structured stock assessment model that incorporates seroprevalence data to directly estimate infection rates and disease-associated mortality, (3) evaluate the accuracy of stock assessment model estimates using simulated seroprevalence data from the epidemiological model, and (4) apply the equations of the new stock assessment model along with real seroprevalence data to estimate VHSV impacts in Prince William Sound herring in between 2008-2019.

3.3 METHODS

Our simulation framework includes an age-structured susceptible-infected-carrier (SIC) operating model that simulated epidemic and population dynamics at daily time steps, and an estimation model that estimated parameters describing interannual outbreaks (Fig. 3.1). We used VHSV in

Pacific herring as a case study for structuring both the operating and estimation models following principles of the disease determined from observations and theory (Anderson & May, 1979; Hershberger et al., 2015). The operating model has a daily time step to simulate infections of fish of different ages and disease stages; and an annual time step to simulate fish population dynamics. Outputs of this operating model included recruitment, disease and natural mortality, infection prevalence, and seroprevalence. Data mimicking those that would be collected during annual scientific surveys of fish populations is subsampled from the operating model output, and included relative fish abundance, age composition, and seroprevalence by age. These data were analyzed using an estimation model—an age-structured stock assessment model of the type commonly used in fisheries management (e.g. Muradian et al. 2017). The estimation model keeps track of annual age-structure, but ignores daily disease transmission; instead, annual disease infection and recovery were estimated directly. Different parameterizations of the operating and estimation models were explored to test whether disease-related parameters could be estimated under different scenarios. Test scenarios included the effects of the quality of data collection through surveys, lagged transmission due to gradual mixing of susceptible herring (e.g. partial recruitment), and ignoring time-varying background natural mortality rates. Performance of the estimation models was evaluated by comparing estimates of annual disease mortality and prevalence to the "true" values output by the operating models. Finally, we included real seroprevalence data of VHSV within the stock assessment of Pacific herring in Prince William Sound, Alaska to test inference on past population-wide infection and disease-associated mortality.

3.3.1 Operating model overview

We adapted a modeling framework similar to Briggs, Vredenburg, Knapp, and Rachowicz (2005), in which within-year disease dynamics were modeled distinctly from between-year population dynamics in an animal population. This model ran a sequence of two events within each year: 1) a fixed period of transmission, recovery, and mortality transitions occurring on daily time-steps, and 2) a single time-step that projects the total population dynamics to the start of the next year (Fig. 3.1). This sequence reflects a life history in which there is a fixed period conducive to disease transmission, for example during annual herring spawning aggregation events. We did not incorporate fishing into this framework, but it is straightforward to adjust the equations to include catches.

3.3.2 Operating model: the fixed period of daily disease, recovery, and mortality transitions

For the period of disease transmission, we adapted a discrete-time model commonly used in epidemiological studies (e.g. Klepac & Caswell, 2011; Metcalf et al., 2012; Winter et al., 2018) to project daily numbers of fish in each of eight age classes (age 0 to '7+', where '7+' is a plus group) and three disease stages: susceptible ($S_{d,a}$), infected ($I_{d,a}$), and carrier ($C_{d,a}$), where subscript d is day and a is fish age in years. This is otherwise known as the SIC model, which is similar to the classical susceptible-infected-recovered (SIR) model (Anderson & May, 1979). The 'recovered' stage is replaced by 'carrier' to account for surviving individuals acquiring immunity, but continuing to host the virus and propagate virions at low levels (Hershberger et al., 2010). Thus, both infected and carrier herring can contribute to transmission.

Not all susceptible individuals are exposed to transmission because they do not overlap in time and space with infected or carrier individuals. For example, fish in the youngest age groups.

We therefore modeled the proportion of susceptible fish that mix with the reservoir population (the population within which the pathogen replicates and infection is transmitted; Haydon, Cleaveland, Taylor, & Laurenson, 2002):

$$S_{y,d=1,a} = v_a T_{y,d=1,a} \quad (1)$$

where $S_{y,d=1,a}$ is the susceptible numbers-at-age in the reservoir population (herein reservoir susceptibles) on day 1, v_a is the fraction of the total susceptible numbers mixing with the reservoir population, and $T_{y,d=1,a}$ is the total susceptible numbers-at-age on day 1 (herein total susceptibles). All fish in the reservoir population are assumed to be homogenously mixed, and non-reservoir susceptible fish do not experience transmission. The quantity v_a is specified by a logistic curve that accounts for the timing of when young fish mix with the reservoir population and first become exposed to natural transmission:

$$v_a = \frac{1}{1 + \exp\left[\frac{-ln(19)\,(a - a_{50}^{mix})}{(a_{95}^{mix} - a_{50}^{mix})}\right]} \quad (2)$$

where a_{50}^{mix} and a_{95}^{mix} are the ages at which 50% and 95% mix with the reservoir susceptibles, respectively. All herring are susceptible to VHSV at birth (Hershberger et al., 2007) and may become sympatric with the reservoir population aggregations after metamorphosis (Hershberger et al., 2015). However, transmission may also occur when age 3 and 4 herring recruit into spawning aggregations for the first time (Hershberger, Kocan, Elder, Meyers, & Winton, 1999). Incorporating $S_{d,a}$ allows for either immediate or delayed pathogen exposure to transmission, which we have explored (see 'Simulation scenarios').

The reservoir numbers-at-age-and-stage on day d, were coded using a single vector for each possible age and stage combination ($\boldsymbol{n}(d)$):

$$\boldsymbol{n}(d) = \left(S_{d,0}, I_{d,0}, C_{d,0}, S_{d,1}, I_{d,1}, C_{d,1,}, \ldots, C_{d,n_a,}\right) \quad (3)$$

To project reservoir stage-specific numbers of a single age group forward one day $\left(S_{d+1,a}, I_{d+1,a}, C_{d+1,a}\right)$, the current day's stage specific number vector was multiplied by an age and day-specific disease transition matrix ($\mathbf{T}_{d,a}$) and background natural mortality (M):

$$\left(S_{d+1,a}, I_{d+1,a}, C_{d+1,a}\right)^T = \left(S_{d,a}, I_{d,a}, C_{d,a}\right)^T \cdot \mathbf{T}_{d,a} \cdot \exp[-M_d] \quad (4)$$

Background natural mortality is assumed constant (M_d = 6.85 x 10^4 d^{-1}, which equates to 0.25 yr^{-1}) for all ages, and $\left(S_{d,a}, I_{d,a}, C_{d,a}\right)^T$ is a transposed vector. The non-reservoir susceptible numbers $(1 - v_a T_{y,d=1,a})$ were not included in Eq. 3-4 because these numbers do not factor into transmission and were projected outside of disease transitions (see Eq. 10). Natural mortality can be constant in all simulated years or can be lognormal over time about an expected value M (see 'Simulation scenarios'). Descriptions of all indices, variables, and parameters from the operating and estimation models are given in Appendix C (Table C1).

The transition matrix $\mathbf{T}_{d,a}$ contains the transition probabilities from one disease stage (in columns) to another (rows) with the order of stages being reservoir susceptible, infected, and carrier:

$$\mathbf{T}_{d,a} = \begin{bmatrix} 1 - \beta_{I,d,a} - \beta_{C,d,a} & 0 & 0 \\ \beta_{I,d,a} + \beta_{C,d,a} & 1 - \alpha - \gamma & 0 \\ 0 & \gamma & 1 \end{bmatrix} \quad (5)$$

where $\beta_{I,d,a}$ is the probability of infection from actively infected ($S_{d,a}$) fish; $\beta_{C,d,a}$ is the probability of infection from carrier ($C_{d,a}$) fish; γ is the probability of recovering from infection; and α is the mortality rate from disease. Dead fish were ignored in transitions because they do not factor into any other equations of this model. The recovery and disease mortality rates used in the simulations were calculated from laboratory experiments infecting Pacific herring with VHSV (Table C1;

Hershberger et al., 2013). For the base model, we assumed that recovery and disease mortality rates are independent of age, day, and year, but it is straightforward to generalize the model to allow time (day and year) variation. Annually-varying rates were explored as sensitivities amongst our simulation scenarios.

Computing the probabilities of infection ($\beta_{I,d,a}$ and $\beta_{C,d,a}$) was more complicated and based on the function below often called the Force of Infection. The Force of Infection acting on age group a is the sum of daily transmission rates from each individual age group and the daily rate at which reservoir susceptible fish become infected:

$$\beta_{I,d,a} = 1 - \exp\left[-\sum\nolimits_j \mu_I I_{d,j} \frac{S_{d,j}}{\sum \boldsymbol{n}(d)}\right] \quad (6)$$

$$\beta_{C,d,a} = 1 - \exp\left[-\sum\nolimits_j \mu_C C_{d,j} \frac{S_{d,j}}{\sum \boldsymbol{n}(d)}\right] \quad (7)$$

where μ_I and μ_C are the mean per capita rates of viral transmission from one infected or one carrier individual, respectively, at age j to all 'contacted' reservoir susceptible fish at age a. These equations assume frequency-dependent transmission, i.e. that transmission does not increase with total population size. For schooling fish like herring this assumption is reasonable since they form more schools and occupy greater area while not increasing in density as numbers increase (en sensu MacCall, 1990), resulting in the number and proximity of individual interactions remaining roughly similar (Murray, 2009).

One alternative to frequency-dependent transmission is density-dependent transmission, where contacts increase with numbers and scale with the absolute number of $S_{d,a}$. This can be modeled by inserting $I_{d,a} \cdot S_{d,a}$ instead of $I_{d,a} \frac{S_{d,a}}{\sum \boldsymbol{n}(d)}$, in Eq. 6 and similarly for the carrier specific numbers in Eq. 7. In our simulations, incorporating density-dependent transmission resulted in

highly unstable disease dynamics where infections rapidly cycled out and the chain-of-transmission ceased, which is another reason why we used frequency-dependent transmission.

Only μ_I and μ_C were specified to fully define $\beta_{I,d,a}$ and $\beta_{C,d,a}$. For simplicity, μ_I and μ_C were assumed constant across ages and years, but with $\mu_C \ll \mu_I$. Reasonable values of μ_I and μ_C were found using projections of the operating model that resulted in irregular outbreaks across years, but outbreaks that were potentially rapid and large when they occurred, as observed in herring populations (Hershberger et al., 2015). In addition, the simulations allowed for transmission to be maintained covertly (i.e. very low infection prevalence rates in some years) by carrier fish. These conditions were met with μ_I = 0.01 day^{-1} and μ_C = 0.000001 day^{-1}.

For the VHSV-herring case study, we chose the length of the fixed period for disease transitions to be 120 days, similar to the period of mass schooling in herring populations (Hay et al., 2001). This fixed period is long enough for rapid epizootics to occur as VHSV epizootics may start and subside within a matter of days (Hershberger et al., 2015). Longer fixed periods (up to 365 days) had negligible impacts on the simulated disease and population dynamics.

At the end of the fixed period (d_{end} = 120), all infected individuals were assumed to recover or die without causing any further transmission in the intervening time between the current year and the next year, resulting in this transition matrix:

$$\mathbf{T}_{d_{end},a} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \gamma/(\alpha+\gamma) & 1 \end{bmatrix} \quad (8)$$

This is justified because most epizootics finish before the 120th day, and more importantly, this is unlikely to occur in nature due to changes in schooling behavior and herd immunity. When the fixed period was extended to the entire year (d_{end} = 365), this resulted in identical annual

probabilities of infection, disease mortality, and recovery but the model took longer to run. Eq. 8 was used to project stage-specific numbers one more day:

$$\left(S_{y,d_{end}+1,a}, I_{y,d_{end}+1,a}, C_{y,d_{end}+1,a}\right)^T = \left(S_{d_{end},a}, I_{d_{end},a}, C_{d_{end},a}\right)^T \cdot \mathbf{T}_{d_{end},a} \quad (9)$$

which was used to calculated to calculate the starting stage structure of the next period of transitions.

3.3.3 Operating model: the single time-step projecting numbers to the next year

The total susceptible fish at the start of the next year ($T_{y+1,d=1,a}$) included all new fish to the population, which was modeled as a stationary mean with serially-correlated deviates (i.e. independent of parental abundance, as stock-recruitment relationships are seldom found in forage fish; Szuwalski et al., 2019; Szuwalski et al., 2015), and the remaining susceptible fish from the current year adjusted for removals from natural mortality that occurred since the 120th day of the previous year:

$$T_{y+1,d=1,a+1} = \begin{cases} \bar{R}\exp\left[\rho\varepsilon_{\mathrm{R},y} + \sqrt{1-\rho^2}\varepsilon_{\mathrm{R},y+1} - 0.5\sigma_{\mathrm{R}}^2\right] & if\ a = 0 \\ (1-v_a)T_{y,d=1,a}\exp[-M_d \cdot 365] + S_{y,d_{end}+1,a}\exp[-M_d \cdot \Delta t] & if\ 0 < a < n_a \end{cases} \quad (10)$$

where $\bar{R}$ is average recruitment in numbers of age 0 fish (millions); ρ is autocorrelation in annual recruitment; $\varepsilon_{\mathrm{R},y}$ are the log-normal annual deviations in recruitment in year *y*; σ_R is the standard deviation of the annual recruitment deviations; and Δt is equal to $365 - d_{end}$, the length of time between the end of the current season and the start of the next.

The reservoir susceptible numbers in the next year ($S_{y+1,d=1,a+1}$) was then calculated from Eq. 1. The number of infected fish was zero in the next year:

$$I_{y+1,d=1,a+1} = 0 \tag{11}$$

There were zero carriers among the new recruits and among older fish the number of carriers ($C_{y+1,d=1,a+1}$) was the number that survived natural mortality from the current year:

$$C_{y+1,d=1,a+1} = \begin{cases} 0 & if\ a = 0 \\ C_{y,d_{end}+1,a}\exp[-M_d \cdot \Delta t] & if\ 0 < a < n_a \end{cases} \tag{12}$$

Stage-specific numbers in the plus group next year were adjusted for the remaining numbers of the plus-age group from the current year (Table C2):

$$\begin{aligned} T_{y+1,d=1,n_a} = \exp[-M_d \cdot 365]\left[\left(1 - v_{n_a-1}\right)T_{y,d=1,n_a-1} + \left(1 - v_{n_a}\right)T_{y,d=1,n_a}\right] \\ + \exp[-M_d \cdot \Delta t]\left[S_{y,d_{end}+1,n_a-1} + S_{y,d_{end}+1,n_a}\right] \end{aligned} \tag{13}$$

$$I_{y+1,d=1,n_a} = 0 \tag{14}$$

$$C_{y+1,d=1,n_a} = \exp[-M_d \cdot \Delta t]\left(C_{y,d_{end}+1,n_a-1} + C_{y,d_{end}+1,n_a}\right) \tag{15}$$

The total numbers-at-age ($N_{y,a}$) is the sum of stage-specific numbers (Eq. 10-15) in each age group.

The age structure in the first year of the operating model was initialized at equilibrium:

$$N_{y=1,a=0} = \bar{R} \tag{16}$$

$$N_{y=1,a+1} = N_{y=1,a}\exp[-M_d \cdot 365] \quad if\ 0 < a < n_a \tag{17}$$

$$N_{y=1,n_a} = N_{y=1,n_a-1}\left(\frac{\exp[-M_d \cdot 365]}{1 - \exp[-M_d \cdot 365]}\right) \tag{18}$$

where $N_{y=1,a}$ is in millions of fish. Each age group was assumed to start with 90% total susceptible, 0% infected, and 10% carriers. The operating model was ran for 50 years to cycle out initial value effects in every individual simulation, and the final 50 years were used for data generation and parameter estimation. The operating model and analyses were coded in R (R Core Team, 2020).

3.3.4 Data simulation

We generated time series of abundance, age-structure, and seroprevalence data from the operating model for input into the estimation model used for stock assessment. First, spawning biomass (B_y) was calculated from:

$$\hat{B}_y = \sum_{a=0}^{n_a} s_a N_{y,a} w_a \tag{19}$$

where s_a is age-specific selectivity of the survey and w_a is weight-at-age (constant across time). For herring and other forage fishes that are sampled during or near spawning activities, s_a is assumed to reflect the maturity schedule (i.e., equal to proportions mature-at-age) and was calculated from a logistic function:

$$s_a = \frac{1}{1 + \exp\left[\frac{-ln(19)\,(a - a_{50}^{survey})}{(a_{95}^{survey} - a_{50}^{survey})}\right]} \tag{20}$$

where a_{50}^{survey} and a_{95}^{survey} are the ages at which 50% and 95% are vulnerable.

Spawning biomass from Eq. 19 was interchangeable with the biomass that was available to surveys which only capture a portion of B_y and have observation error:

$$B_y = qB_y \exp\left[\delta_y - 0.5\sigma_B^2\right] \tag{21}$$

In Eq. 21, q is the survey scalar, and δ_y is annual lognormal error in survey biomass indices where $\delta_y \sim N(0, \sigma_B)$. Survey age-compositions ($\Theta_{y,a}$) were generated from a multinomial with mean proportions obtained from:

$$\Theta_{y,a} = \frac{s_a N_{y,a}}{\sum_{a=0}^{n_a} s_a N_{y,a}} \tag{22}$$

To generate annual seroprevalence values, age-specific compositions of immune fish ($\widetilde{\Omega}_{y,a}$) were first calculated from:

$$\tilde{\Omega}_{y,a} = \frac{C_{y,d=1,a}}{N_{y,a}} \tag{23}$$

which is the proportion of carrier fish in the total population at start of the current year (Eq. 12). Age-specific seroprevalence, or the number of fish with antibodies ($A^{+}_{y,a}$) was then assumed to be assayed from the same samples as the age composition survey (i.e. same survey selectivity s_a):

$$A^{+}_{y,a} = s_a \tilde{\Omega}_{y,a} N_{y,a} \tag{24}$$

Since fish that are not immune were included in these samples, total numbers of fish testing negative for antibodies were calculated as:

$$A^{-}_{y,a} = s_a \left(1 - \tilde{\Omega}_{y,a}\right) N_{y,a} \tag{25}$$

To represent all mutually exclusive outcomes of the sampled seroprevalence, numbers of fish with and without antibodies in each age were combined:

$$\boldsymbol{A_{y,p}} = \left\{A^{+}_{y,a=0}, A^{-}_{y,a=0}, A^{+}_{y,a=1}, A^{-}_{y,a=1}, \ldots, A^{+}_{y,a=n_a}, A^{-}_{y,a=n_a}\right\} \tag{26}$$

Mean proportions of each element of this vector (i.e. $\frac{\boldsymbol{A_{y,p}}}{\sum_{p=1}^{2 \cdot n_a} \boldsymbol{A_{y,p}}}$) were then used to generate multinomial samples used as seroprevalence data. The multinomial assumption meant samples were informed by the age structure as well as the seroprevalence of each age group. Multinomial sample sizes for survey age-composition and seroprevalence data were the same in all years.

All simulated data were generated from spawning numbers-at-age immediately before transmission (i.e. on day 1 of 120), after accounting for survey selectivity.

3.3.5 Estimation model

The estimation model was based on a standard fisheries stock assessment model with an annual time step. It mirrored the population dynamics of the operating model, but simplified disease dynamics into annual groups ignoring daily stage transitions (Fig. 3.1B). Instead, the model

calculated the proportion of each age class that evade infection each year $(1 - v_a)E_{y,a}\omega_y$, where v_a is the age-specific mixing of susceptible fish with the reservoir population, $E_{y,a}$ is the proportion of fish that were susceptible prior to infection, and ω_y is the year-specific proportion of susceptible fish that became infected in the current year. The model also estimated λ, the proportion of infected fish that recover and become carriers. Together, the proportion of fish that evade or survive infection is $1 - v_a E_{y,a}\omega_y(1 - \lambda)$ (see Appendix B for derivation). Inputting this into numbers-at-age equation gave:

$$N_{y+1,a+1} = \exp[-M_y][1 - v_a E_{y,a}\omega_y(1 - \lambda)]N_{y,a} \quad 0 \leq a < n_a \tag{27}$$

where M is annual background mortality (M_y=0.25 yr^{-1} in all years). The estimable disease-related parameters were the annual proportion of susceptible fish that become infected, ω_y, and the time-invariant λ for the proportion of infected fish that recover. Both v_a and $E_{y,a}$ are derived (Eqs. 2 and 29, respectively). It was assumed that no fish remain infected between years (Eq. 27), which is reasonable for VHSV in herring that causes rapid outbreaks and often shows no infection prevalence in samples from wild populations (Hershberger et al., 2015).

Age-specific immunity ($\tilde{\Omega}_{y+1,a+1}$) in the next year is the proportion of all fish (susceptible or recovered) that are immune because they contracted and recovered from disease in year y or further in the past (see Appendix B for derivation):

$$\tilde{\Omega}_{y+1,a+1} = \begin{cases} 0 \quad if\ a = 0 \\ \dfrac{\tilde{\Omega}_{y,a} + v_a E_{y,a}\omega_y\lambda}{1 - v_a E_{y,a}\omega_y(1 - \lambda)} \quad if\ 0 \leq a < n_a \end{cases} \tag{28}$$

These immune proportions were used to calculate the proportion that remain susceptible in each cohort in the current year ($E_{\mathrm{y,a}}$):

$$E_{y,a} = \begin{cases} 1 & if\ a = 0 \\ 1 - \tilde{\Omega}_{y,a} & otherwise \end{cases} \tag{29}$$

Separate equations for the plus-age group ($a = n_a$) adjusted for the existing numbers in n_a from year y (Table C2).

Additional parameters estimated by the model were the parameters of the logistic function for survey selectivity (a_{50}^{survey} and a_{95}^{survey}), parameters of the logistic function for the proportions of young fish mixing with the reservoir population (a_{50}^{mix} and a_{95}^{mix}), lognormal annual recruitment deviations in recruitment (ε_y), mean recruitment ($\bar{R}$), and the lognormal standard deviation in recruitment (σ_R). The values of these parameters and their parameterizations (Eqs. 2 and 20) matched those used in the operating model, except that recruitment in the first year was estimated:

$$N_{y+1,a=0} = \begin{cases} \bar{R}\exp\left[\varepsilon_y - 0.5\sigma_R^2\right]\ if\ y = 1 \\ \bar{R}\exp\left[\rho\varepsilon_y + \sqrt{1-\rho^2}\varepsilon_{y+1} - 0.5\sigma_R^2\right]\ if y > 1 \end{cases} \quad \varepsilon_y \sim N(0, \sigma_R^2) \tag{30}$$

Numbers-at-age in the first year of observations were modeled with estimated log-normal deviations (ε_y) around an estimated mean abundance ($\bar{N}$) and σ_R:

$$N_{y=1,a} = \bar{N}\exp\left[\varepsilon_y - 0.5\sigma_R^2\right]\ where\ \varepsilon_y \sim N(0, \sigma_R^2) \tag{31}$$

Annual recruitment and initial numbers-at-age deviations were treated as random effects that were integrated out in Template Model Builder (TMB) (Kristensen, Nielsen, Berg, Skaug, & Bell, 2016) using the Laplace approximation, which allowed estimation of their respective standard deviations.

The log-likelihood functions for the biomass and age composition survey data were derived from the same distributions used to simulate these data from the operating model. For the age-specific seroprevalence data, a binomial likelihood was used. The multinomial distribution used to simulate seroprevalence is avoided so that age structure information in the seroprevalence samples (i.e. the proportions of the total sample in each age) is not double counted in model fitting.

The negative log-likelihood component for the seroprevalence data sums over the individual binomial likelihoods for each seroprevalence value for each age in each year (Table 3.1):

$$\ell(\hat{A}) = \sum_{y}\sum_{a=0}^{n_a}\left[-k_{y,a}ln\left(\hat{p}_{y,a}\right) - \left(N_{y,a}^{sero} - k_{y,a}\right)ln\left(1-\hat{p}_{y,a}\right)\right] \quad (32)$$

where $k_{y,a}$ is the observed number of fish testing positive for antibodies in age a and year y, $\hat{p}_{y,a}$ is the predicted proportion of seropositive fish in age a and year y ($\hat{p}_{y,a} = \frac{\hat{A}_{y,a}^{+}}{\hat{A}_{y,a}^{+}+\hat{A}_{y,a}^{-}}$ from Eqs. 24-25), and $N_{y,a}^{sero}$ is the total number sampled in each age a and year y.

Predictions of the fitted data ($\hat{B}_y$, $\hat{\Theta}_{y,a}$, and $\widehat{\boldsymbol{A}}_{\boldsymbol{y,p}}$) were calculated from the same equations used for data simulation (Eqs. 19-22, 24-26). The sampling CV for the biomass index (σ_{B}), and effective sample sizes for both age-composition ($Z_{\Theta,y}$) and seroprevalence data ($Z_{A,y}$) were inputs to the estimation model, and were assumed to be known exactly. The estimation model was implemented in TMB (Kristensen et al., 2016) and optimized with 'nlminb' in R (R Core Team, 2020).

3.3.6 Simulation scenarios

We explored the robustness of our estimation model to a variety of scenarios that varied in assumptions about fish biology, epidemiology, and sampling protocols, both within the operating and estimation models (Table 3.2). In each scenario, 500 replicates of the operating model were run, where each replicate differed by randomly generated values for the recruitment deviations (ε_y) and observation errors (e.g. δ_y). The random seeds were preserved amongst scenarios so that scenarios are comparable. Then the estimation model was fitted to the data generated from the operating model to determine which parameters could be estimated accurately.

Ignore disease: This reflects the conventional stock assessment scenario where disease is unaccounted for. Antibody data were not fit, and infection and recovery parameters were ignored in the estimation model.

Incorporate infection prevalence: Matches the method currently used to estimate annual mortality from VHSV in Prince William Sound herring stock assessments: instead of using antibody data (estimating past disease), only the prevalence of VHSV infection ($\#\ VHSV\ positive/\#\ Total$ in a single sample) was included as an annual index of additional natural mortality in the stock assessment (Muradian et al., 2017):

$$\exp\left[-(M_y + \kappa\bar{\Psi}_y)\right] \tag{33}$$

where κ is an estimated parameter that scales infection prevalence to natural mortality and $\bar{\Psi}_y$ is a year-specific infection prevalence index generated from the average of three samples collected each year. Seroprevalence data were ignored. To simulate this scenario within the operating model, the proportion of infected fish in the population available to sampling on a given day d was computed:

$$\Psi_{y,d} = \frac{\sum_{a=0}^{n_a} s_a I_{y,d,a}}{\sum_{a=0}^{n_a} s_a N_{y,d,a}} \tag{34}$$

Eq. 34 was applied to numbers-at-age-and-stage on three random days within a 20-day timeframe during the 120-day transmission season; these conditions assumed clustered sampling at any time during the season. These true infection prevalence values were used as the mean probability in a binomial distribution to randomly generate virus-positive samples from a sample size of 60 per day.

Incorporate seroprevalence: The estimation model was fit to age-specific seroprevalence data with data-rich age-specific sample sizes (200 fish sampled per year for each survey). The assumed

age-specific proportions mixing with the reservoir population matched the maturity and survey selectivity assumed in the operating model ($v_a = s_a$).

Small sample size: Identical to the scenario incorporating seroprevalence data into the estimation model, except only a small sample size (20) was collected for seroprevalence and age composition data each year .

Early mixing of susceptible: In the operating model, susceptible fish mixed with the reservoir population at a younger age (a_{50}^{mix} = 1 yr) than in all other scenarios (a_{50}^{mix} = 3 yr).

Age-specific mixing ignored: Identical to the incorporating seroprevalence, except that the estimation model assumed that all ages were completely mixed in the reservoir population and schooling together ($v_a = 1$).

Time-varying background natural mortality: The operating model included time-varying background natural mortality instead of constant natural mortality as assumed in the other scenarios thus far. Time-varying background mortality (changing M_y) was modeled as being lognormal about the expected value (Table 3.2), but the estimation model did not estimate time-varying background mortality while incorporating seroprevalence data.

Time-varying background natural mortality/ignore disease: The operating model from the 'Time-varying background natural mortality' scenario was paired with an estimation model that ignored disease.

Time-varying disease mortality/recovery: Same as the 'Incorporate seroprevalence' scenario except recovery rates and disease mortality rates varied from one year to the next, since VHSV infections depend on water temperature (Hershberger et al., 2013). Recovery (γ_y) and disease mortality (α_y) transition probabilities were generated from uniform distributions in the operating model for each year. The uniform distributions were bounded by rates observed in VHSV tank

experiments with herring under hot (γ_y = 70 and α_y = 9 yr^{-1}) and cold conditions (γ_y = 20 and α_y = 21 yr^{-1}, Table C1; Hershberger et al. 2013).

3.3.7 Estimation performance

Performance metrics were used to quantify the error in the estimates of model outputs. We use relative error (RE) and median absolute relative error (MARE) to measure bias (by the sign of RE) and precision (by the magnitude of RE and MARE) between the true value x_i and estimated value $\hat{x}_i$ for the simulated replicate i of the output:

$$RE = (\hat{x}_i - x_i)/x_i \tag{35}$$

$$MARE = median(|\hat{x}_i - x_i|/x_i) \tag{36}$$

RE was calculated for the 50 years of estimated spawning stock biomass and recruitment across all scenarios. We also calculate the error in estimates of annual infection rates (ω_y), when included in the estimation model; the true ω_y is the total new infections in a year divided by the number of susceptible fish at the start of the transmission period. Because ω_y is bounded between 0 and 1 and can be zero or close to zero in some years (see 'Results'), annual error was summarized as a simple deviation (DEV) between the true and estimated values:

$$DEV = \hat{\omega}_i - \omega_i \tag{37}$$

MAREs were calculated for spawning biomass and recruitment for the final year of the modelled period and as median of all year-specific MAREs from years 2 through 50. The median DEV of ω_y was calculated over years 2 through 49 (the final year of estimation was excluded as seroprevalence was sampled prior to and did not inform the current year's infection). MAREs are also presented for the main disease-related parameter estimates: the recovery probability (λ), the age at 50% (a_{50}^{trans}) and age at 95% (a_{50}^{trans}) mixing with the reservoir population. The true λ was

derived from the operating model as the daily recovery rate divided by the sum of the daily recovery and disease mortality rates ($\frac{\gamma}{\alpha+\gamma}$) and was constant over all years in most scenarios (see Table C1).

Population assessments primarily rely on accurate estimates of stock status and trends, but are also concerned with the risk of large errors in estimation, here defined as the proportion of simulations where estimated spawning biomass ($\hat{B}_{y,i}$) was either more than 40% above (P_{above}) or less than 40% below the true spawning biomass $B_{y,i}$ (P_{below}). The threshold of 40% has been used previously to represent the risk of large estimation error (e.g. Punt et al. 2018) as a reflection of the average error demonstrated in data-rich stock assessments of various US West Coast groundfish and coastal pelagic species (Ralston, Punt, Hamel, DeVore, & Conser, 2011). We report these probabilities for the final year of the modelled period, as well as the median of the probabilities calculated for each year.

We checked for convergence of all estimation models fit in each scenario from the convergence codes output by the 'nlminb' function within R (R Core Team, 2020). Across all scenarios, 98-100% of the estimation models converged. Results are presented only for converged models.

3.3.8 Application to Prince William Sound herring

The estimation model equations developed above were applied to actual seroprevalence data collected for VHSV in Prince William Sound herring. VHSV effects have been intensively studied in Pacific herring, and likely has a role in the low abundance of Prince William Sound herring that has persisted since the mid-to-late 1990s; however, the extent of this role has not been determined (Elston & Meyers, 2009; Pearson et al., 2012). Age-specific seroprevalence was

obtained from serological surveys using samples collected from the seine and cast net surveys. Seroprevalence was reported as the percent of individuals containing detectable levels of neutralizing antibodies in a plaque neutralization test. Serological methods followed Hart et al. (2017), with the minor exception that endogenous complement was heat inactivated from all plasma samples and replaced with exogenous complement from specific pathogen-free individuals.

We modified the Bayesian age-structured stock assessment model used for this population (Muradian et al., 2017) to fit survey and seroprevalence data through 2020. Briefly, this stock assessment currently fits to six datasets including two hydroacoustic indices of relative abundance, one relative abundance index of milt coverage, an egg deposition survey measuring absolute total female abundance, fishery-dependent age-composition data, and fishery-independent age-composition data from seine and cast net surveys on pre-spawning aggregations of herring (available from 1982-2020; Muradian et al., 2017). For the fishery-dependent and -independent age-composition data, an iterative model fitting procedure is used twice to estimate the effective sample sizes as described in Muradian et al. (2017) and recommended by Stewart and Hamel (2014). For the age-specific seroprevalence data, the original samples for each age group ($N_{y,a}^{sero}$) are used in the binomial likelihood (Table 3.1). While catches were directly removed from the population in the early part of the time frame, fishing has not occurred since 1998 due to low biomass.

To model annual disease effects, we incorporated the equations of the estimation model used in simulation testing (e.g. Equations 24-26, 27-29, and Table 3.1) into the Bayesian stock assessment. The additional estimated parameters were year-specific infection probabilities (ω_y with a uniform prior from 0-1), year-specific recovery probabilities (λ_y with a uniform prior from

0-1), logistic parameters for age-specific mixing with the reservoir population (a_{50}^{mix} and a_{95}^{mix}), and logistic parameters describing age-specific survey selectivity for the seroprevalence samples (a_{50}^{survey} and a_{95}^{survey}, with a uniform priors bounded from 0-3 and 1-5, respectively). The same estimated selectivity curve was used for the observed seroprevalence and age-composition data because the same gear was used to collect both types of samples. It is important to note that we assumed maturity-at-age is equivalent to selectivity because the seine and cast net surveys target aggregations assumed to reflect the spawning population. Using a logistic selectivity/maturity curve was a modification from prior versions of the herring stock assessment that directly estimated the proportion mature at ages 3 and 4 for two time periods, and assumed herring were fully mature by age 5 (Hulson et al., 2007; Muradian et al., 2017). In this modified stock assessment, we also modeled the age-structure starting from age 0 instead of age 3 (in Muradian et al. 2017) so that we could fit to seroprevalence data for ages 1 and 2.

We modified the existing Bayesian stock assessment that had been already coded in AD Model Builder (ADMB; Fournier et al., 2012). Parameter estimation was done using the no-U-turn sampler (NUTS) (Monnahan et al., 2017) with the R package "adnuts" (Monnahan & Kristensen, 2018) inside R statistical software (R Core Team, 2020). The "adnuts" package was specifically developed to apply NUTS to existing ADMB stock assessment models. Three chains of 3,000 samples were generated using a diagonal mass matrix (the default in adnuts) adapted with a warm-up phase of 500 samples and a target acceptance rate of 0.925. The results from all chains were combined. Convergence of the model was supported by all parameter potential scale reduction $\hat{R}$ values being less than 1.1 (<1.1; Gelman et al., 2014a) and zero divergences in all chains.

3.4 RESULTS

3.4.1 Operating model characteristics

A 50-year subset of an example simulation of 500 years with stochastic recruitment shows highly dynamic interannual disease outbreaks (Fig. 3.2). In some years the progression of infections was rapid and in other years slow; and most years had low infection prevalence, interspersed with years that have outbreaks (Fig. 3.2A-C). Cross correlations between annual population characteristics and disease-derived quantities (Fig. 3.3) shown that recruitment from 3-4 years previously is positively correlated with current infection prevalence (ρ = 0.71-0.74); in other words, outbreaks occurred when large year classes became exposed to infection (at age 3-4). However, large differences in the amount of recruitment did not translate to large differences in the magnitude of infection prevalence or survival, suggesting that complete age and disease stage structure regulated the impacts of outbreaks (Fig. 3.2F).

Outbreaks increased seroprevalence in the following years (Fig. 3.2-3). Seroprevalence is measured prior to the transmission period in each year and reflects outbreaks in previous years. With increasing time since the most recent outbreak, seroprevalence decreased as immune cohorts die from natural mortality (Fig. 3.2F). Years with higher infection rates had higher disease-associated mortality, but this was followed by lower total mortality in subsequent years because of acquired population immunity following outbreaks (Fig. 3.3).

Changes in disease stage structure also determined differences in age-specific survival, infection, and seroprevalence within the operating model (Fig. 3.4). Profiles of survival, seroprevalence, and infection in individual cohorts over their life show fish may start acquiring immunity when they begin to mix with the reservoir population at ages 3 to 4, the earliest ages they may also experience disease mortality. Afterward, immunity increases with age as a cohort

experiences multiple outbreaks over its life (Cohorts 50, 56-58; Fig. 3.4B). However, some cohorts experienced infection and acquire immunity late in life as well (Cohorts 54-55). Thus, outbreaks could be predicted from the seroprevalence data: they are more likely to occur when a large cohort mix with the reservoir population (at ages 3 and older) and when older fish have low immunity (Fig. 3.4C).

3.4.2 Simulation analysis: Estimation without seroprevalence data

Population estimates were biased when seroprevalence data were not included in the estimation model. The bias in spawning biomass and recruitment from the 'Ignore disease' scenario stemmed from the estimation model missing the time- and age-varying natural mortality due to additional disease-associated mortality (Fig. 3.5). Similarly, biased estimates resulted when using infection prevalence index to account for disease-associated mortality within the estimation model (scenario 'Incorporate infection prevalence'). In fact, the MAREs for spawning biomass and recruitment are identical between using the infection prevalence index and ignoring disease (Table 3.3). Thus, unless surveys happen to coincide with peak of an epizootic (which was unlikely in our simulations), using an infection prevalence index was akin to disregarding the impact of disease on natural mortality.

3.4.3 Simulation analysis: Estimation with seroprevalence data

The estimation model with seroprevalence data ('Incorporate seroprevalence') produced unbiased estimates (Fig. 3.5) and reduced MAREs (Table 3.3) of annual spawning biomass and recruitment compared to the scenarios without seroprevalence, and accurately estimated annual infection, age-specific mixing, and disease recovery rate. In particular, estimates of spawning biomass in the final

year improved the most (Table 3.3). However, recruitment estimates in the final year had large error (Fig. 3.5) because cohorts partially enter the spawning population at older ages and assessment models require multiple years of data on a single cohort to accurately estimate the total abundance of a single cohort.

3.4.4 Simulation analysis: Estimation with fewer seroprevalence data or infection in unobserved ages

The estimation model still provided nearly unbiased estimates of the spawning biomass and infection rates when seroprevalence sample sizes were small ('Small sample size') or fish became vulnerable at ages earlier than when they are first observed in surveys ('Early mixing of susceptible fish') (Fig. 3.5). As expected, smaller sample sizes for age composition and seroprevalence data resulted in less precise estimates of spawning biomass, recruitment and annual infection rates (Table 3.3a). The model was still able to estimate the age-specific mixing with high precision, but the recovery rate parameter had much greater error than the 'Incorporate seroprevalence' scenario (Table 3.3a). Estimation error in these parameters was even higher in the scenario with early mixing of susceptible fish. Still, the estimation model in these two scenarios provided fairly accurate pictures of annual infection and its overall impact on the population.

3.4.5 Simulation analysis: Estimation with misspecifying age-specific mixing

If the estimation model assumed that fish of all ages are mixing in the reservoir population (scenario 'Age-specific mixing ignored'), the estimation model very poorly estimated both spawning biomass and recruitment, compared to other scenarios (Fig. 3.6). Assuming all new fish are immediately susceptible transmission also assumes disease mortality at the youngest ages, in

which the model estimates substantially greater recruitment on average to compensate for the additional loss of fish that would otherwise explain the biomass and age compositions seen in surveys, despite being untrue with the operating model. The estimated recovery rate also has very large error (Table. 3.3a), which causes bias in the disease mortality experienced by each age group. Altogether, this likely produces the bias pattern seen in spawning biomass, which is consistently underestimated in recent years (Fig. 3.6). Interestingly, annual infection remains mostly unbiased despite having large error (Fig. 3.6). Thus, it is crucial to carefully consider the ages at which fish first mix with (or at least become exposed to transmission from) reservoir populations.

3.4.6 Simulation analysis: Estimation with time-varying rates

When the estimation model assumed a constant background natural mortality while the true background natural mortality varied annually (scenario 'Time-varying background mortality') mostly unbiased estimates of spawning biomass resulted (Fig. 3.6). Recruitment was slightly underestimated (Fig. 3.6). This is likely because annual variation in the true background mortality (M_y) log-normally fluctuated around the same mean assumed fixed in the estimation model. In other words, the estimation model averages out annual deviations in background mortality, but with higher MAREs in spawning biomass and recruitment (Table 3.3a). Parameter estimates related to disease (annual infection, recovery rate, and age-specific mixing) also show low error (Table 3.3a).

Ignoring disease while background natural mortality still varied (scenario 'Time-varying background mortality and ignore disease') produce similar bias patterns to 'Ignoring disease,' but with even higher MAREs in spawning biomass and recruitment (Table 3.3a). Thus, including

seroprevalence provided improvement even when other influences on natural mortality were mis-specified.

Interestingly, the estimation model still performed well when the true recovery rate varied annually but was assumed constant over time in the estimation model (scenario 'Time-varying disease mortality/recovery'). Population and infection estimates were virtually unbiased (Fig. 3.6), and the MAREs of these estimates and disease-related parameters were low and identical to the more simplistic 'Incorporate seroprevalence' scenario (Table 3.3a). While this may suggest a robust estimation model, it is unclear why the estimation model still performs well since not accounting for time-varying disease recovery and thus mortality means the total time- and age-varying mortality was misspecified. Perhaps the already low variation in survival due to disease within the operating model (Figs. 3.2 and 3.4) is not substantially impacted by changes in recovery rate from year to year, which also had relatively low variation over 100 simulated years (annual recovery rate ranged from 0.52-0.87 with a mean of 0.73). In other words, larger variation in disease mortality and recovery over time may result in more pronounced estimation errors. Understanding how and why these rates change, such as from temperature dependence (Hershberger et al., 2013), should help inform more appropriate assumptions about the recovery parameter within the estimation model (i.e. estimating a constant or varying λ_y over time).

3.4.7 Simulation analysis: Risk of substantial bias in estimates under different scenarios

A more practical consideration of the performance of estimation for stock assessment is to evaluate the risk of substantial overestimation or underestimation of stock status (the metric P_{above} and P_{below}). Scenarios that incorporated seroprevalence data for estimation shown no risk that spawning biomass estimates were substantially wrong (Table 3.3b). Scenarios with time-varying

rates (background or disease mortality) also had little risk of substantial overestimation or underestimation. In contrast, ignoring age-specific mixing in the estimation model resulted in severe underestimation of the final year biomass, while ignoring seroprevalence data led to positively biased spawning biomass estimates (Table 3.3b). Generally, including seroprevalence data to account for disease-associated mortality lowered the risk of overfishing and underutilization.

3.4.8 Application of real seroprevalence data in the stock assessment of Prince William Sound herring

When seroprevalence data from 2012-2020 were included in the Prince William Sound herring stock assessment, population immunity estimates trend upward then downward over time with peak immunity in 2015-2016 (Fig. 3.7). Immunity increased in part due to decreasing biomass (Fig. 3.7D), stable recruitment (i.e. no large cohorts of susceptible fish; Fig. 3.7E), and relatively constant infection that potentially spiked in 2014 (Fig. 3.7B). Infection estimates are generally low across this time frame; the upper 95% limit peaks at ~0.21 in 2014 and posterior medians average around 0.06 across all years. The overall disease-associated mortality in spawners is also low (Fig. 3.7C), but the posterior medians equate to approximately 65% mortality of those infected across years, which is a large estimated mortality from infection.

Posterior predictive fits of age-specific seroprevalence are good, but uncertainty intervals for some ages and years are quite large (Fig. 3.8). Generally, low precision is seen in seroprevalence of older ages in many years because of very low sample sizes in these ages, especially in years after 2015 when ages 7-9+ had five or less samples. Additionally, seroprevalence was predicted for age one in some years despite no observed seroprevalences

(except 2018) because the model estimates some age-specific mixing to occur at young ages ($v_{a=1} > 0$). That age one seroprevalence was observed in 2018 also supports early exposure in herring life history. Still, generally low seroprevalence was observed in all age groups and years (e.g. never exceeding 40% except age 7 in 2020), and low contrast in values between years, suggests no large recent VHSV outbreaks in the spawning population.

3.5 DISCUSSION

Disease is an important mechanism regulating populations. In fish populations, impacts of disease are not easily quantified and unaccounted for in the stock assessment models used for fisheries management. We developed and simulation tested using seroprevalence data within fisheries stock assessment for estimating population-level infection and disease mortality rates specifically caused by viral pathogens. The estimation model with seroprevalence information produced unbiased estimates of disease infection and recovery and quantified age- and time-varying natural mortality. This is the first study to demonstrate the use, and evaluate the performance of, seroprevalence data within a fisheries stock assessment model, providing a modelling framework that integrates epidemiological principles within fish population dynamics models to provide a realistic picture of disease impacts and improve estimation for fisheries management.

From our simulation results, we conclude that estimation models with seroprevalence data can inform the prior exposure history to, and infection from, a specific pathogen even in misspecified models. Conversely, estimation of spawning biomass and recruitment strongly depend on model assumptions and show potentially large errors even when changes in survival due to disease were relatively small. In particular, model estimates were sensitive to the parameters specifying age-specific mixing with the reservoir population (v_a) and changes in background

mortality (M_y) and recovery rates over time (λ_y). Our sensitivity analysis fixing different curves for v_a was largely demonstrative and should be expanded to consider different values given how trends in infection and disease mortality changed with v_a (e.g. v_a being <1 at ages 4 and older, or that older ages are not fully mixing with the sympatric population). Additionally, year-specific recovery rates should be estimated (as done in the Prince William Sound herring stock assessment) and other sources of natural mortality should be considered and modeled to obtain the most accurate population estimates.

The application of actual seroprevalence data within the Prince William Sound herring stock assessment model provided an ideal case study because of high quality, long-term population data (Hulson et al., 2007; Muradian et al., 2017) and extensive disease monitoring in this population over the last three decades (Elston & Meyers, 2009; Hershberger et al., 1999; Marty et al., 1998; Marty et al., 2010; Marty et al., 2003; Meyers et al., 1994). Disease monitoring evolved to improve the VHSV information fed to the stock assessment, in particular moving beyond the use of infection prevalence indices which has been advocated (Marty et al., 2010; Marty et al., 2003; Quinn et al., 2001) as well as criticized (Elston & Meyers, 2009; Hershberger et al., 2015). Our simulation results showed that using infection prevalence performed no better than ignoring disease-associated mortality. We therefore recommend discontinuing its use within this stock assessment. However, infection prevalence may be useful for quantifying disease impacts from other pathogens that cause slower transmitting or endemic diseases. For example, the *Ichthyophonus hoferi* parasite, which is found in Pacific herring and other species, causes acute mortality or persistent infections in surviving fish (Hershberger, 2012; Kocan et al., 1999). *Ichthyophonus* prevalence is incorporated into the Prince William Sound herring stock assessment in the same way as VHSV prevalence is (e.g. Eqn. 32; Muradian et al., 2017), although this too

does not accurately quantify mortality from *Ichthyophonus* since not all symptomatic fish die (Hershberger et al., 2016). While our simulation framework used VHSV in herring as the model host-pathogen interaction, it can be readily adapted to test the utility of prevalence indices of *Ichthyophonus* and other pathogens that have well-characterized epidemiology.

Despite adequate performance of the estimation model with seroprevalence in our simulation analysis, posterior predictive fits to real VHSV seroprevalence data had wide intervals. One reason is consistently small samples for older ages, especially after 2015, which would otherwise be informative of past infection. Overdispersion in these data would be better addressed with a different error model, such as the beta-binomial distribution, and doing so may increase uncertainty further. Future applications of seroprevalence data within stock assessment should evaluate different error distributions and rely on expert judgment regarding the reliability of the process used to collect samples for serological testing.

Accounting for age- and time-varying natural mortality remains a key challenge for stock assessment and has been increasingly reviewed and evaluated (Deroba & Schueller, 2013; Jiao, Smith, O'Reilly, & Orth, 2012; Johnson et al., 2014; Lee, Maunder, Piner, & Methot, 2011; Punt et al., 2021). Generally, effects of age-varying natural mortality are secondary to time-varying mortality on model performance (Deroba & Schueller, 2013; Jiao et al., 2012), but ignoring age variation can substantially bias estimates under certain conditions, especially depending on how selectivity is modeled (He, Ralston, & MacCall, 2011). Disease impacts are ignored in these and other studies that estimate varying natural mortality based on theory or empirical studies (e.g. as it relates to size and growth, or the "U-shape" assumption across ages; Chu, Chien, & Lee, 2008; Gislason, Daan, Rice, & Pope, 2010; Lorenzen, 1996). We show disease induces age-varying mortality whose shape changes over time (Fig. 3.4) and ignoring this degrades population

estimates. This is also consequential for errors in reference point estimation which directly relates to the natural mortality error (Punt et al., 2021). Ignoring disease in a host population where a pathogen is more deadly or infectious, causing more pronounced variation in survival, or where model parameters and functions are more complex or misspecified (e.g. dome-shaped or time-varying selectivity, time-variation in the underlying background natural mortality, etc.) may result in even poorer estimation. We also showed that ignoring disease effects may lead to overestimation of spawning biomass in the final year, which is used to determine current stock status or as a starting point for forecasts used for management decisions. Even with the relatively small changes in survival produced by our operating model, consistently overestimated biomass results in a systemic positive bias of stock status that risks decisions leading to overexploitation. While other sources of error in mortality due to time- and age-variation still need consideration, such as predation, our model with seroprevalence data reduces these errors in single-species stock assessments and improves the management of diseased populations.

Although our model incorporating seroprevalence is novel to stock assessment research, epidemiological models like our operating model have long been used to address hypotheses about disease dynamics and questions on disease-control tactics (Anderson, Jackson, May, & Smith, 1981; Anderson & May, 1979; Begon et al., 2009; May & Anderson, 1979). The behavior of these models is largely determined by the form of transmission, which is a complex and uncertain process in host-pathogen interactions (McCallum, Barlow, & Hone, 2001). The assumption of frequency-dependent transmission in our model is a common assumption in terrestrial and aquatic systems, as well as density-dependent transmission, and either has significant implications for pathogen persistence and the dynamic behavior of disease (McCallum et al., 2001; Murray, 2009). For example, frequency-dependent transmission in our operating model allows outbreaks in the

host population when a large proportion of the population is susceptible to infection (Murray, 2009). Simulations from our operating model show how recruitment, which is entirely susceptible, correlates with spikes in infection and dips in survival (Figs. 2-3). This has interesting implications for density-dependent survival in fish, since the absolute numbers of recruits need not be large to cause outbreaks and depress survival (e.g. as long as the recruiting cohort is substantially larger than cohorts from prior years). The magnitudes of peaks in infection and dips in survival are then determined by the disease stage transition probabilities, which were relatively small in our model and produced less pronounced decreases in survival (Fig. 3.2). A larger disease mortality probability (α) relative to the recovery probability (γ) would result in more pronounced changes in survival between years, although other combinations of transition probabilities can also produce similar or even more drastic changes.

Our operating and estimation models are useful in identifying expected patterns in the estimated infection and disease mortality from actual data as they relate to the underlying transmission model. For example, the consistency of outbreaks with recruitment due to frequency-dependent transmission should produce spikes in seroprevalence that follow large recruitments based on our results. However, in our Prince William Sound herring application, a very large recruitment of age 3 fish in 2019 (cohort from 2016) coincided with low estimated infection (Fig. 3.8). This suggests a more complex or different operating model underlying VHSV disease dynamics in Prince William Sound herring. One possibility is a different transmission model. For example, a nonlinear response in the Force of Infection due to the localized accumulation of pathogens in water (Murray, 2009) may more appropriately explain VHSV outbreaks that only occur when herring school in dense aggregations; this was observed in herring temporarily confined to pens in the pound spawn-on-kelp fishery (Hershberger et al., 1999). This form of

transmission could also arise in other settings where herring tightly aggregate, such as in bait balls herded by feeding predators (Similä & Ugarte, 1993). Alternatively, transmission can also occur from other reservoir populations of different species with which herring may come in contact (Hershberger et al., 2015). While functional forms are available for transmission from external reservoirs (McCallum et al., 2001; Murray, 2009), they are challenging to parameterize. Generally, environmental, ecological factors, and fish life history are critical to the determination of the form of transmission, and that this also changes over time. The model presented here can be modified to readily accommodate the various forms of transmission already presented in literature (McCallum et al., 2001; Murray, 2009), as well as novel forms. Altogether, our modeling framework provides a starting point for inferring the dominant form of transmission that causes outbreaks when used in conjunction with seroprevalence data either fit in stock assessment or directly examined on its own.

Despite our focus on VHSV in Pacific herring, it is fairly easy to adapt our models to other viruses and host populations. In addition to suggestions presented throughout this paper, the simplest adjustments to the operating model involve parameter value changes such as the transition probabilities within the daily disease transition matrix (Eq. 5), which can be obtained from lab experiments or regularly measured prevalence in wild populations. The number and types of stages should also be modified to reflect disease progression within individuals; e.g. a stage that reflects latent infection or exposure (where the pathogen is present, but the individual is not yet infectious) is not uncommon. Generalized forms of the Force of Infection function describing transmission have been developed by including powers to allow scaling between different forms (Hopkins, Fleming-Davies, Belden, & Wojdak, 2020). Given the importance of recruitment dynamics in our simulations, more complex extensions of our operating model should explore seasonality in daily

recruitment rates (as opposed to an annual pulse prior to transmission), which have been identified as crucial determinants of the timing of infection and pathogen persistence in wildlife populations (Begon et al., 2009; Peel et al., 2014; Matthew J. Smith et al., 2008). In fish, the fine time-scale processes of juvenile fish recruiting to a reservoir population are unknown. Future research could seek to provide a more in-depth simulation analysis of the impact of recruitment under different assumptions about timing and forms of transmission as noted earlier. Finally, our disease equations in our estimation model circumvent many assumptions regarding transmission because the estimated annual infection rates reflect the emergent dynamics, not the underlying mechanics of the host-pathogen interaction. Therefore, Equations 27-29 should be rather straightforward to incorporate into stock assessments of other species. However, for diseases with substantial numbers of infected individuals carrying over between years, the proportion of fish that remain infected will have to be explicitly accounted for and require slight modifications of Equations 27-28.

Despite the promise of using seroprevalence to model infection and disease-associated mortality, key uncertainties complicate inference. First, age-specific mixing was shown to impact estimates in our simulations and the actual application, and depends on species- and population-specific life history. Ideally, sampling conducted for seroprevalence testing should include relatively large sample sizes for all ages to resolve this issue. For example, seroprevalence samples of age 1 and 2 Pacific herring from Puget Sound show evidence of widespread infection, and antibodies were also present in age-1 fish in 2018 in Prince William Sound herring (P.H. unpub. data). Furthermore, age-specific mixing with the reservoir population is broadly analogous to the timing effect of the recruitment rate mentioned earlier (Begon et al., 2009) and may substantially impact outbreak dynamics; specifically, differing proportions of susceptible individuals from a

large cohort entering into the reservoir populations, but at yearly time steps, could effect a yearly sequence outbreaks with different peaks in each year. Finally, ignoring time variation in age-specific mixing likely biases estimates and could be treated similarly to time-varying fishing selectivity (Martell & Stewart, 2014).

There is some uncertainty associated with seroprevalence tests. The optimized plaque neutralization tests (PNT) used exhibited very low false positives (0%), high sensitivity (90%), and high specificity (Hart et al., 2017). However, other mechanisms of acquiring immunity may be missed because the optimized PNT targets specific neutralizing antibodies. Our estimation and operating models assume that seroprevalence samples reflect all forms of immunity shown in fish that have recovered from infection. If the data do not reflect all immune responses, our estimates of infection rates would be underestimated and disease-associated mortality rates would be overestimated. However, this can be easily addressed by considering the observed seroprevalence reflects a proportion of the total immunity; in other words, multiplying the estimated total immunity by a scalar to predict the observed seroprevalence in the model. Indeed, optimized PNT conducted on wild herring from tank experiments with independently induced outbreaks suggest the test detects a fixed proportion of all immune individuals (unpub. data). More validation studies are needed to develop a more complete understanding of the Pacific herring immune response to VHSV and before using PNT to measure seroprevalence from viruses in other host populations (Hart et al., 2017).

Models of seroprevalence in wildlife populations are a fairly recent development. Such models have been used to quantify exposure of wild ruminant species to Schmallenberg virus in France (Rossi et al., 2017), measure the impact of detection errors on predicting seroprevalence of viruses in wild pigs in the U.S. (Tabak, Pedersen, & Miller, 2019), and identify environmental risk

factors in seroprevalence patterns of an endemic tick-borne pathogen in sheep as transmitted from red grouse and deer in Scotland (Gilbert, Brülisauer, Willoughby, & Cousens, 2020). Our study adds a novel and unique contribution to this emerging field within ecology as a means to integrate seroprevalence data with survey demographic data to jointly estimate population-level infection history and population status. To the field of fisheries science, we provide a novel way to model age- and time-varying survival due to disease to improve fisheries stock assessments of disease-prone populations. This approach requires seroprevalence data that spans all ages and are sufficiently numerous to be useful. Parameter values and model structure will need to be tailored to both host and pathogen as outlined above and require supporting field and lab observations (e.g. Hershberger et al., 2015). Seroprevalence is also based on antibody detection and antibodies are only produced by viral infections. However, there may be other types of permanent markers left by infections that offer the same information on exposure history as seroprevalence. For example, VHSV infections persist in the gills of surviving fish which can be readily measured by qPCR (Cornwell et al., 2013). Our modeling framework can be easily adapted to these type markers, making it more generalizable while circumventing some of the mentioned issues with seroprevalence data.

Monitoring and modeling disease for stock assessment has been hailed as a priority for fisheries management with the status-quo approach being data-intensive mark-recapture models developed outside stock assessment (Groner et al., 2018; Hoenig et al., 2017). Our stock assessment model with novel serological assays offers a formal and possibly more efficient means to directly account for viral diseases within stock assessment. Additionally, our model can potentially be extended to other pathogen types with the appropriate data as outlined above. Infectious marine diseases have and will become even more frequent and severe across marine

taxa due to rapid global change (Burge et al., 2014; Fey et al., 2015; Harvell et al., 2004; Maynard et al., 2016), and our framework provide crucial tools for quantifying their impacts and better equipping fisheries management to incorporate epidemiology.

3.6 ACKNOWLEDGEMENTS

Funding for this work was provided by the Exxon Valdez Oil Trustee Council (EVOSTC). Thank you to Maya Groner and Paul Hershberger for VHSV seroprevalence data and crucial guidance and feedback on disease modeling and VHSV epizootiology. I also thank A. Punt, T. Essington, and O. Shelton for invaluable comments on the manuscript.

3.7 TABLES

Table 3.1. Equations of components from the objective function in the estimation model

Equation	Description
$\delta_{\mathrm{N}} = (n_a - 1)\ln(\sigma_{\mathrm{R}}) + \frac{\sum_{y\epsilon(n_a-1)} \varepsilon_y^2}{2\sigma_{\mathrm{R}}^2}$	Penalty on numbers-at-age deviations
$\delta_{\mathrm{R}} = n_y \ln(\sigma_{\mathrm{R}}) + \frac{\sum_{y\epsilon n_y} \varepsilon_y^2}{2\sigma_{\mathrm{R}}^2}$	Penalty on recruitment deviations
$\ell(\mathrm{B}) = n_y \ln(\sigma_B) + \frac{\sum_{y\epsilon n_y}\left[\ln(\widehat{\mathrm{B}}_y) - \ln(\mathrm{B}_y)\right]^2}{2\sigma_B^2}$	Negative log-likelihood of biomass survey
$\ell(\Theta) = \sum_{y\epsilon n_y}\left[Z_{\Theta,y}\sum_{a\epsilon n_a}\Theta_{y,a}\ln\left(\frac{\widehat{\Theta}_{y,a}}{\Theta_{y,a}}\right)\right]$	Log-likelihood of fishery-dependent age-composition
$\ell(\hat{A}) = \sum_{y\epsilon n_y}\sum_{a\epsilon n_a}\left[-k_{y,a}ln\left(\hat{p}_{y,a}\right) - (N_{y,a}^{sero} - k_{y,a})ln\left(1-\hat{p}_{y,a}\right)\right]$	Log-likelihood of antibody prevalence
$\mathcal{L} = \delta_{\mathrm{N}} + \delta_{\mathrm{R}} + \ell(\mathrm{B}) - \ell(\Theta) - \ell(\hat{A})$	Objective function

Table 3.2. Specifications of the operating and estimation models for simulation scenarios. Gray shaded boxes highlight differences from the 'Ignore disease' scenario.

	Operating model				*Estimation model*				
Name	**Age comp/ antibody sample size**	**Mixed proportion logistic parameters**	**Background mortality (M_y)**	**Mortality & recovery probabilities (α, γ)**	**Infection probability (ω_y)**	**Recovery probability (λ)**	**Mixed proportion parameters**	**Fit antibody data**	**Input infection prevalence**
Ignore disease	200/NA	a_{50}^{mix}=3 a_{95}^{mix}=4	Constant	Constant	Fix (ω_y=0)	Fix (λ =0)	NA	No	No
Incorporate infection prevalence	200/NA	a_{50}^{mix}=3 a_{95}^{mix}=4	Constant	Constant	Fix (ω_y=0)	Fix (λ =0)	NA	No	Yes
Incorporate seroprevalence	200/200	a_{50}^{mix}=3 a_{95}^{mix}=4	Constant	Constant	Est	Est	Est v_a for ages 2-4; fix v_a=0 if a<2, v_a=1 if a>4	Yes	No
Small sample size	20/20	a_{50}^{mix}=3 a_{95}^{mix}=4	Constant	Constant	Est	Est	Same as 'Incorporate seroprevalence'	Yes	No
Early mixing of susceptible fish	200/200	a_{50}^{mix}=1 a_{95}^{mix}=4	Constant	Constant	Est	Est	Est v_a for ages<5; fix v_a=1 if a>4	Yes	No
Age-specific mixing ignored	200/200	a_{50}^{mix}=3 a_{95}^{mix}=4	Constant	Constant	Est	Est	Fix v_a=1	Yes	No
Time-varying background mortality	200/200	a_{50}^{mix}=3 a_{95}^{mix}=4	$\ln(M_y) \sim N(\ln(0.25), 0.25^2)$	Constant	Est	Est	Same as 'Incorporate seroprevalence'	Yes	No
Time-varying background mortality/ ignore disease	200/NA	a_{50}^{mix}=3 a_{95}^{mix}=4	$\ln(M_y) \sim N(\ln(0.25), 0.25^2)$	Constant	Fix (ω_y=0)	Fix (λ =0)	NA	No	No
Time-varying disease mortality/ recovery	200/200	a_{50}^{mix}=3 a_{95}^{mix}=4	Constant	$\alpha_y \sim U(0.025,0.058)$ $\gamma_y \sim U(0.055,0.192)$	Est	Est	Same as 'Incorporate seroprevalence'	Yes	No

Table 3.3a-b. Summary of performance metrics of the estimation model in each scenario. Values are not provided (–) for parameters that are not estimated in a select few scenarios.

	a_{50}^{mix}	a_{95}^{mix}	λ_y	Year 50 SSB	SSB	Recruits	ω_y
Table 3.3a	MARE	MARE	MARE	MARE	Median MARE	Median MARE	Median DEV
Ignore disease	–	–	–	0.18	0.08	0.16	–
Incorporate infection prevalence	–	–	–	0.16	0.07	0.14	–
Incorporate seroprevalence	0.02	0.03	0.07	0.07	0.05	0.10	0.03
Small sample size	0.02	0.07	0.16	0.10	0.07	0.22	0.06
Early mixing of susceptible fish	0.22	0.18	0.13	0.06	0.05	0.13	0.05
Age-specific mixing ignored	–	–	0.89	0.42	0.22	2.69	0.10
Time-varying background mortality	0.02	0.03	0.07	0.12	0.09	0.13	0.02
Time-varying background mortality/ignore disease	–	–	–	0.26	0.11	0.19	–
Time-varying disease mortality/recovery	0.02	0.03	–	0.07	0.05	0.10	0.02

	P_{above}	P_{above}	P_{below}	P_{below}
Table 3.3b	Median	Year 50	Median	Year 50
Ignore disease	0.00	0.05	0.00	0.00
Incorporate infection prevalence	0.00	0.05	0.00	0.00
Incorporate seroprevalence	0.00	0.00	0.00	0.00
Small sample size	0.00	0.01	0.00	0.00
Early mixing of susceptible fish	0.00	0.00	0.00	0.00
Age-specific mixing ignored	0.07	0.00	0.00	0.57
Time-varying background mortality	0.00	0.02	0.00	0.01
Time-varying background mortality/ignore disease	0.00	0.25	0.00	0.00
Time-varying disease mortality/recovery	0.00	0.00	0.00	0.00

3.8 FIGURES

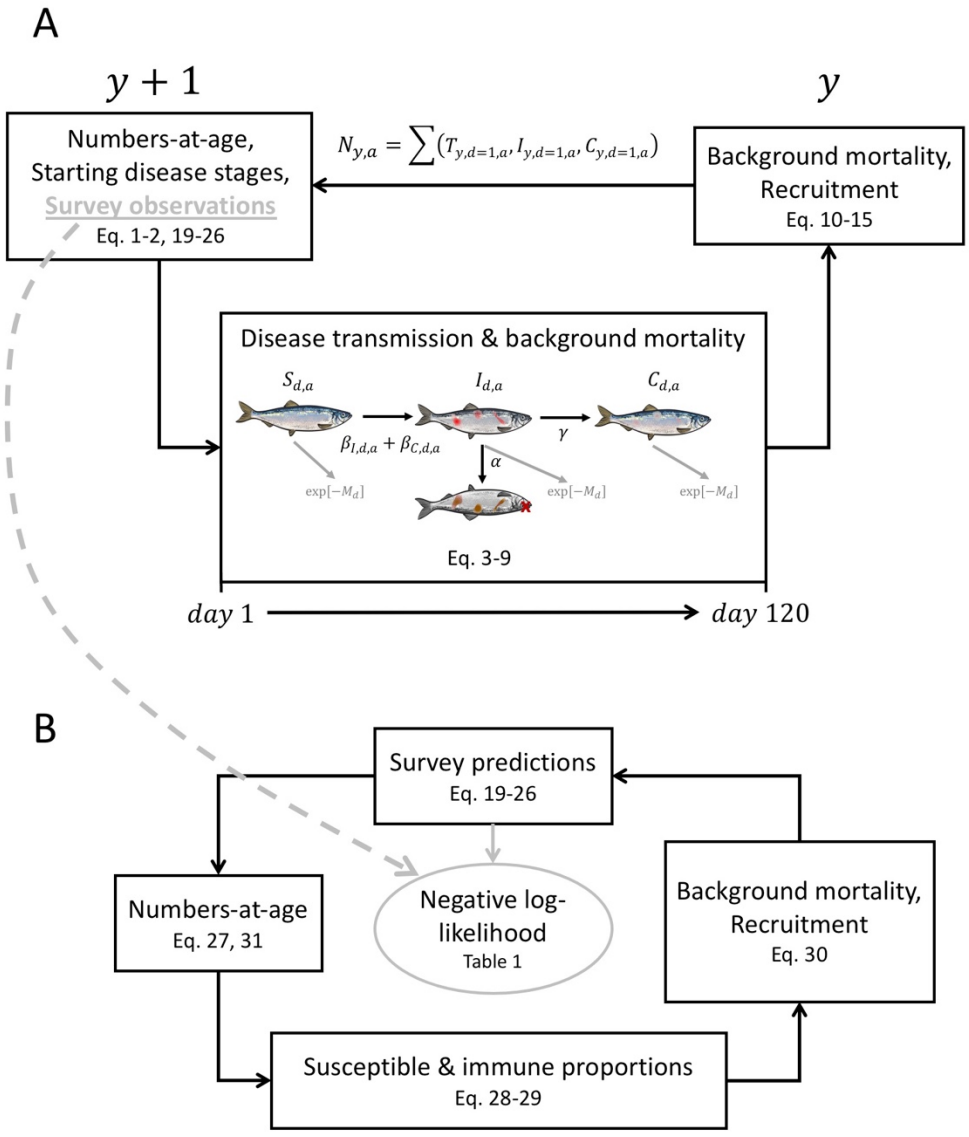

Figure 3.1. Flow diagrams of the disease operating and estimation models used in simulations. A) shows the sequence of equations in the operating model which starts with numbers of fish in each age and disease stage as projected from the previous year (y). These age and stage specific numbers ($S_{d,a}$, $I_{d,a}$, $C_{d,a}$) are projected on daily time steps using transition probabilities ($\beta_{I,d,a}$, $\beta_{C,d,a}$, α, γ) and background mortality ($\exp[-M_d]$) within a discrete-time disease transmission model. After 120 days of transmission, numbers are projected to the next year ($y+1$) after accounting for constant background mortality and recruitment in the intervening time period. B) shows a similar sequence of equations in the estimation model. The key difference between the two models is the bottom box; the estimation model calculates annual susceptible and immune

numbers from estimated annual infection and recovery parameters. Parameters are estimated by fitting survey predictions from the estimation model to survey observations simulated from the operating model (linked by the gray dashed line), and fitting is done by minimizing the negative log-likelihood functions in Table 1. The in-text equations are referenced from the relevant processes labeled in each box.

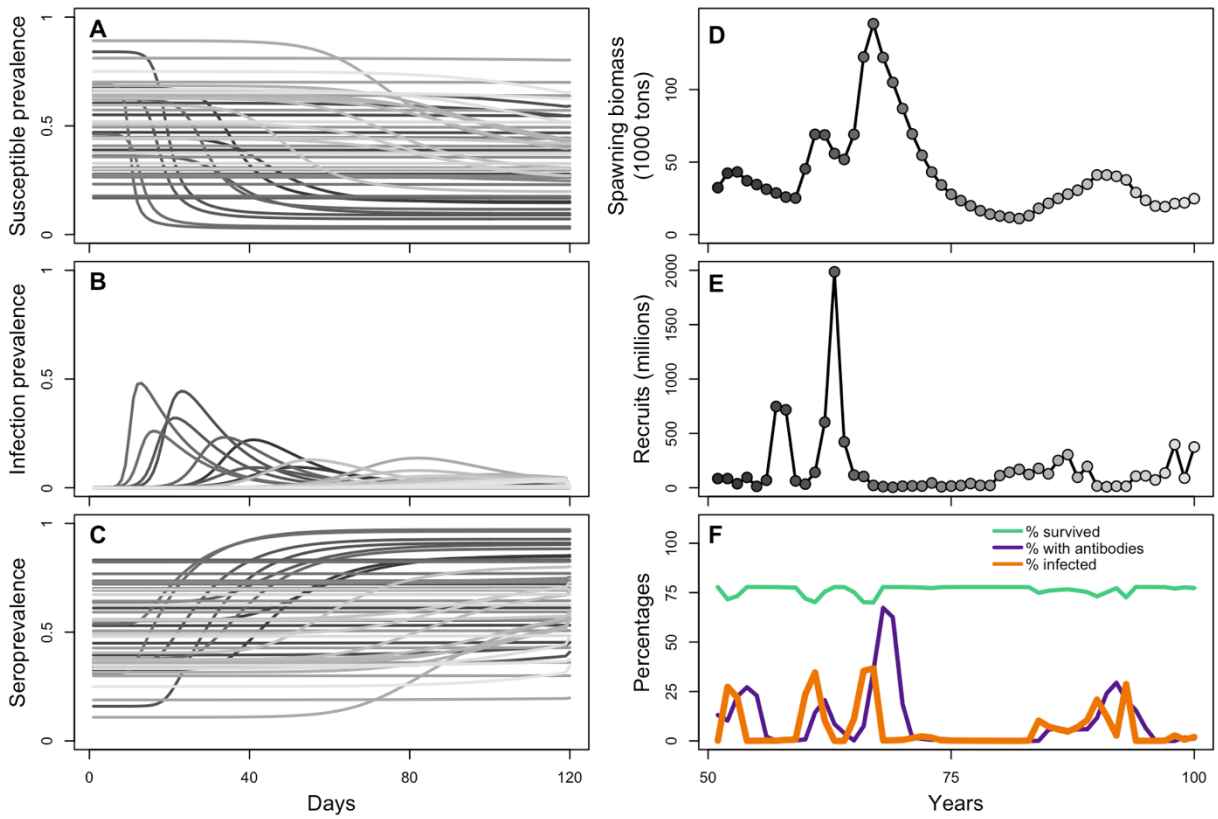

Figure 3.2. Population and epidemiological dynamics over 50 years from a single simulation of the operating model. Daily disease stage transitions are shown as the proportions of the overall population that are (A) susceptible, (B) infected, or (C) recovered and now carriers on each day of the fixed transmission period. Each line is a different year. Grey shading denotes the corresponding years between plot A-C and D-E. The annual population level dynamics are shown in the right column including (D) annual spawning biomass, (E) stochastic recruits as the number of age-0 fish, and (F) the realized population survival (the annual percentage that survive), seroprevalence (the annual percentage that are currently immune), and infection incidence (the percentage that becomes infected) of the population (excluding the plus age group) in each year.

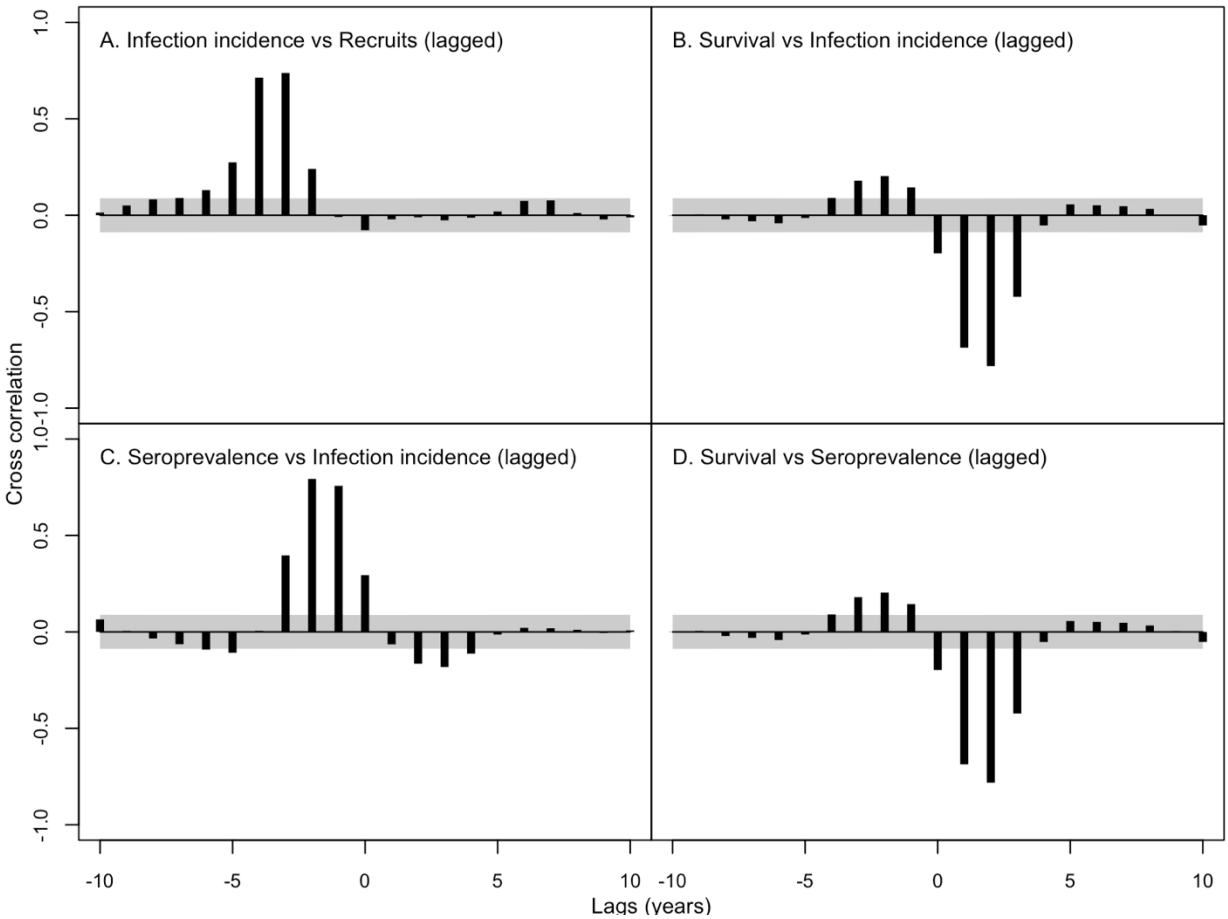

Figure 3.3. Spearman's cross-correlations (black bars) between several population and disease-related quantities calculated for 500 years from a single simulation of the operating model. Correlations are computed at different lags (up to 10 years from the past and the future) between annual infection incidence (A-C), recruitment (A), survival (B, C), and seroprevalence (C-D). Cross correlations are significant if the black bars exceed in either direction the grey shaded area.

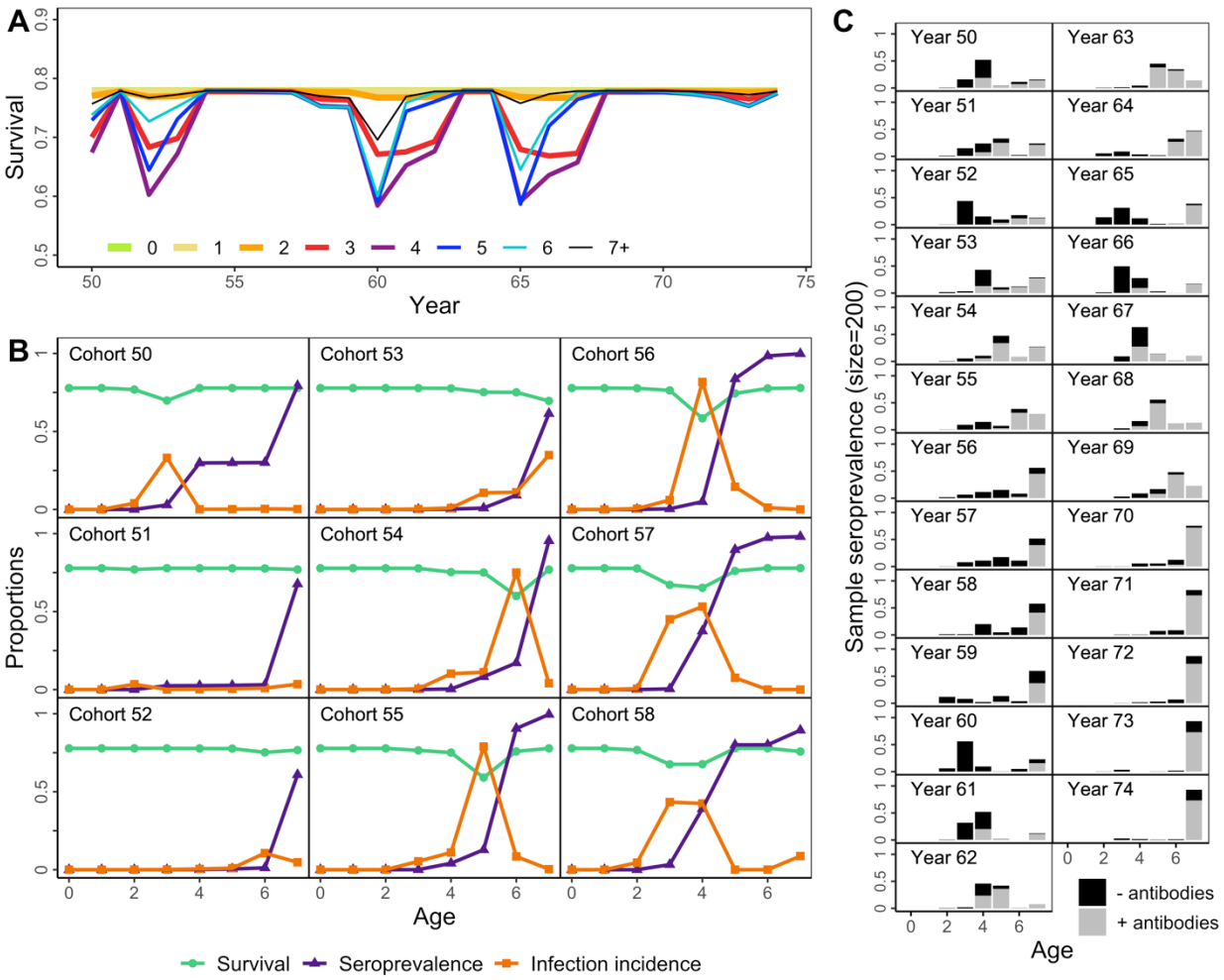

Figure 3.4. Age-specific time series of survival, seroprevalence, and infection in a subset of years. The upper left plot (A) displays age-specific survival over time in years 50 through 74 of the dynamics shown in Fig. 3.2. The lower left plot (B) displays age-specific profiles by cohort (i.e. Cohort 50 is the age 0 fish in year 50, age 1 in year 51, etc.) of survival, seroprevalence, and infection incidence experienced at each age. The right-most plot (C) is a simulated seroprevalence sample from the multinomial distribution using probabilities from the actual age-specific seroprevalence (accounting for survey selectivity) and a sample size of 200. Note the y-axis in plot A starts at 0.5 and ends at 0.9.

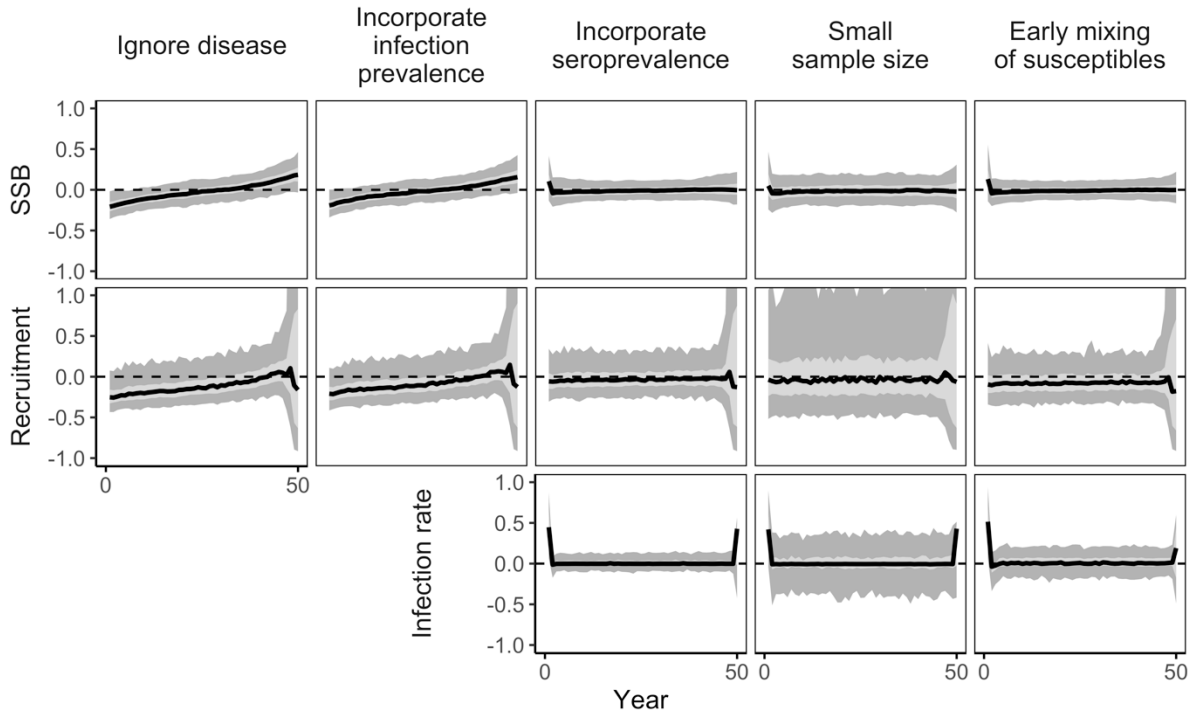

Figure 3.5. Time trajectories of error in estimates of spawning biomass (SSB), recruitment, and annual infection rate (ω_y) across 50 years of simulation from five scenarios (Table 3.2). Relative error (RE) is calculated between the actual values (from the operating model) and the estimated values (from the estimation model) for SSB and Recruitment. Deviation (DEV) is shown for infection rates because these values are scaled between 0 and 1. The black denotes the median relative error or error across converged simulations, the light gray denotes the inner-50th percentiles, and dark denotes inner-95th percentiles.

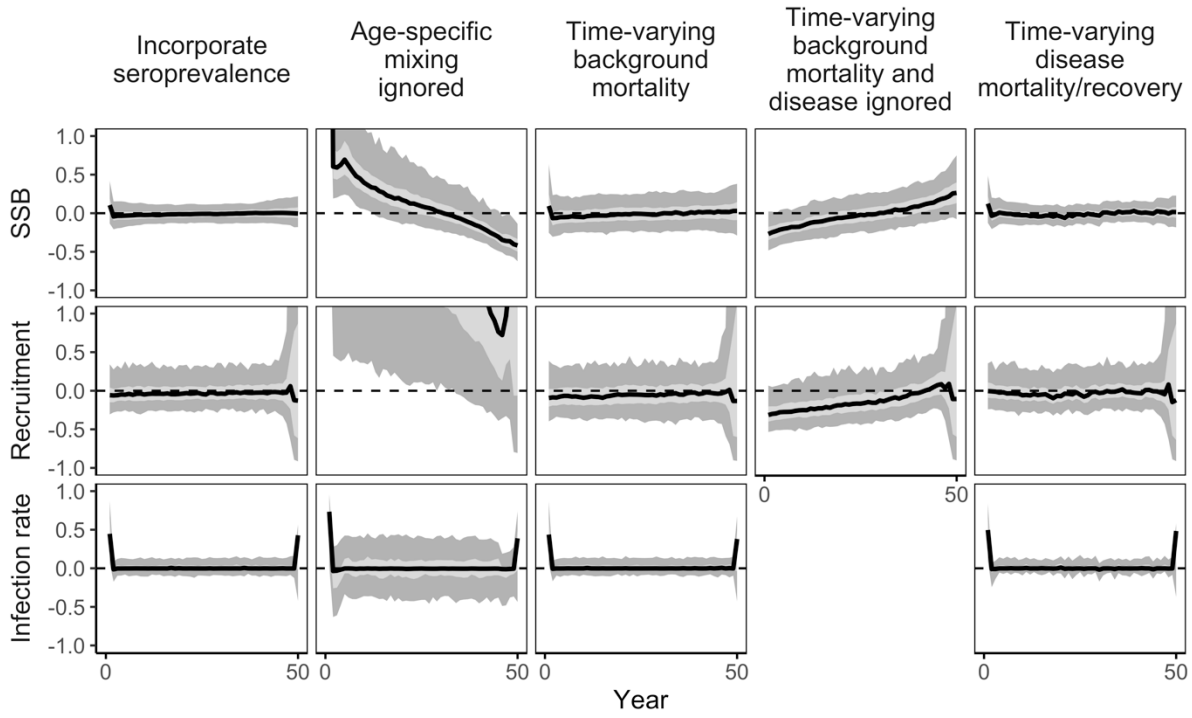

Figure 3.6. Time trajectories of error in estimates of spawning biomass (SSB), recruitment, and annual infection rate (ω_y) across 50 years of simulation from the Incorporate seroprevalence scenario and the remaining four scenarios (Table 3.2).

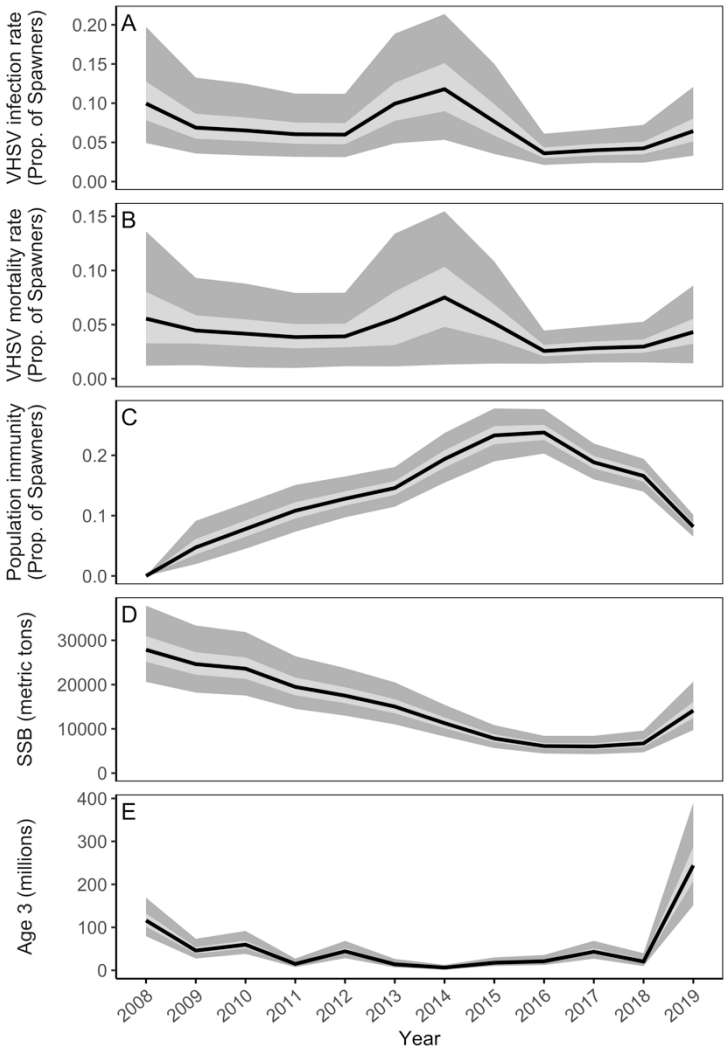

Figure 3.7. Bayesian estimates of disease and population levels over recent years from the Prince William Sound herring stock assessment. From top to bottom are time series estimates of the annual infection rate of VHSV within only the mature numbers, or spawners, of the population (VHSV infection rate), the proportion of spawners that are lost to disease (VHSV mortality rate), proportion of spawners that contain antibodies and have immunity at the beginning of each year (Population immunity), the spawning biomass (SSB), and recruitment numbers at age-3 (Age 3). The black line denotes the median estimate from the posterior distribution, light gray the 50% credibility interval, and dark grey the 95% credibility interval.

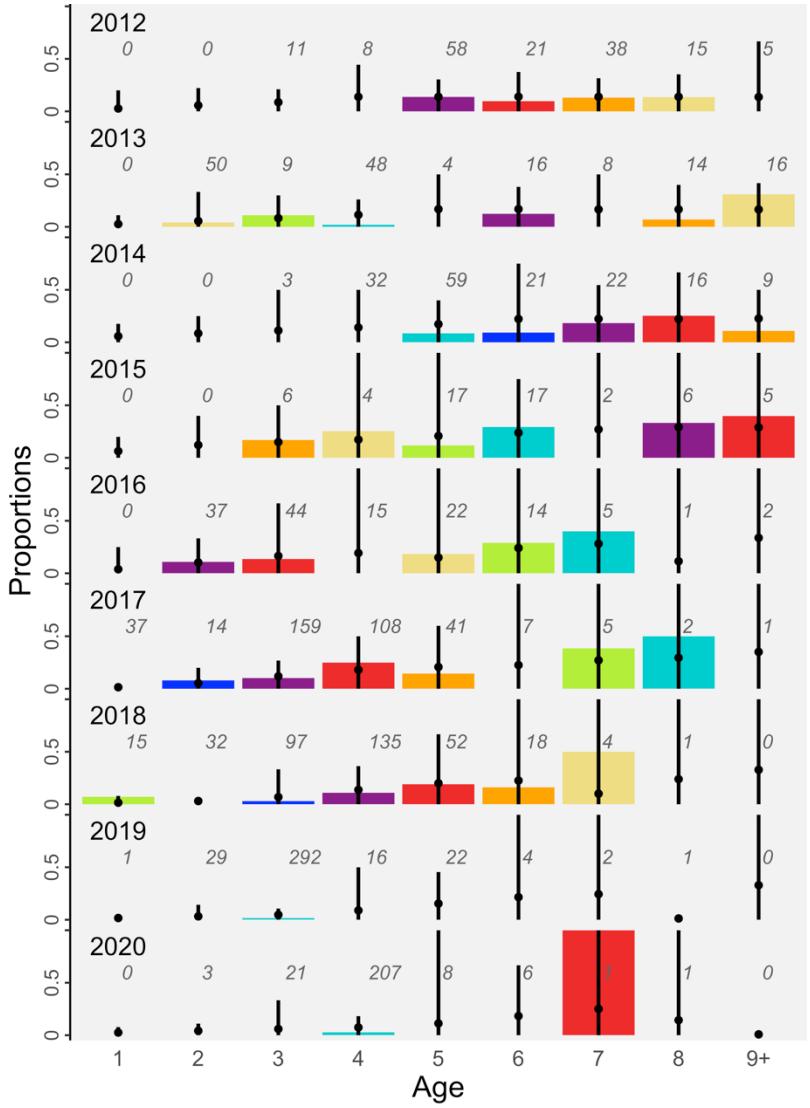

Figure 3.8. Posterior predictive fits to age-specific seroprevalence of samples collected and tested for Prince William Sound herring. Colored bars denote the data where individual colors track cohorts through time. The points show the posterior predictive means with bars spawning the 95% predictive intervals. The italicized numbers above each bar are the sample sizes for each age group

Conclusions

The population dynamics of herring are in part regulated by complex external factors that force large, mostly unpredictable fluctuations in abundance resulting in population collapse and recovery. While herring collapse and recovery may arise from natural processes, such drastic changes challenge effective fisheries management and specifically undermine fixed limits or targets, such as reference points, that may take many years to meet or surpass. Despite a century of research since Hjort (1914), the ability to explain or predict herring collapse and recovery remains a key gap and ongoing priority for management and the stock assessment models that inform it. Generally, fishing, multi-species interaction, and the physical environment are typified as three generic sources of population collapse and recovery. In this dissertation, I explored and evaluated the importance of various factors from these sources amongst herring stocks worldwide and for one specific stock, Prince William Sound herring, in-depth. Additionally, my work continues to advance the stock assessment model of Prince William Sound herring.

It is common for herring stocks to experience low biomass for a decade, and some experience collapse for even decades more, including Prince William Sound herring. In Chapter 1, I found these durations broadly relate to average recruitment and oceanographic variability (SD of SSHA and SST) amongst stocks. In other words, during an instance of collapse, management and fisheries are at the whims of environmental conditions that are favorable for strong recruitment which may be more likely in historically variable environments. Additionally, many Atlantic herring stocks are fished through collapse with no apparent impact on recovery time; conversely, all Pacific herring stocks are closed to fishing during collapse with no impact on the duration of recovery time. While this suggests that fishing may continue at low herring abundance without impacting the time to recovery, this largely depends on the management system and the number,

sizes, and types of fisheries targeting herring. For either species of herring, management should carefully consider the risks of fishing herring at low abundance that depend on population-specific characteristics and environmental conditions. Management strategy evaluations (MSEs) would be very useful in this capacity, where the tradeoffs between objectives focused on both conservation and fishery viability could be tested under fishing during simulated population collapses.

While global patterns in herring recovery times provide necessary context to understanding the severity and uncertainty of population collapse in herring, population-specific case studies provide more tractable information for fisheries management. I showed that the collapse of Prince William Sound herring was an extreme event compared to other populations. That Prince William Sound herring has remained collapsed for as long as it has suggests unusual conditions and factors specific to Prince William Sound may be suppressing the productivity (specifically, recruitment and natural mortality) of this population. Chapters 2 and 3 investigate how to model these factors within the Prince William Sound herring stock assessment. In my exploration and evaluation of various potential covariates in Chapter 2, I found the strongest evidence for an effect of pink salmon on natural mortality, with more ambiguous support for other multi-species (humpback whales and walleye pollock) and broad-scale oceanographic (PDO and NPGO) effects. Disease had been previously suggested as important, particularly from VHSV, but I found no support for substantial effects on time-variation in mortality in Prince William Sound herring, despite improving the realism of the VHSV mortality model in Chapter 3. More broadly, the results from Chapters 2 and 3 highlight the importance of changes in natural mortality for Prince William Sound herring and for the advancement of methodology that can better model these changes within stock assessment. For recruitment, I found some support for long-term shifts in average recruitment (e.g. from either an environmental regime shift or step-wise shift in hatchery releases of juvenile pink

salmon recruitment), but nothing explained interannual variability which is important since large recruitment events are needed for recovery in herring populations. Encouragingly, results from the most recent Prince William Sound herring stock assessment in Chapter 3 show a relatively large cohort of age 3 fish entering the population in 2019, leading to a more than 100% increase in spawning biomass from the previous year. However, recruitment variability remains the greatest enigma to fisheries science, most especially for collapsed populations where recovery is contingent on recruitment. Future work on Prince William Sound herring should focus on improved models of recruitment including accounting for and evaluating hypotheses surrounding pre-recruit survival in ages 0-2 fish.

Results from Chapters 2 reveal caveats in common approaches to modeling ecological effects in single-species stock assessments. Model selection remains a popular methodology for hypothesis evaluation. My demonstration of multiple tools in Bayesian model selection reveals that while it is possible to find some consensus in support for certain models suggesting stronger evidence, it could also mean that multiple criteria are selecting these models for the same "wrong" reasons. In particular, additional diagnostics for the Pareto-smoothed leave-one-out criteria suggest the criterion values may be unreliable either because the models tested may be misspecified, or even just too flexible for this model selection to be reliable. A much more in-depth analysis would be needed, requiring *K-fold* cross-validation (in which it isn't clear how to apply this to BASA with multiple fitted data sets) or exploring other structural assumptions outside mortality and recruitment within the Prince William Sound herring stock assessment that may be misspecified. Additionally, the treatment of ecological covariates as fixed effects or as a latent variables of some underlying process has consequences for my results. I found model selection criteria and estimates of spawning biomass and recruitment greatly differed between the fixed

effect and latent variable models of the same environmental factor, especially when impacting recruitment. This is likely because a substantial portion of the recruitment variation remains unexplained apart from any estimated ecological effect. Altogether, scientists need to carefully consider the inference they want when using alternative criteria for model selection or making specific assumptions about covariates, and evaluate the consequences to the information used to inform fisheries management

Despite the importance of recruitment, models that misspecify natural mortality can lead to substantially biased estimates (Chapter 3 and Punt et al., 2021). My dissertation advances the research on this topic by presenting more robust approaches to modeling time-varying mortality. The synthesis of different methods in model selection and different models of ecological effects in Chapter 2 resulted in identifying potentially substantial effects on natural mortality that could improve overall model predictions compared to modeling predictors of recruitment. The novel models of Chapter 3 provide two key benefits to stock assessment: 1) a more robust and realistic accounting of overall disease-associated mortality using seroprevalence data and 2) a model that informs age- and time-varying effects of disease on natural mortality. While choices in model structure and parameter values were catered to VHSV in Pacific herring, modifications can be made to readily expand the modeling framework to other host populations and pathogens. In particular, the estimation model for stock assessment is generalizable as is, and the disease-related equations can be adapted for any population where there is age-specific data that measures a permanent marker from infection, whether that is the presence of antibodies or histological signs of chronic infection. We anticipate that these modeling tools will become even more essential with the expectation of more frequent and severe disease outbreaks from existing and new pathogens in a rapidly changing ocean.

In summary, I evaluated what factors impact population collapse and recovery in herring and demonstrated how they should be accounted for in fisheries stock assessments. While fishing may make forage fish collapses more common and severe (Essington et al., 2015), environmental conditions ultimately dictate when herring recover. Across stocks, more variable environmental conditions coincide with shorter collapse durations; and within Prince William Sound herring, pink salmon may increase natural mortality, but likely not prevent their recovery. Disease remains a concern for Prince William Sound herring and other species and can now be better addressed with advances in disease monitoring techniques combined with novel models shown in this dissertation. Even as I have found associations between herring population dynamics and ecological factors and potentially better models of these relationships, substantial uncertainty continued to caveat inference on the supporting evidence in these results. The sources of uncertainty generally fall under the major categories often described for population dynamics models, including process error from highly variable and unpredictable herring recruitment; observation errors in the data on ecological covariates and seroprevalence; and model error from misspecified equations or parameters related to recruitment and mortality. Even so, my results continue to highlight the inherent complexity underlying fish population dynamics and modeling approaches for their assessment. These insights into the factors behind population collapse and recovery and the appropriate tools that account for these factors direct future research priorities and provide valuable information to the management of herring fisheries.

Bibliography

Ainsworth, C., Pitcher, T., Heymans, J., & Vasconcellos, M. (2008). Reconstructing historical marine ecosystems using food web models: Northern British Columbia from Pre-European contact to present. *Ecological Modelling, 216*, 354-368. doi:https://doi.org/10.1016/j.ecolmodel.2008.05.005

Alexander, G. M., Michael, O. C., & Brian, J. (2003). A model of spatially evolving herpesvirus epidemics causing mass mortality in Australian pilchard Sardinops sagax. *Diseases of Aquatic Organisms, 54*, 1-14. Retrieved from https://www.int-res.com/abstracts/dao/v54/n1/p1-14/

Anderson, R. M., Jackson, H. C., May, R. M., & Smith, A. M. (1981). Population dynamics of fox rabies in Europe. *Nature, 289*, 765-771. doi:10.1038/289765a0

Anderson, R. M., & May, R. M. (1979). Population biology of infectious diseases: Part I. *Nature, 280*, 361-367. doi:10.1038/280361a0

Anderson, S. C., Cooper, A. B., Jensen, O. P., Minto, C., Thorson, J. T., Walsh, J. C., . . . Selig, E. R. (2017). Improving estimates of population status and trend with superensemble models. *Fish and Fisheries, 18*, 732-741. doi:10.1111/faf.12200

Andrews, A. G., Strasburger, W. W., Farley Jr, E. V., Murphy, J. M., & Coyle, K. O. (2016). Effects of warm and cold climate conditions on capelin (Mallotus villosus) and Pacific herring (Clupea pallasii) in the eastern Bering Sea. *Deep Sea Research Part II: Topical Studies in Oceanography, 134*, 235-246. doi:https://doi.org/10.1016/j.dsr2.2015.10.008

Bakun, A. (1973). Coastal upwelling indices, west coast of North America, 1946-71. *United States National Marine Fisheries Service Special scientific report - fisheries*, 103.

Bakun, A. (1975). Daily and weekly upwelling indices, west coast of North America, 1967-73. *NOAA technical report NMFS SSRF*, 114. Retrieved from ftp://ftp.library.noaa.gov/noaa_documents.lib/NMFS/TR_NMFS_SSRF/TR_NMFS_SSRF_693.pdf Online version in PDF via the NOAA Central Library (5.4 MB)

Bakun, A., & Broad, K. (2003). Environmental 'loopholes' and fish population dynamics: comparative pattern recognition with focus on El Nino effects in the Pacific. *Fisheries Oceanography, 12*, 458-473. doi:https://doi.org/10.1046/j.1365-2419.2003.00258.x

Barnes, C. L., Beaudreau, A. H., Dorn, M. W., Holsman, K. H., & Mueter, F. J. (2020). Development of a predation index to assess trophic stability in the Gulf of Alaska. *Ecological Applications*.

Batten, S., Moffitt, S., Pegau, W., & Campbell, R. (2016). Plankton indices explain interannual variability in Prince William Sound herring first year growth. *Fisheries Oceanography, 25*, 420-432. doi:https://doi.org/10.1111/fog.12162

Baumgartner, T. (1992). Reconstruction of the history of the Pacific sardine and northern anchovy populations over the past two millenia from sediments of the Santa Barbara basin, California. *CalCOFI Rep, 33*, 24-40.

Beaudreau, A. H., Ward, E. J., Brenner, R. E., Shelton, A. O., Watson, J. T., Womack, J. C., . . . Williams, B. C. (2019). Thirty years of change and the future of Alaskan fisheries: Shifts in fishing participation and diversification in response to environmental, regulatory and economic pressures. *Fish and Fisheries, 20*, 601-619. doi:10.1111/faf.12364

Begon, M., Telfer, S., Smith, M. J., Burthe, S., Paterson, S., & Lambin, X. (2009). Seasonal host dynamics drive the timing of recurrent epidemics in a wildlife population. *Proceedings of the Royal Society B: Biological Sciences, 276*, 1603-1610. doi:10.1098/rspb.2008.1732

Bishop, M. A., & Eiler, J. H. (2018). Migration patterns of post-spawning Pacific herring in a subarctic sound. *Deep Sea Research Part II: Topical Studies in Oceanography, 147*, 108-115.

Bishop, M. A., Watson, J. T., Kuletz, K., & Morgan, T. (2015). Pacific herring (Clupea pallasii) consumption by marine birds during winter in Prince William Sound, Alaska. *Fisheries Oceanography, 24*, 1-13. doi:https://doi.org/10.1111/fog.12073

Bjørkvoll, E., Grøtan, V., Aanes, S., Sæther, B.-E., Engen, S., & Aanes, R. (2012). Stochastic population dynamics and life-history variation in marine fish species. *The American Naturalist, 180*, 372-387. doi:https://doi.org/10.1086/666983

Boehlert, G. W., & Mundy, B. C. (1988). Roles of behavioral and physical factors in larval and juvenile fish recruitment to estuarine nursery areas. *Proceedings of the American Fisheries Society Symposium, 3*, 1-67.

Boltaña, S., Sanhueza, N., Aguilar, A., Gallardo-Escarate, C., Arriagada, G., Valdes, J. A., . . . Quiñones, R. A. (2017). Influences of thermal environment on fish growth. *Ecology and Evolution, 7*, 6814-6825. doi:https://doi.org/10.1002/ece3.3239

Boulder, C. C. f. A. R. U. o. C. (2013). Reconstructed Sea Level (Ver. 1) [Data file]. Available from PO.DAAC, CA, USA. Accessed 2017-06-01 at http://dx.doi.org/10.5067/RECSL-000V1.

Briggs, C. J., Vredenburg, V. T., Knapp, R. A., & Rachowicz, L. J. (2005). Investigating the population-level effects of Chytridiomycosis: An emerging infectious disease of amphibians. *Ecology, 86*, 3149-3159. doi:10.1890/04-1428

Brooks, E. N., & Deroba, J. J. (2015). When "data" are not data: the pitfalls of post hoc analyses that use stock assessment model output. *Canadian Journal of Fisheries and Aquatic Sciences, 72*, 634-641.

Brooks, E. N., Thorson, J. T., Shertzer, K. W., Nash, R. D. M., Brodziak, J. K. T., Johnson, K. F., . . . White, J. (2019). Paulik revisited: Statistical framework and estimation performance of multistage recruitment functions. *Fisheries Research, 217*, 58-70. doi:https://doi.org/10.1016/j.fishres.2018.06.018

Brown, E. D., & Norcross, B. L. (2001). Effect of herring egg distribution and ecology on year-class strength and adult distribution: preliminary results from Prince William Sound. *Herring: Expectations for a new millennium. Alaska Sea Grant College Program, Report AK-SG-01-04, University of Alaska, Fairbanks*, 335-345.

Brunel, T., & Dickey-Collas, M. (2010). Effects of temperature and population density on von Bertalanffy growth parameters in Atlantic herring: a macro-ecological analysis. *Marine Ecology Progress Series, 405*, 15-28. doi:https://doi.org/10.3354/meps08491

Burge, C. A., Mark Eakin, C., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., . . . Harvell, C. D. (2014). Climate Change Influences on Marine Infectious Diseases: Implications for Management and Society. *Annual Review of Marine Science, 6*, 249-277. doi:10.1146/annurev-marine-010213-135029

CAPAM. (2017). In: Sharma, R., Porch, C., Babcock, E., Maunder, M., Punt, A. (Eds.), Recruitment: Theory, Estimation, and Application in Fishery Stock Assessment Models. Center for the Advancement of Population Assessment Methodology, Miami, FL, October 30th-November 3rd, 2017 8901 La Jolla Shores Dr., La Jolla, CA 92037.

doi:http://www.capamresearch.org/sites/default/files/CAPAM_Workshop_Series_Report_3.pdf.

Caputi, N., Fletcher, W., Pearce, A., & Chubb, C. (1996). Effect of the Leeuwin Current on the recruitment of fish and invertebrates along the Western Australian coast. *Marine and Freshwater Research, 47*, 147-155. doi:https://doi.org/10.1071/MF9960147

Carls, M. G., Rice, S. D., & Hose, J. E. (1999). Sensitivity of fish embryos to weathered crude oil: Part I. Low-level exposure during incubation causes malformations, genetic damage, and mortality in larval pacific herring (Clupea pallasi). *Environmental Toxicology and Chemistry, 18*, 481-493.

Checkley, D., Alheit, J., Oozeki, Y., & Roy, C. (2009). *Climate change and small pelagic fish*: Cambridge University Press Cambridge.

Chu, C. Y. C., Chien, H.-K., & Lee, R. D. (2008). Explaining the optimality of U-shaped age-specific mortality. *Theoretical Population Biology, 73*, 171-180. doi:https://doi.org/10.1016/j.tpb.2007.11.005

Clark, W. G. (1999). Effects of an erroneous natural mortality rate on a simple age-structured stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences, 56*, 1721-1731. doi:10.1139/f99-085

Cornwell, E. R., Bellmund, C. A., Groocock, G. H., Wong, P. T., Hambury, K. L., Getchell, R. G., & Bowser, P. R. (2013). Fin and gill biopsies are effective nonlethal samples for detection of Viral hemorrhagic septicemia virus genotype IVb. *Journal of Veterinary Diagnostic Investigation, 25*, 203-209. doi:10.1177/1040638713476865

Corten, A. (2001). Northern distribution of North Sea herring as a response to high water temperatures and/or low food abundance. *Fisheries Research, 50*, 189-204. doi:https://doi.org/10.1016/S0165-7836(00)00251-4

Crone, P. R., Maunder, M. N., Lee, H., & Piner, K. R. (2019). Good practices for including environmental data to inform spawner-recruit dynamics in integrated stock assessments: Small pelagic species case study. *Fisheries Research, 217*, 122-132.

Cushing, D. (1971). The dependence of recruitment on parent stock in different groups of fishes. *ICES Journal of Marine Science, 33*, 340-362. doi:https://doi.org/10.1093/icesjms/33.3.340

Cushing, D. (1990). Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in marine biology, 26*, 249-293. doi:https://doi.org/10.1016/S0065-2881(08)60202-3

Deriso, R. B., Maunder, M. N., & Pearson, W. H. (2008). Incorporating covariates into fisheries stock assessment models with application to Pacific herring. *Ecol Appl, 18*, 1270-1286. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18686586

Deriso, R. B., Maunder, M. N., & Skalski, J. R. (2007). Variance estimation in integrated assessment models and its importance for hypothesis testing. *Canadian Journal of Fisheries Aquatic Sciences, 64*, 187-197.

Deroba, J. J., Gaichas, S. K., Lee, M.-Y., Feeney, R. G., Boelke, D., & Irwin, B. J. (2018). The dream and the reality: meeting decision-making time frames while incorporating ecosystem and economic models into management strategy evaluation. *Canadian Journal of Fisheries and Aquatic Sciences, 76*, 1112-1133. doi:10.1139/cjfas-2018-0128

Deroba, J. J., & Schueller, A. M. (2013). Performance of stock assessments with misspecified age- and time-varying natural mortality. *Fisheries Research, 146*, 27-40. doi:https://doi.org/10.1016/j.fishres.2013.03.015

Dey, S., Delampady, M., & Gopalaswamy, A. M. (2019). Bayesian model selection for spatial capture–recapture models. *Ecology and Evolution, 9*, 11569-11583.

Deyle, E. R., Fogarty, M., Hsieh, C. H., Kaufman, L., MacCall, A. D., Munch, S. B., . . . Sugihara, G. (2013). Predicting climate effects on Pacific sardine. *Proc Natl Acad Sci U S A, 110*, 6430-6435. doi:https://doi.org/10.1073/pnas.1215506110

Di Lorenzo, E., Cobb, K. M., Furtado, J. C., Schneider, N., Anderson, B. T., Bracco, A., . . . Vimont, D. J. (2010). Central Pacific El Niño and decadal climate change in the North Pacific Ocean. *Nature Geoscience, 3*, 762-765. doi:10.1038/ngeo984

Di Lorenzo, E., Schneider, N., Cobb, K. M., Franks, P., Chhak, K., Miller, A. J., . . . Curchitser, E. (2008). North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophysical Research Letters, 35*.

Dickey-Collas, M., Nash, R. D., Brunel, T., Van Damme, C. J., Marshall, C. T., Payne, M. R., . . . Hatfield, E. M. (2010). Lessons learned from stock collapse and recovery of North Sea herring: a review. *ICES Journal of Marine Science*, 1875-1886. doi:https://doi.org/10.1093/icesjms/fsq033

Doll, P., Kaspar, F., & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology, 270*, 105-134. doi:https://doi.org/10.1016/S0022-1694(02)00283-4

Dormann, C. F., Calabrese, J. M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., . . . Hartig, F. (2018). Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecological monographs, 88*, 485-504. doi:10.1002/ecm.1309

Dorn, M., Aydin, K., Fissel, B., Jones, D., McCarthy, A., Palsson, W., & Spalinger, K. (2017). Assessment of the walleye pollock stock in the Gulf of Alaska. *NOAA*.

Elston, R. A., & Meyers, T. R. (2009). Effect of viral hemorrhagic septicemia virus on Pacific herring in Prince William Sound, Alaska, from 1989 to 2005. *Diseases of Aquatic Organisms, 83*, 223-246.

Essington, T. E., Moriarty, P. E., Froehlich, H. E., Hodgson, E. E., Koehn, L. E., Oken, K. L., . . . Stawitz, C. C. (2015). Fishing amplifies forage fish population collapses. *Proceedings of the National Academy of Sciences, 112*, 6648-6652. doi:https://doi.org/10.1073/pnas.1422020112

Fey, S. B., Siepielski, A. M., Nusslé, S., Cervantes-Yoshida, K., Hwan, J. L., Huber, E. R., . . . Carlson, S. M. (2015). Recent shifts in the occurrence, cause, and magnitude of animal mass mortality events. *Proceedings of the National Academy of Sciences, 112*, 1083. doi:10.1073/pnas.1414894112

Fortier, L., & Gagné, J. A. (1990). Larval Herring (Clupea harengus) Dispersion, Growth, and Survival in the St, Lawrence Estuary: Match/Mismatch or Membership/Vagrancy? *Canadian Journal of Fisheries and Aquatic Sciences, 47*, 1898-1912. doi:https://doi.org/10.1139/f90-214

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., . . . Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software, 27*, 233-249.

Fox, C., Paquet, P., & Reimchen, T. (2018). Pacific herring spawn events influence nearshore subtidal and intertidal species. *Marine Ecology Progress Series, 595*, 157-169. doi:https://doi.org/10.3354/meps12539

Fox, C. H., Paquet, P. C., & Reimchen, T. E. (2015). Novel species interactions: American black bears respond to Pacific herring spawn. *BMC ecology, 15*, 14. doi:https://doi.org/10.1186/s12898-015-0045-9

Fréon, P., Cury, P., Shannon, L., & Roy, C. (2005). Sustainable exploitation of small pelagic fish stocks challenged by environmental and ecosystem changes: a review. *Bulletin of Marine Science, 76*, 385-462.

Gargett, A. E. (1997). The optimal stabilitywindow': a mechanism underlying decadal fluctuations in North Pacific salmon stocks? *Fisheries Oceanography, 6*, 109-117. doi:https://doi.org/10.1046/j.1365-2419.1997.00033.x

Garver, K. A., Traxler, G. S., Hawley, L. M., Richard, J., Ross, J., & Lovy, J. (2013). Molecular epidemiology of viral haemorrhagic septicaemia virus (VHSV) in British Columbia, Canada, reveals transmission from wild to farmed fish. *Diseases of Aquatic Organisms, 104*, 93-104. Retrieved from https://www.int-res.com/abstracts/dao/v104/n2/p93-104/

Gauvreau, A. M., Lepofsky, D., Rutherford, M., & Reid, M. (2017). "Everything revolves around the herring" the Heiltsuk–herring relationship through time. *Ecology and Society, 22*. doi:https://doi.org/10.5751/ES-09201-220210

Gelfand, A. E., & Ghosh, S. K. (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika, 85*, 1-11.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014a). *Bayesian Data Analysis* (Vol. 2). Chapman: Boca Raton, FL.

Gelman, A., Hwang, J., & Vehtari, A. (2014b). Understanding predictive information criteria for Bayesian models. *Statistics computing, 24*, 997-1016.

Gende, S. M., & Sigler, M. F. (2006). Persistence of forage fish 'hot spots' and its association with foraging Steller sea lions (Eumetopias jubatus) in southeast Alaska. *Deep Sea Research Part II: Topical Studies in Oceanography, 53*, 432-441. doi:https://doi.org/10.1016/j.dsr2.2006.01.005

Gilbert, D. (1997). Towards a new recruitment paradigm for fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences, 54*, 969-977. doi:https://doi.org/10.1139/f96-272

Gilbert, L., Brülisauer, F., Willoughby, K., & Cousens, C. (2020). Identifying Environmental Risk Factors for Louping Ill Virus Seroprevalence in Sheep and the Potential to Inform Wildlife Management Policy. *Frontiers in Veterinary Science, 7*, 377. Retrieved from https://www.frontiersin.org/article/10.3389/fvets.2020.00377

Gislason, H., Daan, N., Rice, J. C., & Pope, J. G. (2010). Size, growth, temperature and the natural mortality of marine fish. *Fish and Fisheries, 11*, 149-158. doi:https://doi.org/10.1111/j.1467-2979.2009.00350.x

Gray, B. P., Bishop, M. A., & Powers, S. (2019). Structure of winter groundfish feeding guilds in Pacific herring Clupea pallasii and walleye pollock Gadus chalcogrammus nursery fjords. *Journal of Fish Biology, 95*, 527-539. doi:https://doi.org/10.1111/jfb.13984

GRDC. (2014). Global Freshwater Fluxes into the World Oceans. Koblenz: Federal Institute of Hydrology (BfG). Retrieved from http://www.bafg.de/GRDC/EN/03_dtprdcts/31_FWFLX/freshflux_node.html.

Groner, M. L., Hoenig, J. M., Pradel, R., Choquet, R., Vogelbein, W. K., Gauthier, D. T., & Friedrichs, M. A. M. (2018). Dermal mycobacteriosis and warming sea surface temperatures are associated with elevated mortality of striped bass in Chesapeake Bay. *Ecology and Evolution, 8*, 9384-9397. doi:https://doi.org/10.1002/ece3.4462

Hamada, S. (2015). Ainu geographic games and an indigenous history of the herring in Hokkaido, Japan. *The Canadian Journal of Native Studies, 35*, 2.

Hamlington, B. D., Leben, R. R., Strassburg, M. W., & Kim, K.-Y. (2014). Cyclostationary empirical orthogonal function sea-level reconstruction. *Geoscience Data Journal, 1*, 13-19. doi:https://doi.org/10.1002/gdj3.6

Hammill, M., Stenson, G., Proust, F., Carter, P., & McKinnon, D. (2007). Feeding by grey seals in the Gulf of St. Lawrence and around Newfoundland. *NAMMCO Scientific Publications, 6*, 135-152. doi:https://doi.org/10.7557/3.2729

Hardman-Mountford, N., Richardson, A., Boyer, D., Kreiner, A., & Boyer, H. (2003). Relating sardine recruitment in the Northern Benguela to satellite-derived sea surface height using a neural network pattern recognition approach. *Progress in Oceanography, 59*, 241-255. doi:https://doi.org/10.1016/j.pocean.2003.07.005

Hare, S. R., & Mantua, N. J. (2000). Empirical evidence for North Pacific regime shifts in 1977 and 1989. *Progress in Oceanography, 47*, 103-145.

Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L., & Rosati, R. A. (1982). Evaluating the yield of medical tests. *Jama, 247*, 2543-2546. doi:https://doi.org/10.1001/jama.247.18.2543

Hart, L. M., MacKenzie, A., Purcell, M. K., Powers, R. L., & Hershberger, P. K. (2017). Optimization of a Plaque Neutralization Test (PNT) to Identify the Exposure History of Pacific Herring to Viral Hemorrhagic Septicemia Virus (VHSV). *Journal of Aquatic Animal Health, 29*, 74-82. doi:10.1080/08997659.2017.1285369

Harvell, D., Aronson, R., Baron, N., Connell, J., Dobson, A., Ellner, S., . . . Ward, J. (2004). The rising tide of ocean diseases: unsolved problems and research priorities. *Frontiers in Ecology and the Environment, 2*, 375-382. doi:https://doi.org/10.1890/1540-9295(2004)002[0375:TRTOOD]2.0.CO;2

Haught, S., & Moffitt, S. (2018). Scales as growth history records for Pacific herring in Prince William Sound. Accessed 03/25/2020 at. Retrieved from https://portal.aoos.org/gulf-of-alaska#metadata/35fd35d8-f6f1-4762-9cf0-8e2e970755c4/project/folder_metadata/3115471.

Hay, D. (1985). Reproductive biology of Pacific herring (Clupea harengus pallasi). *Canadian Journal of Fisheries and Aquatic Sciences, 42*, s111-s126. doi:https://doi.org/10.1139/f85-267

Hay, D., & McCarter, P. (1997). Larval distribution, abundance, and stock structure of British Columbia herring. *Journal of Fish Biology, 51*, 155-175. doi:https://doi.org/10.1111/j.1095-8649.1997.tb06098.x

Hay, D. E., Rose, K. A., Schweigert, J., & Megrey, B. A. (2008). Geographic variation in North Pacific herring populations: Pan-Pacific comparisons and implications for climate change impacts. *Progress in Oceanography, 77*, 233-240. doi:https://doi.org/10.1016/j.pocean.2008.03.015

Hay, D. E., Toresen, R., Stephenson, R., Thompson, M., Claytor, R., Funk, F., . . . Wheeler, J. (2001). *Taking stock: an inventory and review of world herring stocks in 2000*. Paper presented at the Lowell Wakefield Symposium, Anchorage, AK.

Haydon, D. T., Cleaveland, S., Taylor, L. H., & Laurenson, M. K. (2002). Identifying reservoirs of infection: a conceptual and practical challenge. *Emerging infectious diseases, 8*, 1468-1473. doi:10.3201/eid0812.010317

He, X., Ralston, S., & MacCall, A. D. (2011). Interactions of age-dependent mortality and selectivity functions in age-based stock assessment models. *Fishery Bulletin, 109*, 198-216.

Hershberger, P., Gregg, J., Pacheco, C., Winton, J., Richard, J., & Traxler, G. (2007). Larval Pacific herring, Clupea pallasii (Valenciennes), are highly susceptible to viral haemorrhagic septicaemia and survivors are partially protected after their metamorphosis to juveniles. *Journal of Fish Diseases, 30*, 445-458.

Hershberger, P. K. (2012). *Ichthyophonus* Disease (Ichthyophoniasis). In: *Fish Health Section Blue Book: Suggested Procedures for the Detection and Identification of Certain Finfish and Shellfish Pathogens*. In: American Fisheries Society Bethesda, Maryland.

Hershberger, P. K., Garver, K. A., & Winton, J. R. (2015). Principles underlying the epizootiology of viral hemorrhagic septicemia in Pacific herring and other fishes throughout the North Pacific Ocean. *Canadian Journal of Fisheries and Aquatic Sciences, 73*, 853-859. doi:10.1139/cjfas-2015-0417

Hershberger, P. K., Gregg, J. L., Grady, C. A., Taylor, L., & Winton, J. R. (2010). Chronic and persistent viral hemorrhagic septicemia virus infections in Pacific herring. *Diseases of Aquatic Organisms, 93*, 43-49. Retrieved from https://www.int-res.com/abstracts/dao/v93/n1/p43-49/

Hershberger, P. K., Gregg, J. L., Hart, L. M., Moffitt, S., Brenner, R., Stick, K., . . . Meyers, T. R. (2016). The parasite Ichthyophonus sp. in Pacific herring from the coastal NE Pacific. *Journal of Fish Diseases, 39*, 395-410. doi:https://doi.org/10.1111/jfd.12370

Hershberger, P. K., Kocan, R. M., Elder, N. E., Meyers, T. R., & Winton, J. R. (1999). Epizootiology of viral hemorrhagic septicemia virus in Pacific herring from the spawn-on-kelp fishery in Prince William Sound, Alaska, USA. *Diseases of Aquatic Organisms, 37*, 23-31. Retrieved from https://www.int-res.com/abstracts/dao/v37/n1/p23-31/

Hershberger, P. K., Purcell, M. K., Hart, L. M., Gregg, J. L., Thompson, R. L., Garver, K. A., & Winton, J. R. (2013). Influence of temperature on viral hemorrhagic septicemia (Genogroup IVa) in Pacific herring, Clupea pallasii Valenciennes. *Journal of Experimental Marine Biology and Ecology, 444*, 81-86. doi:https://doi.org/10.1016/j.jembe.2013.03.006

Hilborn, R. (2002). The dark side of reference points. *Bulletin of Marine Science, 70*, 403-408.

Hilborn, R., Amoroso, R. O., Bogazzi, E., Jensen, O. P., Parma, A. M., Szuwalski, C., & Walters, C. J. (2017). When does fishing forage species affect their predators? *Fisheries Research, 191*, 211-221. doi:https://doi.org/10.1016/j.fishres.2017.01.008

Hilborn, R., Hively, D. J., Jensen, O. P., & Branch, T. A. (2014). The dynamics of fish populations at low abundance and prospects for rebuilding and recovery. *ICES Journal of Marine Science, 71*, 2141-2151.

Hilborn, R., & Walters, C. J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty.* Norwell, Massachusetts, USA.

Hjort, J. (1914). Fluctuations in the great fisheries of northern Europe viewed in the light of biological research. *Rapports et Procès-Verbaux des Réunions Conseil International pour l'Exploration de la Mer, 20*, 228.

Hoenig, J. M., Groner, M. L., Smith, M. W., Vogelbein, W. K., Taylor, D. M., Landers Jr, D. F., . . . Shields, J. D. (2017). Impact of disease on the survival of three commercially fished species. *Ecological Applications, 27*, 2116-2127. doi:https://doi.org/10.1002/eap.1595

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65-70.

Hooten, M. B., & Hobbs, N. (2015). A guide to Bayesian model selection for ecologists. *Ecological monographs, 85*, 3-28.

Hopkins, S. R., Fleming-Davies, A. E., Belden, L. K., & Wojdak, J. M. (2020). Systematic review of modelling assumptions and empirical evidence: Does parasite transmission increase nonlinearly with host density? *Methods in Ecology and Evolution, 11*, 476-486. doi:https://doi.org/10.1111/2041-210X.13361

Hsieh, C.-h., Reiss, C. S., Hunter, J. R., Beddington, J. R., May, R. M., & Sugihara, G. (2006). Fishing elevates variability in the abundance of exploited species. *Nature, 443*, 859. doi:https://doi.org/10.1038/nature05232

Hulson, P.-J. F., Miller, S. E., Dressel, S. C., Quinn II, T. J., & Van Kirk, K. (2018). Time-dependent Parameterization of Natural Mortality, Maturity, and Fishing Selectivity: A Case Study Using the Pacific Decadal Oscillation for Pacific Herring (Clupea pallasii) in Sitka Sound, Alaska. In F. J. Mueter, M. R. Baker, S. C. Dressel, & A. B. Hollowed (Eds.), *Impacts of a Changing Environment on the Dynamics of High-latitude Fish and Fisheries*. Alaska Sea Grant: University of Alaska Fairbanks.

Hulson, P.-J. F., Miller, S. E., Quinn, T. J., Marty, G. D., Moffitt, S. D., & Funk, F. (2007). Data conflicts in fishery models: incorporating hydroacoustic data into the Prince William Sound Pacific herring assessment model. *ICES Journal of Marine Science, 65*, 25-43.

Hutchings, J. A. (2000). Collapse and recovery of marine fishes. *Nature, 406*, 882-885.

Hutchings, J. A. (2001). Influence of population decline, fishing, and spawner variability on the recovery of marine fishes. *Journal of Fish Biology, 59*, 306-322. doi:https://doi.org/10.1111/j.1095-8649.2001.tb01392.x

Hutchings, J. A., & Reynolds, J. D. (2004). Marine fish population collapses: Consequences for recovery and extinction risk. *Bioscience, 54*, 297-309. doi:https://doi.org/10.1641/0006-3568(2004)054[0297:MFPCCF]2.0.CO;2

Ianelli, J., Holsman, K. K., Punt, A. E., & Aydin, K. (2016). Multi-model inference for incorporating trophic and climate uncertainty into stock assessments. *Deep Sea Research Part II: Topical Studies in Oceanography, 134*, 379-389. doi:https://doi.org/10.1016/j.dsr2.2015.04.002

Ibrahim, J. G., & Laud, P. W. (1994). A Predictive Approach to the Analysis of Designed Experiments. *Journal of the American Statistical Association, 89*, 309-319. doi:10.1080/01621459.1994.10476472

ICES. (2014). *Report of the Herring Assessment Working Group for the Area South of 62ºN (HAWG)*. Retrieved from ICES HQ, Copenhagen, Denmark: http://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/acom/2014/HAWG/01%20HAWG%20Report%202014.pdf

Inchausti, P., & Halley, J. (2003). On the relation between temporal variability and persistence time in animal populations. *Journal of Animal Ecology, 72*, 899-908. doi:https://doi.org/10.1046/j.1365-2656.2003.00767.x

Ishwaran, H., & Kogalur, U. B. (2016). Random Forests for Survival, Regression, and Classification (RF-SRC) (Version R package version 2.2.0).

Ishwaran, H., Kogalur, U. B., Blackstone, E. H., & Lauer, M. S. (2008). Random survival forests. *The annals of applied statistics*, 841-860. doi:https://doi.org/10.1214/08-AOAS169

Ito, S., Rose, K. A., Megrey, B. A., Schweigert, J., Hay, D., Werner, F. E., & Aita, M. N. (2015). Geographic variation in Pacific herring growth in response to regime shifts in the North Pacific Ocean. *Progress in Oceanography, 138*, 331-347. doi:https://doi.org/10.1016/j.pocean.2015.05.022

Jiao, Y., Smith, E. P., O'Reilly, R., & Orth, D. J. (2012). Modelling non-stationary natural mortality in catch-at-age models. *ICES Journal of Marine Science, 69*, 105-118. doi:10.1093/icesjms/fsr184

Johnson, K. F., Monnahan, C. C., McGilliard, C. R., Vert-pre, K. A., Anderson, S. C., Cunningham, C. J., . . . Punt, A. E. (2014). Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. *ICES Journal of Marine Science, 72*, 137-150. doi:10.1093/icesjms/fsu055

Jones, R., Rigg, C., & Pinkerton, E. (2017). Strategies for assertion of conservation and local management rights: A Haida Gwaii herring story. *Marine Policy, 80*, 154-167. doi:https://doi.org/10.1016/j.marpol.2016.09.031

Kaeriyama, M., Nakamura, M., Yamaguchi, M., Ueda, H., Anma, G., Takagi, S., . . . Myers, K. W. (2000). Feeding ecology of sockeye and pink salmon in the Gulf of Alaska. *N. Pac. Anadr. Fish Comm. Bull, 2*, 55-63.

Kai, M., & Yokoi, H. (2019). Performance evaluation of information criteria for estimating a shape parameter in a Bayesian state-space biomass dynamics model. *Fisheries Research, 219*. doi:https://doi.org/10.1016/j.fishres.2019.105326

Kaplan, E. L., & Meier, P. (1958). Nonparametric-Estimation from Incomplete Observations. *Journal of the American Statistical Association, 53*, 457-481. doi:https://doi.org/10.1080/01621459.1958.10501452

Kemp, I. M., Beauchamp, D. A., Sweeting, R., & Cooper, C. (2013). Potential for competition among herring and juvenile salmon species in Puget Sound, Washington. *North Pacific Anadromous Fish Commission technical report*, 139-143.

Kilduff, D. P., Di Lorenzo, E., Botsford, L. W., & Teo, S. L. H. (2015). Changing central Pacific El Niños reduce stability of North American salmon survival rates. *Proceedings of the National Academy of Sciences, 112*, 10962-10966. doi:10.1073/pnas.1503190112

Klepac, P., & Caswell, H. (2011). The stage-structured epidemic: linking disease and demography with a multi-state matrix approach model. *Theoretical Ecology, 4*, 301-319. doi:10.1007/s12080-010-0079-8

Kocan, R. M., Hershberger, P., Mehl, T., Elder, N., Bradley, M., Wildermuth, D., & Stick, K. (1999). Pathogenicity of Ichthyophonus hoferi for laboratory-reared Pacific herring Clupea pallasi and its early appearance in wild Puget Sound herring. *Diseases of Aquatic Organisms, 35*, 23-29. Retrieved from https://www.int-res.com/abstracts/dao/v35/n1/p23-29/

Kristensen, K., Nielsen, A., Berg, C., Skaug, H., & Bell, B. (2016). TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software, 70*, 1-21. doi:10.18637/jss.v070.i05

Lam, M. E., Pitcher, T. J., Surma, S., Scott, J., Kaiser, M., White, A. S., . . . Ward, L. M. (2019). Value-and ecosystem-based management approach: the Pacific herring fishery conflict. *Marine Ecology Progress Series, 617*, 341-364. doi:https://doi.org/10.3354/meps12972

Lance, M. M., Chang, W.-Y., Jeffries, S. J., Pearson, S. F., & Acevedo-Gutiérrez, A. (2012). Harbor seal diet in northern Puget Sound: implications for the recovery of depressed fish

stocks. *Marine Ecology Progress Series, 464*, 257-271. doi:https://doi.org/10.3354/meps09880

Laud, P. W., & Ibrahim, J. G. (1995). Predictive Model Selection. *Journal of the Royal Statistical Society: Series B (Methodological), 57*, 247-262. doi:https://doi.org/10.1111/j.2517-6161.1995.tb02028.x

Lee, H.-H., Maunder, M. N., Piner, K. R., & Methot, R. D. (2012). Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? *Fisheries Research, 125-126*, 254-261. doi:https://doi.org/10.1016/j.fishres.2012.03.001

Lee, H.-H., Maunder, M. N., Piner, K. R., & Methot, R. D. J. F. R. (2011). Estimating natural mortality within a fisheries stock assessment model: an evaluation using simulation analysis based on twelve stock assessments. *109*, 89-94.

Lindegren, M., Checkley, D. M., Rouyer, T., MacCall, A. D., & Stenseth, N. C. (2013). Climate, fishing, and fluctuations of sardine and anchovy in the California Current. *Proc Natl Acad Sci U S A, 110*, 13672-13677. doi:https://doi.org/10.1073/pnas.1305733110

Litzow, M. A., Ciannelli, L., Puerta, P., Wettstein, J. J., Rykaczewski, R. R., & Opiekun, M. (2018). Non-stationary climate–salmon relationships in the Gulf of Alaska. *Proceedings of the Royal Society B, 285*, 20181855.

Litzow, M. A., Ciannelli, L., Puerta, P., Wettstein, J. J., Rykaczewski, R. R., & Opiekun, M. (2019). Nonstationary environmental and community relationships in the North Pacific Ocean. *Ecology, 100*, e02760. doi:https://doi.org/10.1002/ecy.2760

Litzow, M. A., Hunsicker, M. E., Bond, N. A., Burke, B. J., Cunningham, C. J., Gosselin, J. L., . . . Zador, S. G. (2020). The changing physical and ecological meanings of North Pacific Ocean climate indices. *Proceedings of the National Academy of Sciences, 117*, 7665-7671. doi:https://doi.org/10.1073/pnas.1921266117

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of Fish Biology, 49*, 627-642. doi:https://doi.org/10.1111/j.1095-8649.1996.tb00060.x

MacCall, A. D. (1990). *Dynamic geography of marine fish populations*. Seattle, WA: Washington Sea Grant Program.

Magera, A. M., Flemming, J. E. M., Kaschner, K., Christensen, L. B., & Lotze, H. K. (2013). Recovery trends in marine mammal populations. *PloS one, 8*, e77908. doi:https://doi.org/10.1371/journal.pone.0077908

Mantua, N. J., & Hare, S. R. (2002). The Pacific decadal oscillation. *Journal of oceanography, 58*, 35-44.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., & Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the american Meteorological Society, 78*, 1069-1080.

Martell, S., & Stewart, I. (2014). Towards defining good practices for modeling time-varying selectivity. *Fisheries Research, 158*, 84-95. doi:https://doi.org/10.1016/j.fishres.2013.11.001

Marty, G. D., Freiberg, E. F., Meyers, T. R., Wilcock, J., Farver, T. B., & Hinton, D. E. (1998). Viral hemorrhagic septicemia virus, Ichthyophonus hoferi, and other causes of morbidity in Pacific herring Clupea pallasi spawning in Prince William Sound, Alaska, USA. *Diseases of Aquatic Organisms, 32*, 15-40.

Marty, G. D., Hulson, P.-J. F., Miller, S. E., Quinn II, T. J., Moffitt, S. D., & Merizon, R. A. (2010). Failure of population recovery in relation to disease in Pacific herring. *Diseases of Aquatic Organisms, 90*, 1-14.

Marty, G. D., Ii, T. J. Q., Carpenter, G., Meyers, T. R., & Willits, N. H. (2003). Role of disease in abundance of a Pacific herring (Clupea pallasi) population. *Canadian Journal of Fisheries and Aquatic Sciences, 60*, 1258-1265.

Maunder, M. N., & Deriso, R. B. (2003). Estimation of recruitment in catch-at-age models. *Canadian Journal of Fisheries and Aquatic Sciences, 60*, 1204-1216.

Maunder, M. N., & Deriso, R. B. (2010). Dealing with missing covariate data in fishery stock assessment models. *Fisheries Research, 101*, 80-86.

Maunder, M. N., Deriso, R. B., & Hanson, C. H. (2015). Use of state-space population dynamics models in hypothesis testing: advantages over simple log-linear regressions for modeling survival, illustrated with application to longfin smelt (Spirinchus thaleichthys). *Fisheries Research, 164*, 102-111.

Maunder, M. N., & Thorson, J. T. (2019). Modeling temporal variation in recruitment in fisheries stock assessment: A review of theory and practice. *Fisheries Research, 217*, 71-86. doi:https://doi.org/10.1016/j.fishres.2018.12.014

Maunder, M. N., & Watters, G. M. (2003). A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and example. *Fishery Bulletin, 101*, 89-99.

May, R. M., & Anderson, R. M. (1979). Population biology of infectious diseases: Part II. *Nature, 280*, 455-461. doi:10.1038/280455a0

Maynard, J., van Hooidonk, R., Harvell, C. D., Eakin, C. M., Liu, G., Willis, B. L., . . . Shields, J. D. (2016). Improving marine disease surveillance through sea temperature monitoring, outlooks and projections. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*, 20150208. doi:10.1098/rstb.2015.0208

McCallum, H., Barlow, N., & Hone, J. (2001). How should pathogen transmission be modelled? *Trends in Ecology & Evolution, 16*, 295-300. doi:https://doi.org/10.1016/S0169-5347(01)02144-9

McClatchie, S., Hendy, I., Thompson, A., & Watson, W. (2017). Collapse and recovery of forage fish populations prior to commercial exploitation. *Geophysical Research Letters, 44*, 1877-1885. doi:https://doi.org/10.1002/2016GL071751

McKechnie, I., Lepofsky, D., Moss, M. L., Butler, V. L., Orchard, T. J., Coupland, G., . . . Lertzman, K. (2014). Archaeological data provide alternative hypotheses on Pacific herring (*Clupea pallasii*) distribution, abundance, and variability. *Proceedings of the National Academy of Sciences, 111*, E807-E816. doi:10.1073/pnas.1316072111

McVicar, A. (2011). *Ichthyophonus. In Fish Diseases and Disorders, Vol. 3: Viral, Bacterial, and Fungal Infections,* (P. Woo & D. Bruno Eds. 2nd ed.). Cambridge, MA: CAB.

McWhinnie, S. F. (2009). The tragedy of the commons in international fisheries: An empirical examination. *Journal of Environmental Economics and Management, 57*, 321-333. doi:https://doi.org/10.1016/j.jeem.2008.07.008

Melvin, G. D., & Stephenson, R. L. (2006). The dynamics of a recovering fish stock: Georges Bank herring. *ICES Journal of Marine Science, 64*, 69-82. doi:https://doi.org/10.1093/icesjms/fsl018

Menzies, C. R. (2016). *People of the Saltwater: An Ethnography of Git lax m'oon*: U of Nebraska Press.

Metcalf, C. J. E., Lessler, J., Klepac, P., Morice, A., Grenfell, B. T., & Bjørnstad, O. N. (2012). Structured models of infectious disease: Inference with discrete data. *Theoretical Population Biology, 82*, 275-282. doi:https://doi.org/10.1016/j.tpb.2011.12.001

Meyers, T., Short, S., Lipson, K., Batts, W., Winton, J., Wilcock, J., & Brown, E. (1994). Association of viral hemorrhagic septicemia virus with epizootic hemorrhages of the skin in Pacific herring Clupea harengus pallasi from Prince William Sound and Kodiak Island, Alaska, USA. *Diseases of Aquatic Organisms, 19*, 27-37.

Meyers, T. R., & Winton, J. R. (1995). Viral hemorrhagic septicemia virus in North America. *Annual Review of Fish Diseases, 5*, 3-24.

Miller, K. M., Teffer, A., Tucker, S., Li, S., Schulze, A. D., Trudel, M., . . . Hinch, S. G. (2014). Infectious disease, shifting climates, and opportunistic predators: cumulative factors potentially impacting wild salmon declines. *Evolutionary Applications, 7*, 812-855. doi:10.1111/eva.12164

Miller, T. J., Hare, J. A., & Alade, L. A. (2016). A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to southern New England yellowtail flounder. *Canadian Journal of Fisheries and Aquatic Sciences, 73*, 1261-1270.

Mohn, R. (1999). The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. *ICES Journal of Marine Science, 56*, 473-488. doi:10.1006/jmsc.1999.0481

Möllmann, C., Kornilovs, G., Fetter, M., & Köster, F. W. (2005). Climate, zooplankton, and pelagic fish growth in the central Baltic Sea. *ICES Journal of Marine Science, 62*, 1270-1280.

Monnahan, C. C., Branch, T. A., Thorson, J. T., Stewart, I. J., & Szuwalski, C. S. (2019). Overcoming long Bayesian run times in integrated fisheries stock assessments. *ICES Journal of Marine Science, 76*, 1477-1488.

Monnahan, C. C., & Kristensen, K. (2018). No-U-turn sampling for fast Bayesian inference in ADMB and TMB: Introducing the adnuts and tmbstan R packages. *PloS one, 13*.

Monnahan, C. C., Thorson, J. T., & Branch, T. A. (2017). Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods in Ecology and Evolution, 8*, 339-348. doi:10.1111/2041-210X.12681

Moran, J. R., Heintz, R. A., Straley, J. M., & Vollenweider, J. J. (2018a). Regional variation in the intensity of humpback whale predation on Pacific herring in the Gulf of Alaska. *Deep Sea Research Part II: Topical Studies in Oceanography, 147*, 187-195. doi:https://doi.org/10.1016/j.dsr2.2017.07.010

Moran, J. R., O'Dell, M. B., Arimitsu, M. L., Straley, J. M., & Dickson, D. M. (2018b). Seasonal distribution of Dall's porpoise in Prince William Sound, Alaska. *Deep Sea Research Part II: Topical Studies in Oceanography, 147*, 164-172. doi:https://doi.org/10.1016/j.dsr2.2017.11.002

Mueter, F. J., Ware, D. M., & Peterman, R. M. (2002). Spatial correlation patterns in coastal environmental variables and survival rates of salmon in the north-east Pacific Ocean. *Fisheries Oceanography, 11*, 205-218. doi:10.1046/j.1365-2419.2002.00192.x

Mullon, C., Fréon, P., & Cury, P. (2005). The dynamics of collapse in world fisheries. *Fish and Fisheries, 6*, 111-120. doi:https://doi.org/10.1111/j.1467-2979.2005.00181.x

Muradian, M. L., Branch, T. A., Moffitt, S. D., & Hulson, P.-J. F. (2017). Bayesian stock assessment of Pacific herring in Prince William Sound, Alaska. *PloS one, 12*, e0172153.

Murray, A. G. (2009). Using simple models to review the application and implications of different approaches used to simulate transmission of pathogens among aquatic animals. *Preventive Veterinary Medicine, 88*, 167-177. doi:https://doi.org/10.1016/j.prevetmed.2008.09.006

Myers, R. A. (1998). When do environment-recruitment correlations work? *Reviews in Fish Biology and Fisheries, 8*, 285-305. doi:Doi 10.1023/A:1008828730759

Myers, R. A., & Barrowman, N. J. (1996). Is fish recruitment related to spawner abundance? *Fishery Bulletin, 94*, 707-724.

Nagasawa, K. (2001). Long-term variations in abundance of Pacific herring (Clupea pallasi) in Hokkaido and Sakhalin related to changes in environmental conditions. *Progress in Oceanography, 49*, 551-564. doi:Doi 10.1016/S0079-6611(01)00040-4

Naumenko, N. I. (2002). *Temporal variations in size-at-age of the western Bering Sea herring*. Retrieved from Globec International Program on Climate Change and Carrying Capacity, North Pacific Marine Science Organization (PICES), Sidney, B.C., Canada: https://pices.int/publications/scientific_reports/Report20/Rep20_REX_Naumenko.pdf

Needle, C. L. (2001). Recruitment models: diagnosis and prognosis. *Reviews in Fish Biology and Fisheries, 11*, 95-111. doi:10.1023/A:1015208017674

Neubauer, P., Jensen, O. P., Hutchings, J. A., & Baum, J. K. (2013). Resilience and recovery of overexploited marine populations. *Science, 340*, 347-349. doi:10.1126/science.1230441

Okamoto, D. K., Poe, M. R., Francis, T. B., Punt, A. E., Levin, P. S., Shelton, A. O., . . . Woodruff, J. (2020). Attending to spatial social–ecological sensitivities to improve trade-off analysis in natural resource management. *Fish and Fisheries, 21*, 1-12. doi:10.1111/faf.12409

Oken, K. L., Essington, T. E., & Fu, C. (2018). Variability and stability in predation landscapes: A cross-ecosystem comparison on the potential for predator control in temperate marine ecosystems. *Fish and Fisheries, 19*, 489-501. doi:https://doi.org/10.1111/faf.12269

Okey, T. A., & Pauly, D. (1999). *Trophic mass-balance model of Alaska's Prince William Sound ecosystem, for the post-spill period 1994-1996. Final Report, Restoration Project 99330-1, Exxon Valdez Oil Spill Trustee Council, Anchorage*. Retrieved from https://evostc.state.ak.us/media/2274/1999-99330baa-final.pdf

Olesiuk, P. F. (1999). *An assessment of the status of harbour seals (Phoca vitulina) in British Columbia* (ISSN 1480–4883). Retrieved from DFO Canadian Science Advisory Secretariat Research Document, 99/33: https://waves-vagues.dfo-mpo.gc.ca/Library/244725.pdf

Olesiuk, P. F. (2008). *Preliminary assessment of the recovery potential of northern fur seals (Callorhinus ursinus) in British Columbia*. Retrieved from DFO Canadian Science Advisory Secretariat, 2007/076:

Overholtz, W., & Link, J. (2006). Consumption impacts by marine mammals, fish, and seabirds on the Gulf of Maine–Georges Bank Atlantic herring (*Clupea harengus*) complex during the years 1977–2002. *ICES Journal of Marine Science, 64*, 83-96. doi:https://doi.org/10.1093/icesjms/fsl009

Pearce, A., & Phillips, B. (1988). ENSO events, the Leeuwin Current, and larval recruitment of the western rock lobster. *ICES Journal of Marine Science, 45*, 13-21. doi:https://doi.org/10.1093/icesjms/45.1.13

Pearson, W., Elston, R., Bienert, R., Drum, A., & Antrim, L. (1999). Why did the Prince William Sound, Alaska, Pacific herring (*Clupea pallasi*) fisheries collapse in 1993 and 1994?

Review of hypotheses. *Canadian Journal of Fisheries and Aquatic Sciences, 56*, 711-737. doi:https://doi.org/10.1139/f98-207

Pearson, W. H., Deriso, R. B., Elston, R. A., Hook, S. E., Parker, K. R., & Anderson, J. W. (2012). Hypotheses concerning the decline and poor recovery of Pacific herring in Prince William Sound, Alaska. *Reviews in Fish Biology and Fisheries, 22*, 95-135.

Peck, M. A., Neuenfeldt, S., Essington, T. E., Trenkel, V. M., Takasuka, A., Gislason, H., . . . Vestergaard, N. (2014). Forage fish interactions: a symposium on "Creating the tools for ecosystem-based management of marine resources". *ICES Journal of Marine Science: Journal du Conseil, 71*, 1-4.

Peck, M. A., Neuenfeldt, S., Essington, T. E., Trenkel, V. M., Takasuka, A., Gislason, H., . . . Rice, J. C. (2014). Forage Fish Interactions: a symposium on "Creating the tools for ecosystem-based management of marine resources". *ICES Journal of Marine Science, 71*, 1-4. doi:https://doi.org/10.1093/icesjms/fst174

Peel, A. J., Pulliam, J. R. C., Luis, A. D., Plowright, R. K., O'Shea, T. J., Hayman, D. T. S., . . . Restif, O. (2014). The effect of seasonal birth pulses on pathogen persistence in wild mammal populations. *Proceedings of the Royal Society B: Biological Sciences, 281*, 20132962. doi:10.1098/rspb.2013.2962

Pepin, P. (2015). Reconsidering the impossible—linking environmental drivers to growth, mortality, and recruitment of fish. *Canadian Journal of Fisheries and Aquatic Sciences, 73*, 205-215. doi:https://doi.org/10.1139/cjfas-2015-0091

Petitgas, P., Secor, D. H., McQuinn, I., Huse, G., & Lo, N. (2010). Stock collapses and their recovery: mechanisms that establish and maintain life-cycle closure in space and time. *ICES Journal of Marine Science, 67*, 1841-1848. doi:https://doi.org/10.1093/icesjms/fsq082

Piironen, J., & Vehtari, A. (2017). Comparison of Bayesian predictive methods for model selection. *Statistics and Computing, 27*, 711-735.

Pikitch, E., Boersma, P. D., Boyd, I., Conover, D., Cury, P., Essington, T., . . . Pauly, D. (2012). Little fish, Big impact: Managing a crucial link in ocean food webs. Retrieved from https://www.lenfestocean.org/~/media/legacy/Lenfest/PDFs/littlefishbigimpact_revised_12june12.pdf?la=en

Pikitch, E. K., Rountos, K. J., Essington, T. E., Santora, C., Pauly, D., Watson, R., . . . Munch, S. B. (2014). The global contribution of forage fish to marine fisheries and ecosystems. *Fish and Fisheries, 15*, 43-64. doi:https://doi.org/10.1111/faf.12004

Pinsky, M. L., & Byler, D. (2015). Fishing, fast growth and climate variability increase the risk of collapse. *Proceedings of the Royal Society B, 282*, 20151053. doi:https://doi.org/10.1098/rspb.2015.1053

Pinsky, M. L., Jensen, O. P., Ricard, D., & Palumbi, S. R. (2011). Unexpected patterns of fisheries collapse in the world's oceans. *Proc Natl Acad Sci U S A, 108*, 8317-8322. doi:https://doi.org/10.1073/pnas.1015313108

Pitcher, T. J. (1995). The impact of pelagic fish behaviour on fisheries. *Scientia Marina (Espana), 59*, 295-306.

Polovina, J., Mitchum, G., & Evans, G. (1996). Decadal and basin-scale variation in mixed layer depth and the impact on biological production in the central and North Pacific, 1960-88. *Oceanographic Literature Review, 6*, 581.

Pörtner, H.-O., & Knust, R. (2007). Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science, 315*, 95-97. doi:https://doi.org/10.1126/science.1135471

Pörtner, H.-O., & Peck, M. A. (2010). Climate change effects on fishes and fisheries: towards a cause-and-effect understanding. *Journal of Fish Biology, 77*, 1745-1779. doi:https://doi.org/10.1111/j.1095-8649.2010.02783.x

Puerta, P., Ciannelli, L., Rykaczewski, R. R., Opiekun, M., & Litzow, M. A. (2019). Do Gulf of Alaska fish and crustacean populations show synchronous non-stationary responses to climate? *Progress in Oceanography, 175*, 161-170.

Punt, A. E., A'mar, T., Bond, N. A., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A., . . . Szuwalski, C. (2013). Fisheries management under climate and environmental uncertainty: control rules and performance simulation. *ICES Journal of Marine Science, 71*, 2208-2220. doi:https://doi.org/10.1093/icesjms/fst057

Punt, A. E., Castillo-Jordán, C., Hamel, O. S., Cope, J. M., Maunder, M. N., & Ianelli, J. N. (2021). Consequences of error in natural mortality and its estimation in stock assessment models. *Fisheries Research, 233*, 105759. doi:https://doi.org/10.1016/j.fishres.2020.105759

Punt, A. E., & Hilborn, R. (1997). Fisheries stock assessment and decision analysis: the Bayesian approach. *Reviews in Fish Biology, 7*, 35-63.

Punt, A. E., Hurtado-Ferro, F., & Whitten, A. R. (2014). Model selection for selectivity in fisheries stock assessments. *Fisheries Research, 158*, 124-134. doi:https://doi.org/10.1016/j.fishres.2013.06.003

Punt, A. E., Szuwalski, C. S., & Stockhausen, W. (2014). An evaluation of stock–recruitment proxies and environmental change points for implementing the US Sustainable Fisheries Act. *Fisheries Research, 157*, 28-40. doi:https://doi.org/10.1016/j.fishres.2014.03.015

Pyper, B. J., & Peterman, R. M. (1998). Comparison of methods to account for autocorrelation in correlation analyses of fish data. *Canadian Journal of Fisheries and Aquatic Sciences, 55*, 2127-2140. doi:https://doi.org/10.1139/f98-104

Quinn, T., Marty, G. D., Wilcock, J., & Willette, M. (2001). *Disease and population assessment of Pacific herring in Prince William Sound, Alaska*. Paper presented at the Lowell Wakefield Symposium, Anchorage, AK.

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ralston, S., Punt, A. E., Hamel, O. S., DeVore, J. D., & Conser, R. J. (2011). A meta-analytic approach to quantifying scientific uncertainty in stock assessments. *Fishery Bulletin, 109*, 217-232.

Read, A. J., & Brownstein, C. R. (2003). Considering Other Consumers: Fisheries, Predators, and Atlantic Herring in the Gulf of Maine. *Conservation Ecology, 7*. doi:https://doi.org/10.5751/ES-00474-070102

Reum, J. C., Essington, T. E., Greene, C. M., Rice, C. A., & Fresh, K. L. (2011). Multiscale influence of climate on estuarine populations of forage fish: the role of coastal upwelling, freshwater flow and temperature. *Marine Ecology Progress Series, 425*, 203-215. doi:https://doi.org/10.3354/meps08997

Ricard, D., Minto, C., Baum, J., & Jensen, O. (2011). Assessing the stock assessment knowledge and status of marine fisheries with the new RAM Legacy database. *Fish and Fisheries, 13*, 380-398. doi:https://doi.org/10.1111/j.1467-2979.2011.00435.x

Ricard, D., Zimmermann, F., & Heino, M. (2016). Are negative intra-specific interactions important for recruitment dynamics? A case study of Atlantic fish stocks. *Marine Ecology Progress Series, 547*, 211-217. doi:https://doi.org/10.3354/meps11625

Rice, J., & Browman, H. I. (2014). Where has all the recruitment research gone, long time passing? *ICES Journal of Marine Science, 71*, 2293-2299. doi:10.1093/icesjms/fsu158

Rice, S., & Carls, M. (2007). *Prince William Sound herring: an updated synthesis of population declines and lack of recovery, Exxon Valdez Oil Spill Restoration Project Final Report (Restoration Project 050794)*. Retrieved from National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Auke Bay Laboratory, Juneau, Alaska: http://www.evostc.state.ak.us/Store/FinalReports/2005-050794-Final.pdf

Rossi, S., Viarouge, C., Faure, E., Gilot-Fromont, E., Gache, K., Gibert, P., . . . Bréard, E. (2017). Exposure of Wildlife to the Schmallenberg Virus in France (2011–2014): Higher, Faster, Stronger (than Bluetongue)! *Transboundary and Emerging Diseases, 64*, 354-363. doi:https://doi.org/10.1111/tbed.12371

Royer, T. C. (1982). Coastal fresh water discharge in the northeast Pacific. *Journal of Geophysical Research: Oceans, 87*, 2017-2021.

Schirripa, M. J., Goodyear, C. P., & Methot, R. M. (2009). Testing different methods of incorporating climate data into the assessment of US West Coast sablefish. *ICES Journal of Marine Science, 66*, 1605-1613. doi:10.1093/icesjms/fsp043

Schrimpf, M. B., Parrish, J. K., & Pearson, S. F. (2012). Trade-offs in prey quality and quantity revealed through the behavioral compensation of breeding seabirds. *Marine Ecology Progress Series, 460*, 247-259.

Schweigert, J. F., Boldt, J. L., Flostrand, L., & Cleary, J. S. (2010). A review of factors limiting recovery of Pacific herring stocks in Canada. *ICES Journal of Marine Science, 67*, 1903-1913. doi:https://doi.org/10.1093/icesjms/fsq134

Sewall, F., Norcross, B., Mueter, F., & Heintz, R. (2017). Empirically based models of oceanographic and biological influences on Pacific Herring recruitment in Prince William Sound. *Deep Sea Research Part II: Topical Studies in Oceanography, 147*, 127-137.

Sigler, M. F., Gende, S. M., & Csepp, D. J. (2017). Association of foraging Steller sea lions with persistent prey hot spots in southeast Alaska. *Marine Ecology Progress Series, 571*, 233-243. doi:https://doi.org/10.3354/meps12145

Similä, T., & Ugarte, F. (1993). Surface and underwater observations of cooperatively feeding killer whales in northern Norway. *Canadian Journal of Zoology, 71*, 1494-1499. doi:10.1139/z93-210

Skall, H. F., Olesen, N. J., & Mellergaard, S. (2005). Viral haemorrhagic septicaemia virus in marine fish and its implications for fish farming – a review. *Journal of Fish Diseases, 28*, 509-529. doi:10.1111/j.1365-2761.2005.00654.x

Smith, A. D., Brown, C. J., Bulman, C. M., Fulton, E. A., Johnson, P., Kaplan, I. C., . . . Shannon, L. J. (2011). Impacts of fishing low–trophic level species on marine ecosystems. *Science, 333*, 1147-1150. doi:https://doi.org/10.1126/science.1209395

Smith, M. J., White, A., Sherratt, J. A., Telfer, S., Begon, M., & Lambin, X. (2008). Disease effects on reproduction can cause population cycles in seasonal environments. *Journal of Animal Ecology, 77*, 378-389. doi:https://doi.org/10.1111/j.1365-2656.2007.01328.x

Smith, T. M., Reynolds, R. W., Peterson, T. C., & Lawrimore, J. (2008). Improvements to NOAA's Historical Merged Land–Ocean Surface Temperature Analysis (1880–2006). *Journal of Climate, 21*, 2283-2296. doi:https://doi.org/10.1175/2007JCLI2100.1

Somarakis, S., Tsoukali, S., Giannoulaki, M., Schismenou, E., & Nikolioudakis, N. (2018). Spawning stock, egg production and larval survival in relation to small pelagic fish recruitment. *Marine Ecology Progress Series: Advance View, 10*. doi:https://doi.org/10.3354/meps12642

Southward, A., Hawkins, S., & Burrows, M. (1995). Seventy years' observations of changes in distribution and abundance of zooplankton and intertidal organisms in the western English Channel in relation to rising sea temperature. *Journal of thermal Biology, 20*, 127-155. doi:https://doi.org/10.1016/0306-4565(94)00043-I

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B, 64*, 583-639.

Spies, I., Aydin, K., Ianelli, J. N., & Palsson, W. (2017). Assessment of the arrowtooth flounder stock in the Gulf of Alaska. *NOAA*.

Stachura, M. M., Essington, T. E., Mantua, N. J., Hollowed, A. B., Haltuch, M. A., Spencer, P. D., . . . Doyle, M. J. (2014). Linking Northeast Pacific recruitment synchrony to environmental variability. *Fisheries Oceanography, 23*, 389-408. doi:https://doi.org/10.1111/fog.12066

Stewart, I. J., & Hamel, O. S. (2014). Bootstrapping of sample sizes for length-or age-composition data used in stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences, 71*, 581-588.

Stick, K. C., Lindquist, A., & Lowry, D. (2014). *2012 Washington State Herring Stock Status Report* (FPA 14-09). Retrieved from Washington Department of Fish and Wildlife: https://wdfw.wa.gov/sites/default/files/publications/01628/wdfw01628.pdf

Stocker, M., Haist, V., & Fournier, D. (1985). Environmental variation and recruitment of Pacific herring (Clupea harengus pallasi) in the Strait of Georgia. *Canadian Journal of Fisheries and Aquatic Sciences, 42*, s174-s180.

Stocking, J., Bishop, M. A., & Arab, A. (2018). Spatio-temporal distributions of piscivorous birds in a subarctic sound during the nonbreeding season. *Deep Sea Research Part II: Topical Studies in Oceanography, 147*, 138-147. doi:https://doi.org/10.1016/j.dsr2.2017.07.017

Straley, J. M., Moran, J. R., Boswell, K. M., Vollenweider, J. J., Heintz, R. A., Quinn II, T. J., . . . Rice, S. D. (2017). Seasonal presence and potential influence of humpback whales on wintering Pacific herring populations in the Gulf of Alaska. *Deep Sea Research Part II: Topical Studies in Oceanography, 147*, 173-186. doi:https://doi.org/10.1016/j.dsr2.2017.08.008

Sturdevant, M. V., Brenner, R., Fergusson, E. A., Orsi, J. A., & Heard, B. (2013). Does predation by returning adult pink salmon regulate pink salmon or herring abundance. *Pac. Anadr. Fish Comm. Tech. Rep, 9*, 153-164.

Surma, S., Pakhomov, E. A., & Pitcher, T. J. (2018a). Energy-based ecosystem modelling illuminates the ecological role of Northeast Pacific herring. *Marine Ecology Progress Series, 588*, 147-161. doi:https://doi.org/10.3354/meps12430

Surma, S., & Pitcher, T. J. (2015). Predicting the effects of whale population recovery on Northeast Pacific food webs and fisheries: an ecosystem modelling approach. *Fisheries Oceanography, 24*, 291-305. doi:https://doi.org/10.1111/fog.12109

Surma, S., Pitcher, T. J., Kumar, R., Varkey, D., Pakhomov, E. A., & Lam, M. E. (2018b). Herring supports Northeast Pacific predators and fisheries: Insights from ecosystem

modelling and management strategy evaluation. *PloS one, 13*, e0196307. doi:https://doi.org/10.1371/journal.pone.0196307

Sveegaard, S., Nabe-Nielsen, J., Stæhr, K.-J., Jensen, T. F., Mouritsen, K. N., & Teilmann, J. (2012). Spatial interactions between marine predators and their prey: herring abundance as a driver for the distributions of mackerel and harbour porpoise. *Marine Ecology Progress Series, 468*, 245-253. doi:https://doi.org/10.3354/meps09959

Szuwalski, C. S., Britten, G. L., Licandeo, R., Amoroso, R. O., Hilborn, R., & Walters, C. (2019). Global forage fish recruitment dynamics: A comparison of methods, time-variation, and reverse causality. *Fisheries Research, 214*, 56-64. doi:https://doi.org/10.1016/j.fishres.2019.01.007

Szuwalski, C. S., Vert-Pre, K. A., Punt, A. E., Branch, T. A., & Hilborn, R. (2015). Examining common assumptions about recruitment: a meta-analysis of recruitment dynamics for worldwide marine fisheries. *Fish and Fisheries, 16*, 633-648. doi:https://doi.org/10.1111/faf.12083

Tabak, M. A., Pedersen, K., & Miller, R. S. (2019). Detection error influences both temporal seroprevalence predictions and risk factors associations in wildlife disease models. *Ecology and Evolution, 9*, 10404-10414. doi:https://doi.org/10.1002/ece3.5558

Teerlink, S. F., von Ziegesar, O., Straley, J. M., Quinn, T. J., Matkin, C. O., & Saulitis, E. L. (2015). First time series of estimated humpback whale (Megaptera novaeangliae) abundance in Prince William Sound. *Environmental and Ecological Statistics, 22*, 345-368. doi:https://doi.org/10.1007/s10651-014-0301-8

Thomas, G. L., & Thorne, R. E. (2003). Acoustical-optical assessment of Pacific herring and their predator assemblage in Prince William Sound, Alaska. *Aquatic Living Resources, 16*, 247-253.

Thorne, R. (2008). Walleye pollock as predator and prey in the Prince William Sound ecosystem. *Resiliency of gadid stocks to fishing and climate change. Alaska Sea Grant, University of Alaska, Fairbanks (16 p.)*.

Thornton, T. F., Butler, V., Funk, F., Moss, M., Hebert, J., Elder, T., . . . Hamada, S. (2010). *Herring Synthesis: Documenting and modeling herring spawning areas within socioecological systems over time in the Southeastern Gulf of Alaska* (728). Retrieved from North Pacific Research Board: https://uas.alaska.edu/research/herringsynthesis/final_docs/HerringSynthesisFINAL102710.pdf

Thornton, T. F., & Kitka, H. (2015). An indigenous model of a contested Pacific Herring fishery in Sitka, Alaska. *International Journal of Applied Geospatial Research (IJAGR), 6*, 94-117. doi:https://doi.org/10.4018/ijagr.2015010106

Tjelmeland, S., & Lindstrøm, U. (2005). An ecosystem element added to the assessment of Norwegian spring-spawning herring: implementing predation by minke whales. *ICES Journal of Marine Science, 62*, 285-294. doi:https://doi.org/10.1016/j.icesjms.2004.12.011

Toresen, R., & Østvedt, O. J. (2000). Variation in abundance of Norwegian spring-spawning herring (Clupea harengus, Clupeidae) throughout the 20th century and the influence of climatic fluctuations. *Fish and Fisheries, 1*, 231-256.

Trites, A., & Donnelly, C. (2003). The decline of Steller sea lions *Eumetopias jubatus* in Alaska: a review of the nutritional stress hypothesis. *Mammal review, 33*, 3-28.

Trochta, J. T., & Branch, T. A. (2018). Meta-analysis of Global Herring Population Dynamics. Research Workspace. https://doi.org/10.24431/rw1k1i.

Trochta, J. T., Branch, T. A., Shelton, A. O., & Hay, D. E. (2020). The highs and lows of herring: A meta-analysis of patterns and factors in herring collapse and recovery. *Fish and Fisheries, 21*, 639-662. doi:https://doi.org/10.1111/faf.12452

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., & Gelman, A. (2020). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*, 1413-1432. doi:10.1007/s11222-016-9696-4

Vehtari, A., & Ojanen, J. (2012). A survey of Bayesian predictive methods for model assessment, selection and comparison. *Statist. Surv., 6*, 142-228. doi:10.1214/12-SS102

Vert-pre, K. A., Amoroso, R. O., Jensen, O. P., & Hilborn, R. (2013). Frequency and intensity of productivity regime shifts in marine fish stocks. *Proc Natl Acad Sci U S A, 110*, 1779-1784. doi:https://doi.org/10.1073/pnas.1214879110

Vetter, E. F. (1988). Estimation of natural mortality in fish stocks: a review. *Fish. Bull., 86*, 25-43.

Vogelbein, W. K., Hoenig, J., Gauthier, D., Smith, M., Sadler, P., Kator, H., & Rhodes, M. (2012). The Role of Mycobacteriosis in Elevated Natural Mortality of Chesapeake Bay Striped Bass: Developing Better Models for Stock Assessment and Management : a final report. *Virginia Institute of Marine Science, William & Mary*. doi:https://doi.org/10.25773/v5-d39g-3056

Walters, C. J., & Ludwig, D. (1981). Effects of measurement errors on the assessment of stock–recruitment relationships. *Canadian Journal of Fisheries and Aquatic Sciences, 38*, 704-710. doi:https://doi.org/10.1139/f81-093

Wang, J., Zamar, R., Marazzi, A., Yohai, V., Salibian-Barrera, M., Maronna, R., . . . Konis, K. (2017). Package 'robust'. *R package version 0.4-18*.

Ward, E. J., Adkison, M., Couture, J., Dressel, S. C., Litzow, M. A., Moffitt, S., . . . Brenner, R. (2017). Evaluating signals of oil spill impacts, climate, and species interactions in Pacific herring and Pacific salmon populations in Prince William Sound and Copper River, Alaska. *PloS one, 12*, e0172898.

Ware, D. M., & Thomson, R. E. (2005). Bottom-Up Ecosystem Trophic Dynamics Determine Fish Production in the Northeast Pacific. *Science, 308*, 1280-1284. doi:10.1126/science.1109049

Watanabe, S. (2013). A widely applicable Bayesian information criterion. *Journal of Machine Learning Research, 14*, 867-897.

Wilberg, M. J., & Bence, J. R. (2008). Performance of deviance information criterion model selection in statistical catch-at-age analysis. *Fisheries Research, 93*, 212-221. doi:https://doi.org/10.1016/j.fishres.2008.04.010

Williams, E. H., & Quinn, T., J. (2000). Pacific herring, *Clupea pallasi*, recruitment in the Bering Sea and north-east Pacific Ocean, II: relationships to environmental variables and implications for forecasting. *Fisheries Oceanography, 9*, 300-315. doi:https://doi.org/10.1046/j.1365-2419.2000.00146.x

Willson, M. F., & Womble, J. N. (2006). Vertebrate exploitation of pulsed marine prey: a review and the example of spawning herring. *Reviews in Fish Biology and Fisheries, 16*, 183-200. doi:https://doi.org/10.1007/s11160-006-9009-7

Wilson, A., Goldberg, T., Marcquenski, S., Olson, W., Goetz, F., Hershberger, P., . . . Toohey-Kurth, K. (2014). Development and Evaluation of a Blocking Enzyme-Linked Immunosorbent Assay and Virus Neutralization Assay To Detect Antibodies to Viral Hemorrhagic Septicemia Virus. *Clinical and Vaccine Immunology, 21*, 435. doi:10.1128/CVI.00675-13

Winter, A. K., Martinez, M. E., Cutts, F. T., Moss, W. J., Ferrari, M. J., McKee, A., . . . Metcalf, C. J. E. (2018). Benefits and Challenges in Using Seroprevalence Data to Inform Models for Measles and Rubella Elimination. *The Journal of Infectious Diseases, 218*, 355-364. doi:10.1093/infdis/jiy137

Womble, J. N., & Sigler, M. F. (2006). Seasonal availability of abundant, energy-rich prey influences the abundance and diet of a marine predator, the Steller sea lion Eumetopias jubatus. *Marine Ecology Progress Series, 325*, 281-293. doi:https://doi.org/10.3354/meps325281

Worm, B., Hilborn, R., Baum, J. K., Branch, T. A., Collie, J. S., Costello, C., . . . Zeller, D. (2009). Rebuilding global fisheries. *Science, 325*, 578-585. doi:https://doi.org/10.1126/science.1173146

Yang, M.-S. (1993). Food habits of the commercially important groundfishes in the Gulf of Alaska in 1990.

Yeh, S.-W., Kang, Y.-J., Noh, Y., & Miller, A. J. (2011). The North Pacific Climate Transitions of the Winters of 1976/77 and 1988/89. *Journal of Climate, 24*, 1170-1183. doi:10.1175/2010JCLI3325.1

Zimmermann, F., Claireaux, M., & Enberg, K. (2019). Common trends in recruitment dynamics of north-east Atlantic fish stocks and their links to environment, ecology and management. *Fish and Fisheries, 20*, 518-536. doi:https://doi.org/10.1111/faf.12360

Appendix A

Incomplete time series

There were alternative catch time series originating from different sources and spanning different time periods for some of the western Pacific populations; the Korf-Karaginsk, Okhotsk, and Hokkaido-Sakhalin populations had overlapping time periods that did not exactly agree. We used the catch series reported most frequently and extended this back in time using the older time series for non-overlapping years, with the resulting combined series displayed in Fig. 1.1. Still, most of the overlapping catches had a similar scale and followed similar trajectories (except Gizigha-Kamchatka for which we used the time series referenced in Hay et al. 2001). The conjoining of catch series had no bearing on our results since all of these populations did not have associated biomass and recruitment time series during the time frame we analyzed (most recent 30 years).

Stock definitions

There is no generally agreed upon definition of what constitutes a herring stock or population (Begg, Friedland, & Pearce, 1999), terms that we use interchangeably here. The data available for herring may be for individual spawning groups, aggregated stock complexes, or a variety of other local to regional groupings. We followed the population definitions used or otherwise recommended by management in the various regions. For example, in Puget Sound, herring schools are surveyed annually at multiple known spawning locations, but there are only three genetically distinct populations that maintain some degree of geographical separation and are recognized by management (Stick, Lindquist, & Lowry, 2014). We therefore used these three groupings (Cherry Point, Squaxin Pass, rest of Puget Sound) for our analyses. For Alaska, herring stocks are identified by management based on spawning area. The catch series from Dutch Harbor is an exception, as it represents landings from mixed-stock bait fishery that does not occur during spawning time. While the herring caught at Dutch harbor are primarily comprised of Togiak herring, herring from other Eastern Bering Sea spawning stocks are also captured (Rowell, Brannian, & Funk, 1989). Similarly, while the Gulf of Maine-Georges Bank is acknowledged to be a complex of stocks with several distinct spawning groups (McQuinn, 1997), management applies to the stock complex as a whole and time series are only available at this level of aggregation (see Table A1 for data providers).

Zero-inflated negative binomial linear mixed effects model

We fit zero-inflated negative binomial regression models to collapsed biomass and high recruitment year counts. Many biomass time series do not extend below certain collapse thresholds (i.e. 0.2 or 0.15), leading to an excess of zeros in the observed counts that may ultimately bias regression estimates. This motivates our use of the zero-inflated negative binomial regression models. The zero-inflation model is bounded between 0 and 1 with a logistic link and can also include linear predictors to further explain changes in the probability of extra zeros. Given that the population units may describe stock-complexes, sub-populations within a complex, or spawning aggregations, we also fit models with a random intercept by management region. We use stock definitions used by management to identify regional groupings for random effects, since management agencies often survey various established spawning aggregations within a larger management division (sources for these definitions are found in Table A1). Thus, while time series exist for smaller spawning groups, they may share ecosystem, environment,

and fishing factors that we are unable to capture in our predictors, but still may influence correlations between these groups.

A zero-inflated negative binomial model is developed for each predictor, whose scaled effect (0 centers with S.D.=1) is incorporated in both the linear and zero-inflated components of the model. Significance of predictor's effects is tested using a parametric bootstrapping procedure that simulates zero-inflated counts from the reduced random effect model (i.e. region intercept only) and fits both the reduced and each of the full models. The difference in deviance between the full and reduced models is calculated, and this simulation is iterated 1,000 times to obtain a null distribution of deviance differences with which the observed deviance difference is compared to obtain a *p*-value. Since we fit many models, we adjust these bootstrapped *p*-values using the Holm–Bonferroni method (Holm, 1979) and select the top predictors based on the Holm–Bonferroni significance.

Random Survival Forests

Random survival forests are performed on survival data or event-time data (e.g. times-to-death). For our analysis, the event in question is the end of a period with low biomass or low recruitment, and the data are the measured times up until that end (i.e. population recovery or high recruitment) or the end of the observational period (i.e. right-censored). A random forest analysis is an machine learning method that constructs many regression trees from bootstrapped samples of the original data set. Each tree contains multiple nodes that splits the response data based on conditional rules involving predictor values that are chosen to optimize a splitting criterion (Breiman, 2001). With Random Survival Forests, this splitting criterion finds the split on a predictor that maximizes differences in the hazard function between the daughter nodes (Ishwaran, Kogalur, Blackstone, & Lauer, 2008). The hazard function for our data describes the change in number of recoveries or strong recruitment out of a number of ongoing low periods over time. Node splits in a Random Survival Forest maximize the sum of the differences in the number of recoveries or strong recruitments and number of individual low periods that are still ongoing over a range of time points. The specific splitting criterion used at each node is a log-rank test, which is a common hypothesis test for survival distributions as well as a robust splitting rule in survival regression trees (LeBlanc & Crowley, 1993). Variable importance (VIMP) and partial dependence plots of predictors are based on ensemble estimates of all trees within the forest and used to rank variables and assess their relationship. Readers should refer to Ishwaran et al. (2008) for details on data, splitting criterion, and other model notation applied in Random Survival Forest analysis.

The same predictors from the zero-inflated negative binomial GLMM are used in the Random Survival Forest analysis, with the addition of the year when the collapse or low recruitment period starts to determine potential decadal differences. Predictors that are derived from time series are calculated over the 30 most recent years. Additionally, stocks with shorter low biomass or recruitment times typically show multiple occurrences of these times during their 30 year period. Since a single predictor value is assigned to each stock, rather than each low biomass or recruitment time, there is pseudo-replication of these values in the data which may bias results. To reduce this pseudo-replication, we only take the maximum low biomass or recruitment time from each stock, since we are interested in explaining the longest times. If a stock shows multiple times of equal duration (whether it is censored or not), then a random time is selected to be included.

Partial dependence describes the change in the response variable with respect to a single predictor and all other predictors integrated out. Collapse probability is the response variable in evaluating partial dependence. Changes in collapse probability (or probability of high recruitment for the recruitment data) can alternatively be interpreted as changes in the relative number of recoveries over a range of predictor values. Random Survival Forest provides an ensemble estimate of the cumulative hazard over time (i.e. a sequence of values equal to the probability of recovery at specific durations of time) as opposed to a point prediction for each predictor value. To account for the time dimension of the cumulative hazard ensemble estimate, point predictions are made at specific times from a series of predictor values allowing us to evaluate partial dependence at specific low biomass durations.

As a machine learning method rather than a statistical regression, random forest analyses can address multicollinearity, though simulation studies have revealed comparable performance between the two techniques when high collinearity is present (Dormann et al., 2013). We explored variable cross-correlations and subsequently ran Random Survival Forest analyses with and without collinear variables to explore the potential impacts of collinearity on our results (Fig. A7). Variable cross-correlations show that several variable pairs have collinearity (i.e. $|r| > 0.7$, as suggested by Dormann et al. (2013)); SSHA standard deviation with mean age 5 weight, recruitment CV with biomass CV, recruitment CV with recruitment median (also in biomass CV and median), mean SSHA with SSHA standard deviation, latitude with mean SSHA, and SST trend with fish predators trend (Fig. A7). Using the variable pair with the highest collinearity, we removed Mean age 5 weight and re-ran the Random Survival Forest; SSHA standard deviation remained the most important variable. Conversely, removing SSHA standard deviation had not impact on the importance ranking of Mean age 5 weight.

Note on calculating relative biomass

We recognize that using the mean of the 90th percentiles from time series to normalize biomass, recruits, and catches is prone to being influenced by time series length. In other words, the 90th percentile will include more data points as time series length increases and potentially bias the resulting mean values. We alternatively applied a 30-year moving window calculation (i.e. the shortest time series considered for analysis) to determine the means of the 90th percentile of each possible window within a time series and then selected the maximum mean for normalizing the rest of the time series. This alternative scheme with the same thresholds did not change the overall conclusions from the GLMM analysis.

Sensitivity checks

Choices on which data to analyze were made to improve the interpretability of our results. For example, we only considered low biomass and recruitment durations greater than two years. Durations from one year and greater allowed more noise in the data and results since all stocks regularly experience one year low level events that are unlikely to reflect true population collapses. Another choice involved using the 30 most recent years to standardize comparisons of collapse durations; using all years from each time series for the collapse durations in the Random Survival Forest, with length of time series as a predictor, resulted in poor predictive accuracy. The hierarchy of stock definitions, from spawning groups, to stock complexes, to all herring within a sea basin, also complicated our definitions of individual observations. We used random effects in a GLMM framework to account for hierarchy among potential stock groupings (color coded in Fig. 1.1), but this did not improve nor change the main results. No hierarchy could be incorporated in the

Random Survival Forests while each stock can also have multiple low biomass or recruitment events longer than two years. This is problematic because more low biomass or recruitment events correlate with shorter event durations and predictor values would be pseudo-replicated if each event was explicitly considered (predictors are for stocks, not the individual collapses). Thus, we chose to focus on the longest low biomass or recruitment duration of each stock. The robustness of this choice was evaluated by randomly sampling a collapse duration per stock if a stock had multiple events; this did not change our results.

We checked the sensitivity of zero-inflated negative binomial regression results to outliers in the predictors. For low biomass, a single outlier is visually evident in both peak fishing as expressed by the maximum relative catch to biomass ratio and the log maximum catch (Fig. 1.4), but their removal did not change the significance nor sign of the effect of these predictors. For high recruitment, outliers were visually identified in the maximum relative catch to biomass ratios (1 outlier), log maximum catch (1), SST trend (2), SSHA standard deviation (3), and mean SST (3). Removing these did not alter the significance of bootstrapped *p*-values for SST trend (p=0.411), mean SST (p=0.271), and log maximum catch (p=0.588). SSHA standard deviation (p=0.006, lower 95th confidence interval=0.11, upper 95th confidence interval=0.42). For the two SST predictors, we justified the removal of several values because most values represented stocks located in small enclosed bays or straits (San Francisco Bay, Strait of Georgia, West Newfoundland, and Gulf of St. Lawrence, Icelandic summer spawners). The SST values associated with these stocks are likely inaccurate as they could not be extracted at the spawning locations approximated in Fig. 1.1, but the nearest open water pixel further from the coast which are far removed from nearshore, shallow conditions. This also applies to Icelandic herring, which spawn in fjords around the island where water conditions are likely isolated from the open water SST we extracted. The catch to biomass ratio became less significant too (p=0.031, lower 95th confidence interval=-0.53, upper 95th confidence interval=-.04). The bootstrapped significance of the SD of SSHA increased (p=0.007), although the adjusted significance still ranks this model lower compared to our top variables (median and CV of relative biomass). Finally, the main results and estimated effect sizes were robust to different thresholds for categorizing low biomass (less than 20% of Mean High Biomass) and high recruitment years (greater than 60% of Mean High Recruitment).

Rerunning the Random Survival Forest analysis on low biomass times after excluding visually identified outliers (one outlier in recruitment CV from Prince William Sound; two outliers in SSHA trend from Bothnian Bay and North Sea autumn spawners) still ranked the same top variables (SD of SSHA, linear trend in SSHA, CV of recruitment) and did not noticeably alter their Variable Importance values. When using a lower collapse threshold of 0.2, the out-of-bag error rate considerably worsened (0.65), suggesting that our results are sensitive to the collapse definition used.

Increasing the recruitment threshold to 0.6 did not alter predictive accuracy (error=0.23) and maintained the top two predictors (median biomass and SD of SSHA) as the most important. Decreasing the threshold to 0.4 slightly worsened accuracy (error=0.28) while changing the ranking of variable importance: the top variables (based on VIMP) were mean freshwater influx, predatory fish trends, SST trends, highest catch-to-biomass ratio, SD of SSHA, and median relative.

An alternative check for statistical differences in low biomass or low recruitment durations across predictor values is to calculate Harrell's concordance index (Harrell, Califf, Pryor, Lee, & Rosati, 1982). The concordance index indicates how well two groups of durations (split on the

predictor values) rank the collapse probability of any two stocks between the groups; values greater than 0.5 indicate that one group more consistently shows higher proportions of stocks collapsed at longer collapse times. We split biomass collapse times at breakpoints in the top ranked predictors as determined from their partial dependencies. For example, splitting collapse times at SD of SSHA of 3.5 produced a concordance index of 0.76 indicating highly discriminated differences between collapse times above and below 3.5 (Fig A10). Differences in collapse times are also shown for splits on CV of and median of relative recruitment (0.74 and 0.68 concordance indices, respectively), but less so for trends in SSHA and SST. For recruitment failure, splitting collapse times on median relative biomass (0.5) and SD of SSHA (3.0) resulted in the most distinct groups (concordance index=0.67), with higher biomass and SSHA SD associated with shorter recruitment failure (Fig. A11). Splitting on trends in predator fish and the highest catch-to-biomass ratio resulted in only less discernable differences in recruitment failure (Fig. A11). High concordance indices in our top predictors verify the results from the Random Survival Forest.

Non-significant predictors may inaccurately represent the underlying hypotheses we aimed to test

Non-significant factors may inaccurately represent the underlying hypotheses we aimed to test. Latitude has previously explained differences in spawn timing, maturity, and growth of eastern Pacific herring as a proxy for the climatic gradient in the northern hemisphere (Hay, 1985; Hay, Rose, Schweigert, & Megrey, 2008). However, the eastern Pacific climatic gradient does not match the gradient in the Atlantic since it is influenced by different atmospheric pressure systems and circulation patterns. Freshwater inputs characterize the physical processes and habitat quality of estuaries in which herring spawn and their progeny survive and grow to maturity (Fortier & Gagné, 1990; Hay & McCarter, 1997). However, these effects occur at a much finer spatial and temporal scale than represented by the freshwater influx values used here (e.g. decadal means over 5° of latitude). The earliest age at which herring mature cannot accurately capture potential differences in regeneration times since most herring start maturing after 2-3 years, while maturity schedules also likely time-vary depending on stock size. The mean weight at age 5 suffers an inadequate consideration of the time-varying nature of growth and body size, and will bias the mean weights of stocks depending on the time frames over which they are averaged and the external conditions controlling either size or growth in each stock. Maximum catch reflects fishing effort and availability of herring biomass to fisheries, which no doubt differs between stocks in light of unique fishing histories and the environmental conditions acting when maximum catch occurs. Consequently, true differences in latent population size are likely masked by fishing effects.

Outlying stocks in the Baltic Sea

Among the stocks we examined, those in the Baltic Sea experience the largest variability, mean, and fastest positive trends in SSHA. Previous analyses have found that these herring stocks are influenced by changes in food supply and competition with sprat (Möllmann, Kornilovs, Fetter, & Köster, 2005) and by recruitment variability in the earliest life states (Gröger, Hinrichsen, & Polte, 2014), which are affected by the Baltic Sea Index (BSI), a measure of atmospheric pressure reflecting localized and broader climate forces. Greater interannual hydrographic variability may promote frequent year-classes and shorter times of failed recruitment because of greater deviations from average conditions, such as average circulation patterns, that contribute to low relative recruitment or year-class failure (Houde, 2008). These deviations could involve mechanisms

alluding to variety of recruitment hypotheses, including Member/Vagrant (Iles & Sinclair, 1982), Stable Ocean (Lasker, 1978), or Optimal Environmental Window hypotheses (Cury & Roy, 1989), which can only be validated by closely investigating each herring stock's population dynamics and local conditions.

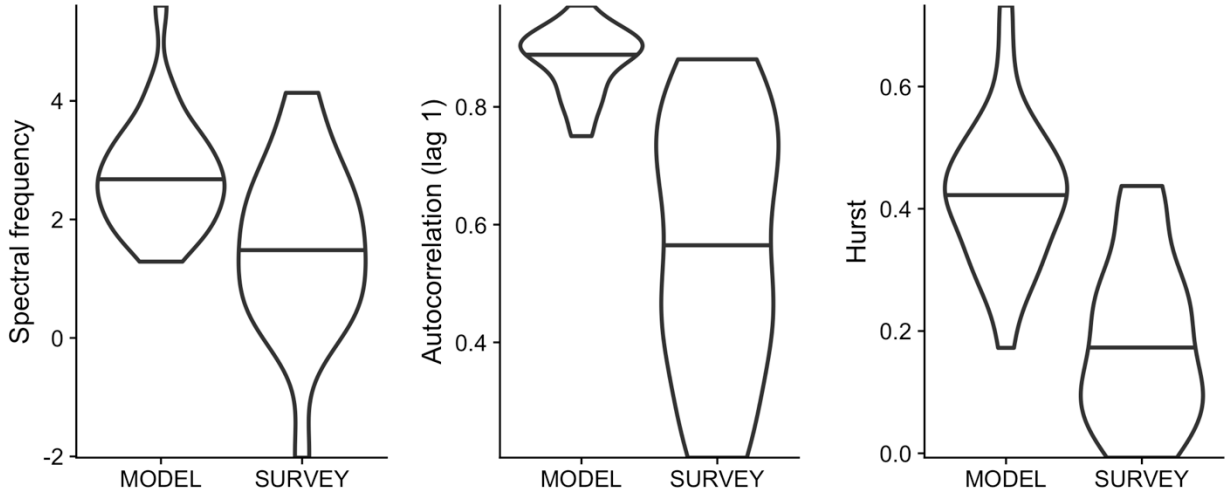

Fig. A1: Distributions of time series statistics comparing herring stock assessment (n=27) and **raw** survey (n=20) estimates. Seven time series of survey estimates are omitted due to large gaps in the times series (>2 years) and/or less than ten total observations overall (Cape Avinof, Goodnews Bay, Nelson Island, Norton Sound, Security Cove, Humboldt Bay, and Squaxin Pass).

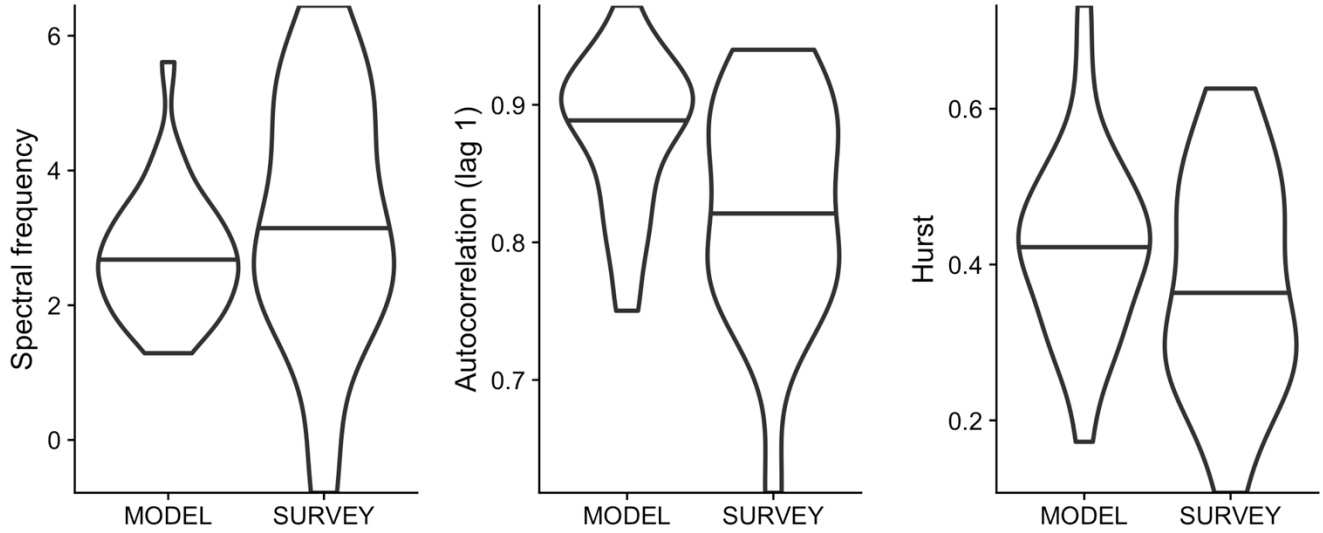

Fig. A2: Distributions of time series statistics comparing herring stock assessment (n=27) and **filtered** survey (n=20) estimates.

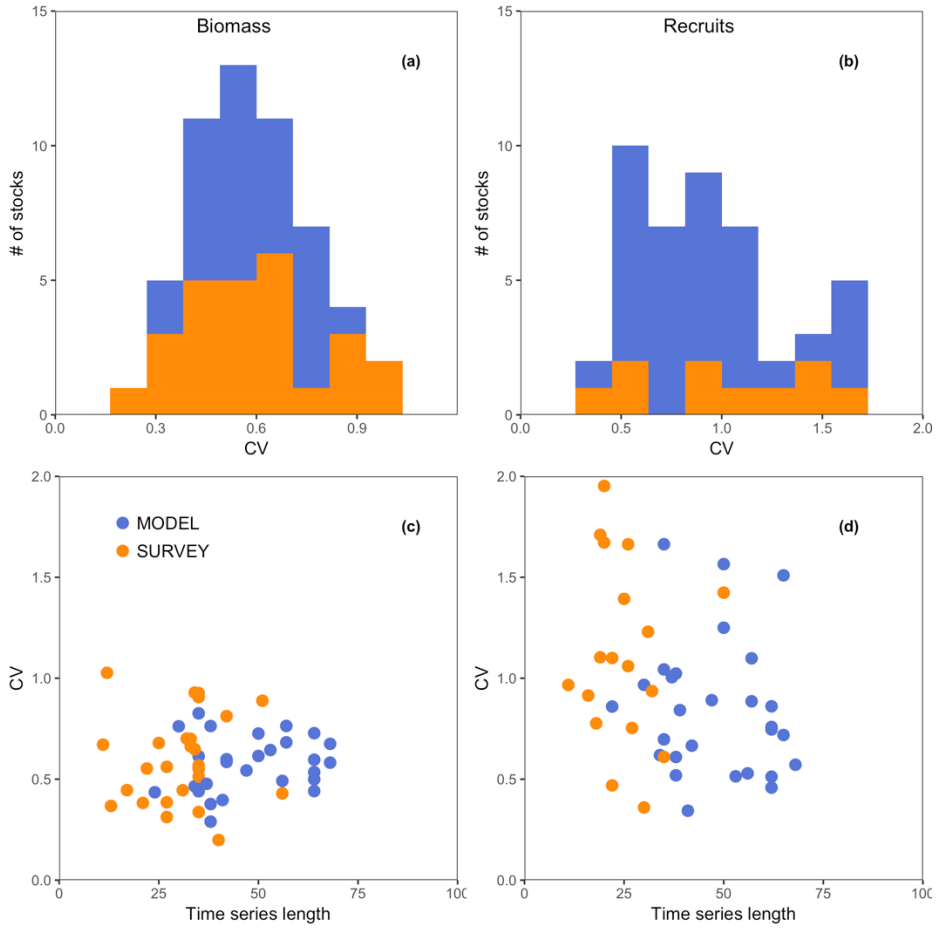

Fig. A3: Histograms of CV values for all herring biomass and recruitment time series and CV plotted against the lengths of herring time series. Data is colored coded by estimate type.

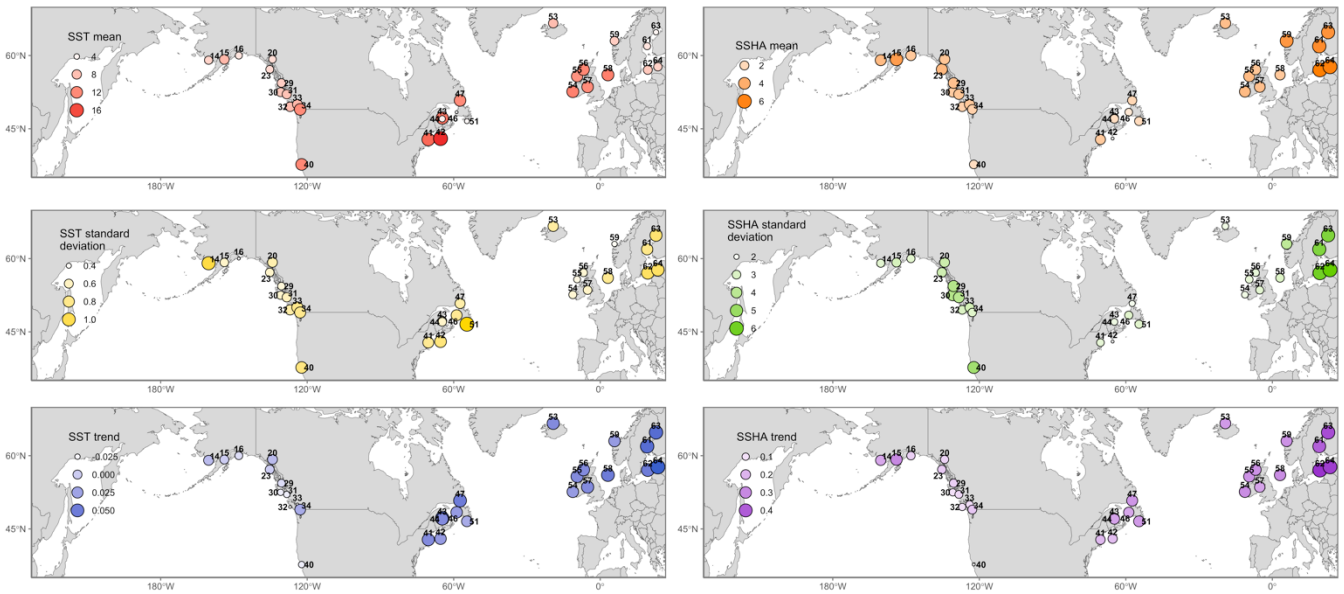

Fig. A4: Extracted locations and values for several metrics (mean, standard deviation, and trend) on sea surface temperature and sea surface height anomalies **during** each stock's estimated peak spawning month. The values for each metric are represented by both point size and color. Locations differ between the two environmental variables and with herring locations in Fig. 1.1 due to different grid resolutions in the original data sets, in which the cell with observations closest to each stock's location (Fig. 1.1) is shown. Higher mean SST is shown in the Northwest Atlantic because these stocks spawn in early autumn (September-October) when coastal waters are still near summer peak SST. The specific values and time frames over which they are calculated are shown in Table A3.

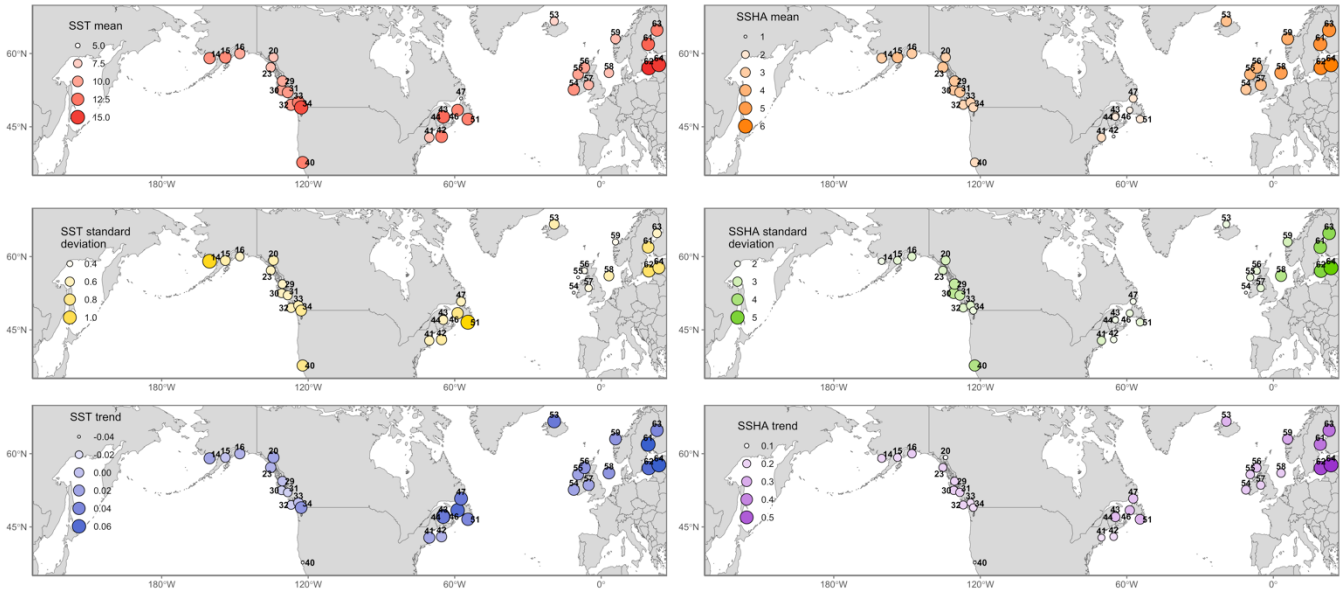

Fig. A5: Extracted locations and values for several metrics (mean, standard deviation, and trend) on sea surface temperature and sea surface height anomalies **following** each stock's estimated peak spawning month. The values for each metric are represented by both point size and color. Locations differ between the two environmental variables and with herring locations in Fig. 1.1 due to different grid resolutions in the original data sets, in which the cell with observations closest to each stock's location (Fig. 1.1) is shown. The specific values and time frames over which they are calculated are shown in Table A4.

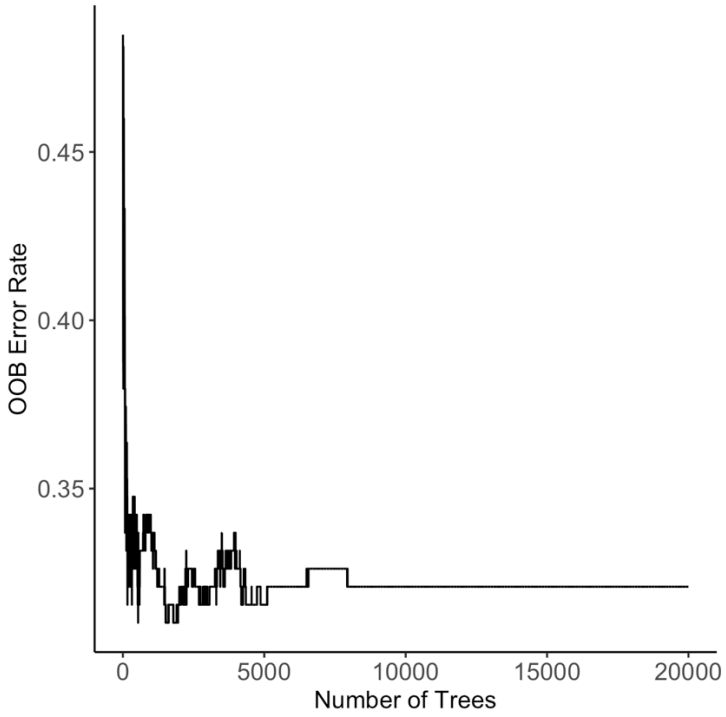

Fig. A6. Out-of-bag (OOB) prediction error for **biomass durations less than 30%** of Mean High Biomass as a function of the size of the random forest, or the number of trees (n=20,000) contained within. Prediction error is measured with Harrell's concordance index (Harrell et al. 1982), the proportion of available observation combinations where predicted event times agree with the inequality of the observed event times. Values close to 0.5 indicate 50/50 random predictive accuracy (i.e. no better than guessing), while 0 is complete accuracy.

Fig. A7. Pearson correlation coefficients between environmental, fishing, and life history predictors used in either the generalized linear model regression and random survival forest analysis. Predictor values that served as metrics for time series were calculated over the 30 most recent years, including the environmental series. We only used populations for which catches, recruitment, and adult biomass estimates were concurrently available (n = 30). Encircled values denote correlations > |0.70|, with values close to one showing significant collinearity between predictors.

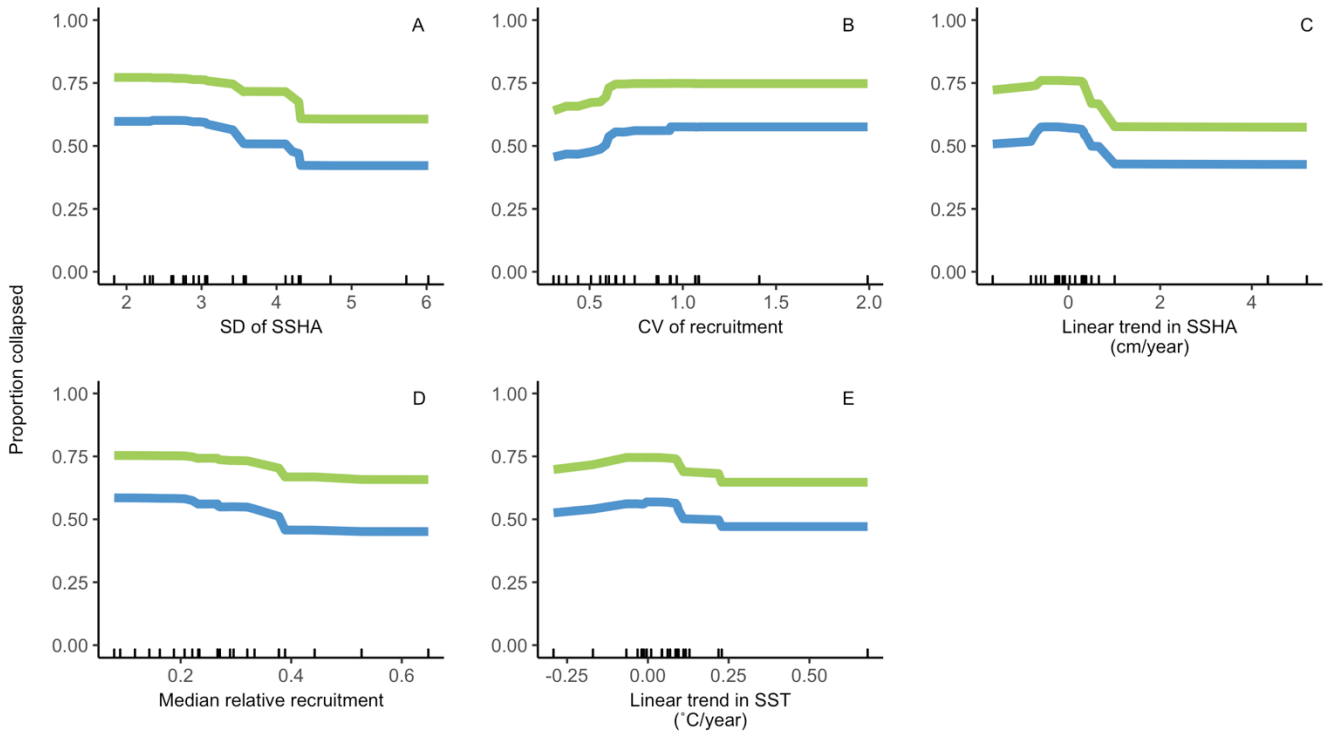

Fig. A8. Predicted partial dependence in the observations of the most important predictors of biomass collapse probabilities after 5 years (green) and 10 years (blue). Collapsed values correspond to the proportion of events that remain collapsed and are point predictions of the observed data (n=24) that is also conditioned on the other covariates.

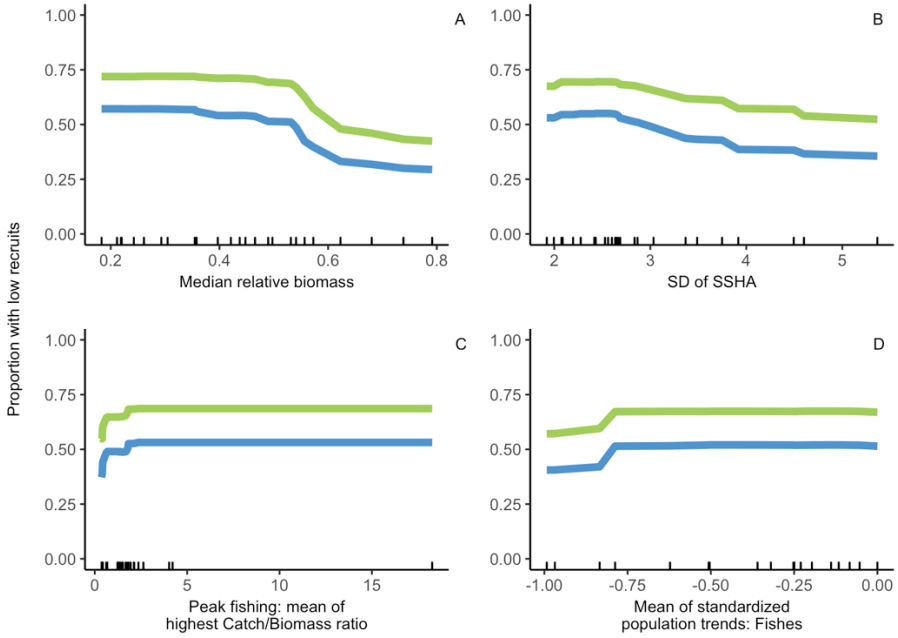

Fig. A9. Predicted partial dependence in the observations of the most important predictors of low recruit probabilities after 5 years (green) and 10 years (blue). Collapsed values correspond to the proportion of events that remain collapsed and are point predictions of the observed data (n=30) that is also conditioned on the other covariates.

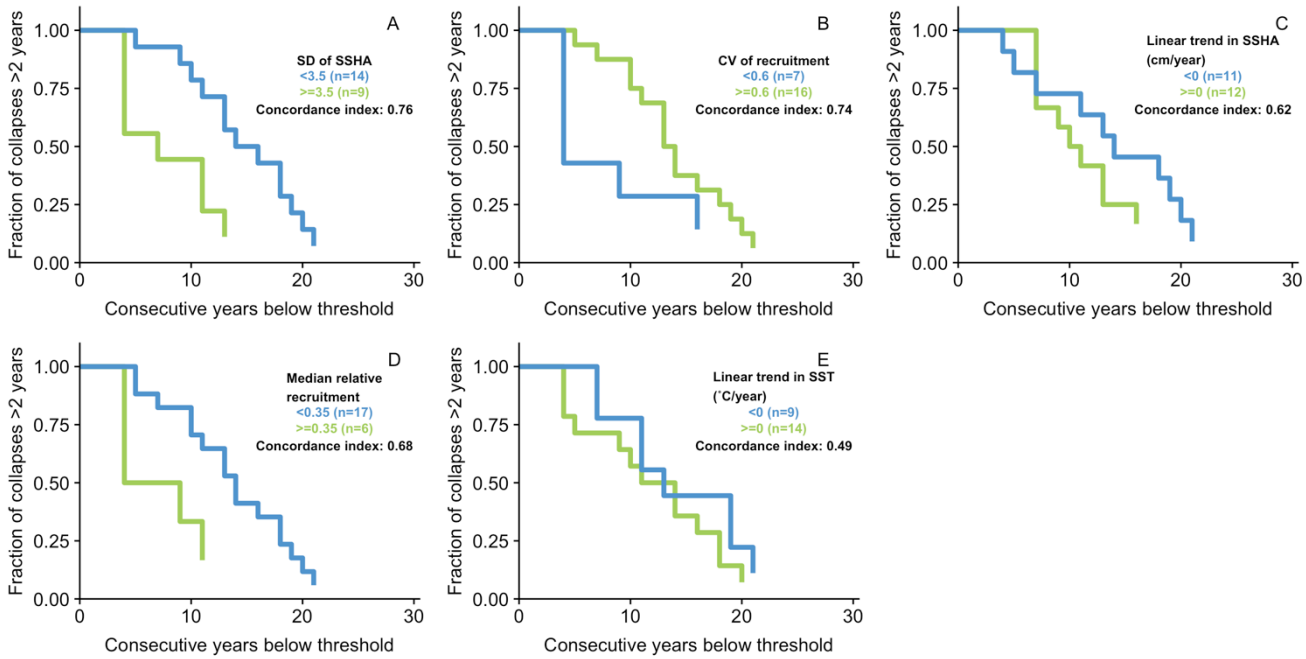

Fig. A10. Frequency of **biomass collapse** events of stocks grouped by splits on the top predictor values. The collapse time curves of the group below (blue) and above (green) the breakpoint used to split the data are shown along with the associated concordance index.

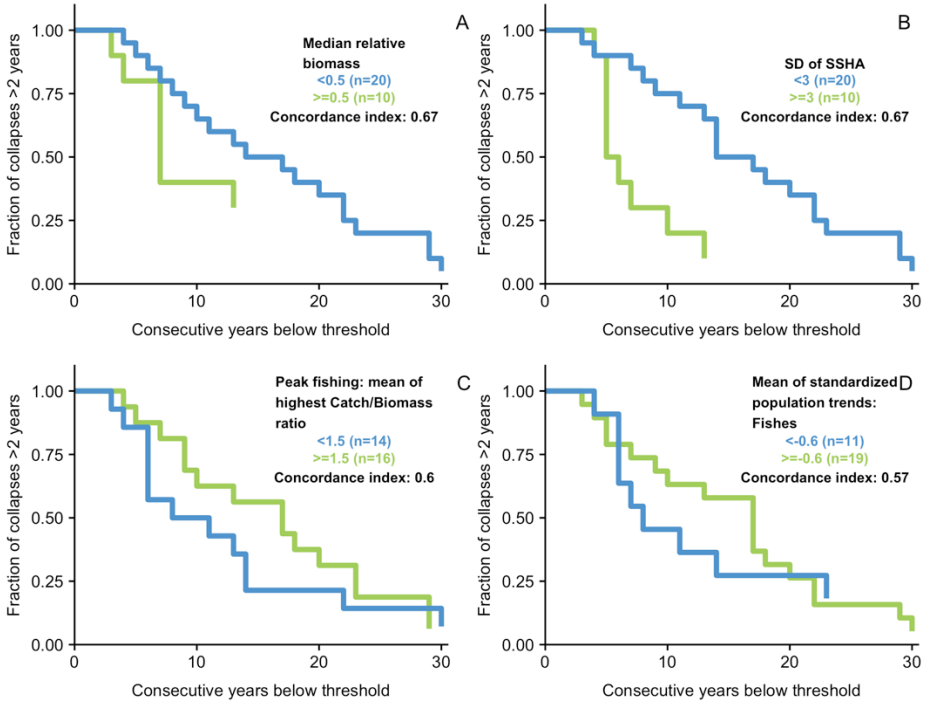

Fig. A11. Frequency of **recruitment failure** events of stocks grouped by splits on the top predictor values. The recruitment failure time curves of the group below (blue) and above (green) the breakpoint used to split the data are shown along with the associated concordance index.

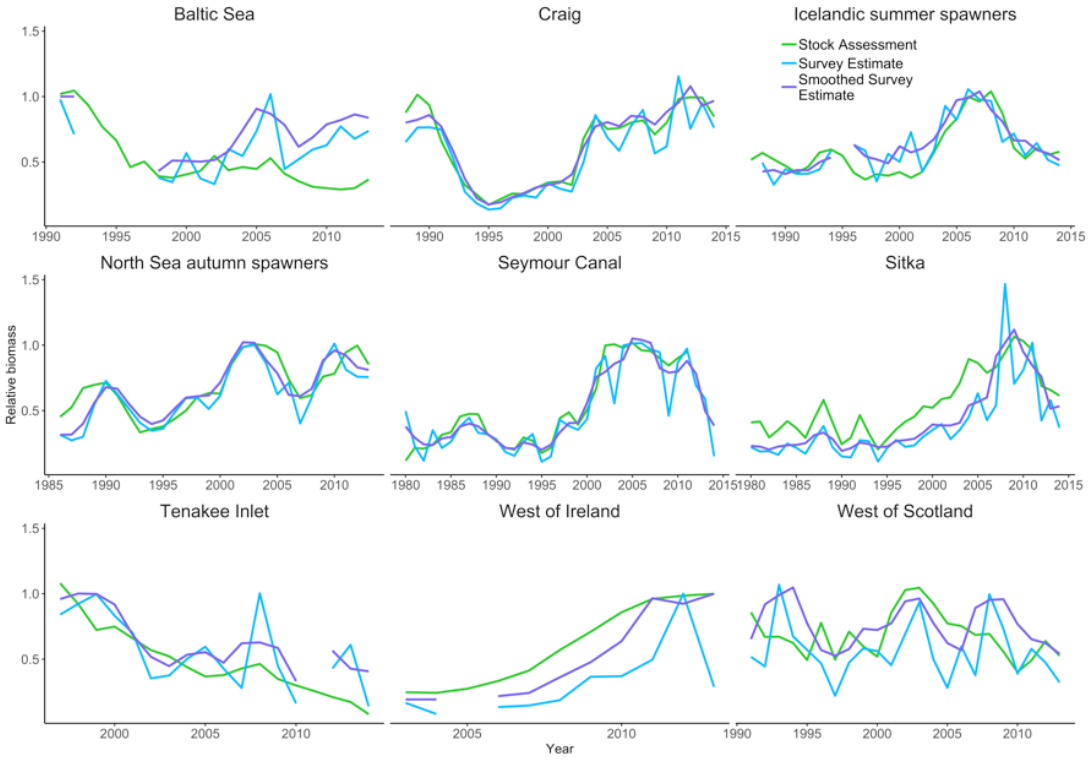

Fig. A12. Comparison between survey, smoothed survey, and assessment estimates for select stocks with both surveys and assessments in our dataset. The estimates show the expected close agreement between all three measures.

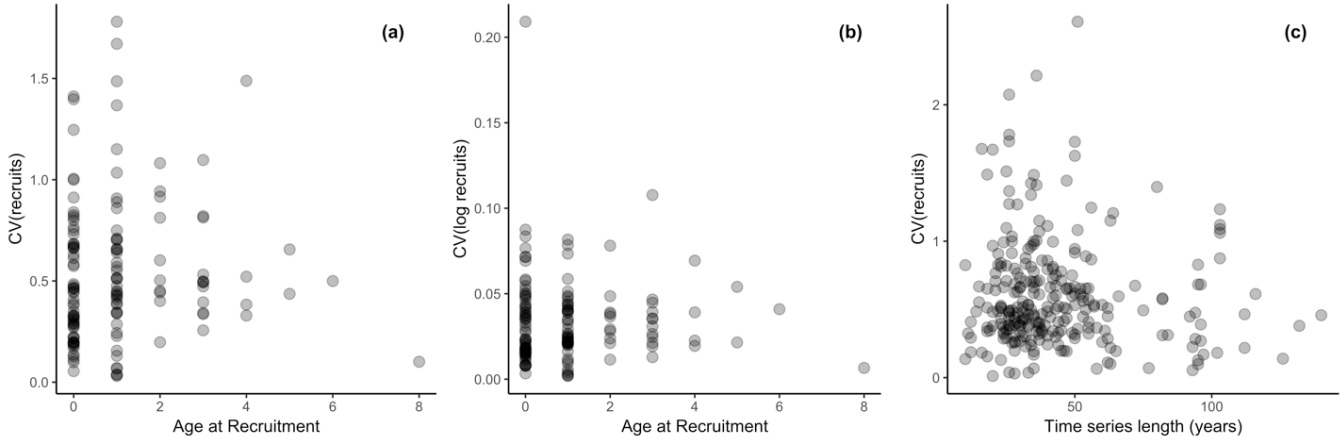

Fig. A13. The CV of recruitment vs. age at recruitment and CV of recruitment vs. time series length **of non-forage fish species in the RAM Legacy database**. No significant effect of age nor time series length on CV values is found, although the spread of CV values appears to narrow at older ages and longer time series length (i.e. fewer stocks at the larger values).

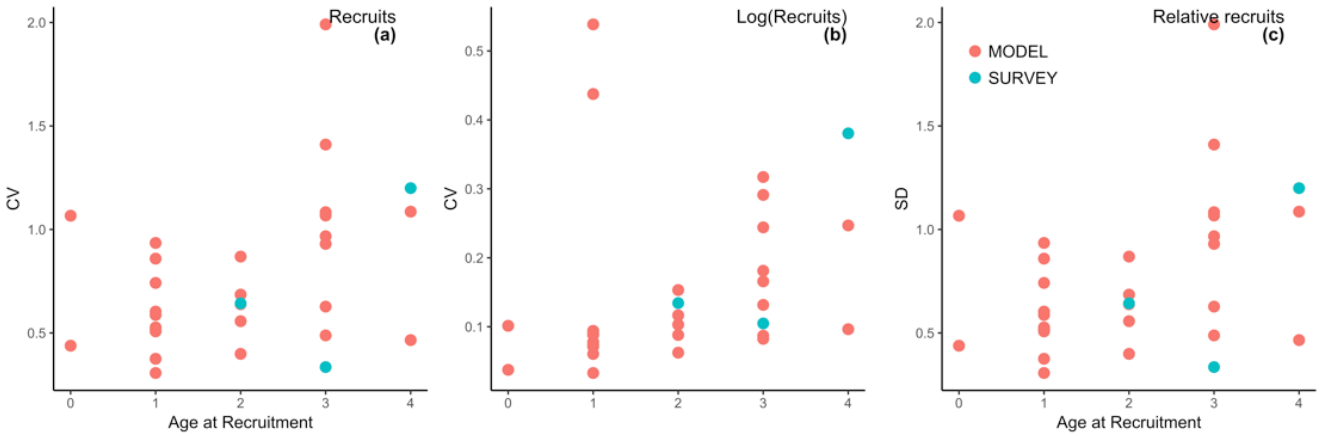

Fig. A14. The CV of recruitment, CV of log-recruitment, and SD of relative recruits (recruitment divided by Mean High Recruitment) plotted against age at recruitment **of herring stocks in our dataset**. The points are colored by the method used to derive recruitment estimates.

References

Begg, G. A., Friedland, K. D., & Pearce, J. B. (1999). Stock identification and its role in stock assessment and fisheries management: an overview. *Fisheries Research, 43*, 1-8.

Breiman, L. E. O. (2001). Random forests. *Machine Learning, 45*, 5-32.

Cury, P., & Roy, C. (1989). Optimal environmental window and pelagic fish recruitment success in upwelling areas. *Canadian Journal of Fisheries and Aquatic Sciences, 46*, 670-680.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*, 27-46.

Fortier, L., & Gagné, J. A. (1990). Larval Herring (Clupea harengus) Dispersion, Growth, and Survival in the St, Lawrence Estuary: Match/Mismatch or Membership/Vagrancy? *Canadian Journal of Fisheries and Aquatic Sciences, 47*, 1898-1912. doi:https://doi.org/10.1139/f90-214

Gröger, J. P., Hinrichsen, H.-H., & Polte, P. (2014). Broad-scale climate influences on spring-spawning herring (*Clupea harengus*) recruitment in the Western Baltic Sea. *PloS one, 9*, e87525.

Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L., & Rosati, R. A. (1982). Evaluating the yield of medical tests. *Jama, 247*, 2543-2546. doi:https://doi.org/10.1001/jama.247.18.2543

Hay, D. (1985). Reproductive biology of Pacific herring (Clupea harengus pallasi). *Canadian Journal of Fisheries and Aquatic Sciences, 42*, s111-s126. doi:https://doi.org/10.1139/f85-267

Hay, D., & McCarter, P. (1997). Larval distribution, abundance, and stock structure of British Columbia herring. *Journal of Fish Biology, 51*, 155-175. doi:https://doi.org/10.1111/j.1095-8649.1997.tb06098.x

Hay, D. E., Rose, K. A., Schweigert, J., & Megrey, B. A. (2008). Geographic variation in North Pacific herring populations: Pan-Pacific comparisons and implications for climate change impacts. *Progress in Oceanography, 77*, 233-240. doi:https://doi.org/10.1016/j.pocean.2008.03.015

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65-70.

Houde, E. D. (2008). Emerging from Hjort's shadow. *Journal of Northwest Atlantic Fishery Science, 41*, 53-70.

Iles, T., & Sinclair, M. (1982). Atlantic herring: stock discreteness and abundance. *Science, 215*, 627-633.

Ishwaran, H., Kogalur, U. B., Blackstone, E. H., & Lauer, M. S. (2008). Random survival forests. *The annals of applied statistics*, 841-860. doi:https://doi.org/10.1214/08-AOAS169

Lasker, R. (1978). The relation between oceanographic conditions and larval anchovy food in the California Current: identification of factors contributing to recruitment failure. *Rapp. P.-V. Reun. Cons. Int. Explor. Mer., 173*, 212-230.

LeBlanc, M., & Crowley, J. (1993). Survival trees by goodness of split. *Journal of the American Statistical Association, 88*, 457-467.

McQuinn, I. H. (1997). Metapopulations and the Atlantic herring. *Reviews in Fish Biology and Fisheries, 7*, 297-329.

Möllmann, C., Kornilovs, G., Fetter, M., & Köster, F. W. (2005). Climate, zooplankton, and pelagic fish growth in the central Baltic Sea. *ICES Journal of Marine Science, 62*, 1270-1280.

Rowell, K. A., Brannian, L. K., & Funk, F. (1989). Forecast of the Pacific herring biomass in Togiak district, Bristol Bay, 1990. *Alaska Department of Fish and Game, Regional Informati on Report 2D89-, Anchorage*.

Stick, K. C., Lindquist, A., & Lowry, D. (2014). *2012 Washington State Herring Stock Status Report* (FPA 14-09). Retrieved from Washington Department of Fish and Wildlife: https://wdfw.wa.gov/sites/default/files/publications/01628/wdfw01628.pdf

Table A1. Summary of herring populations used in analysis

Code	*Name*	*Alternative Name*	*Catches*	*Spawning biomass*	*Recruitment*	*Estimate type*	*Region*	*Reference*	*Contact*
BALRIG	Gulf of Riga	ICES 28	1977-2013	1977-2014	1977-2014	Extended survivor analysis (age-based)	Baltic	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Subdivision 28.1 (Gulf of Riga)/2014. ICES, Copenhagen.	
BALBOB	Bothnian Bay	ICES 31	1980-2013	1980-2013	1980-2013	Extended survivor analysis (age-based)	Baltic	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Subdivision 31 (Bothnian Bay)/2014. ICES, Copenhagen.	
BALBAS	Baltic Sea	ICES 25-32	1974-2013	1974-2014	1974-2014	Extended survivor analysis (age-based)	Baltic	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Subdivisions 25 - 29 (excluding Gulf of Riga) and 32/2014. ICES, Copenhagen.	
BALBOS	Bothnian Sea	ICES 30	1973-2013	1973-2014	1973-2014	State-space assessment model (age-based)	Baltic	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Subdivision 30 (Bothnian Sea)/2014. ICES, Copenhagen.	
BALWSP	Western Baltic spring spawners	ICES 22-24-IIIa	1991-2013	1991-2014	1991-2014	State-space assessment model (age-based)	Baltic	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Division IIIa and Subdivisions 22 - 24 (Western Baltic spring spawners)/2014. ICES, Copenhagen.	
NORSPR	Norwegian spring spawners		1907-2013	1907-2014	1950-2014	Virtual population analysis	Norwegian Sea	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Subareas I. II. V and Divisions IVa and XIVa (Norwegian spring-spawning herring)/2014. ICES, Copenhagen.	
NOSEAU	North Sea autumn spawners		1947-2013	1947-2014	1947-2014	State-space assessment model (age-based)	North Sea	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Subarea IV and Divisions IIIa and VIId (North Sea autumn spawners)/2014. ICES, Copenhagen.	
CELIRS	Irish Sea	ICES VIIa	1961-2013	1961-2013	1961-2013	State-space assessment model (age-based)	Celtic/Irish	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Division VIIa North of 52Â° 30â€™ N (Irish Sea)/2014. ICES, Copenhagen.	

WESCOT	West of Scotland	ICES VIa	1957-2013	1957-2013	1957-2013	Integrated catch-at-age analysis/separable virtual population analysis	Celtic/Irish	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Division VIa (North)/2014. ICES, Copenhagen.	
CELNWI	Northwest Ireland	ICES VIa-VIIb-VIIc	1957-2013	1957-2013	1957-2013	Integrated catch-at-age analysis/separable virtual population analysis	Celtic/Irish	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Divisions VIa (South) and VIIb.c/2014. ICES, Copenhagen.	
CELWIR	West of Ireland	ICES VIIa-g-h-j	1958-2013	1958-2013	1958-2014	State-space assessment model (age-based)	Celtic/Irish	ICES Stock Database, Extraction date: 2015/01/24 of Herring in Division VIIa South of 52Â° 30â€™ N and VIIg.h.j.k (Celtic Sea and South of Ireland)/2014. ICES, Copenhagen.	
ICESUM	Icelandic summer spawners		1947-2013	1947-2014	1950-2014	Virtual population analysis	Iceland	**1987-2014**: ICES Stock Database, Extraction date: 2015/01/24 of Herring in Division Va (Icelandic summer-spawners)/2014. ICES, Copenhagen. **1947-1986**: G. Stefánsson and J. Jakobsson 1992, Marine Research Institute, Iceland, unpublished data.	Gudmunder Óskarsson, Marine Research Institute, Skulagata, gjos@hafro.is
NFCBSS	Conception Bay & Southern Shore	NAFO 3KL and 3Ps	1966-2014			NA	East Southeast Newfoundland	Fisheries and Oceans Canada (DFO)	Bourne, C. M., Fisheries & Oceans Canada, Christina.Bourne@dfo-mpo.gc.ca
NFBBTB	Bonavista & Trinity Bays	NAFO 3KL and 3Ps	1966-2014	1988-2014	1988-2013	Survey	East Southeast Newfoundland	Fisheries and Oceans Canada (DFO)	Bourne, C. M., Fisheries & Oceans Canada, Christina.Bourne@dfo-mpo.gc.ca
NFSMPB	St Mary’s & Placentia Bays	NAFO 3KL and 3Ps	1966-2014	1982-2012	1982-2012	Survey	East Southeast Newfoundland	Fisheries and Oceans Canada (DFO)	Bourne, C. M., Fisheries & Oceans Canada, Christina.Bo

									urne@dfo-mpo.gc.ca
NFWNDB	White & Notre Dame Bays	NAFO 3KL and 3Ps	1966-2014	1988-2012	1988-2012	Survey	East Southeast Newfoundland	Fisheries and Oceans Canada (DFO)	Bourne, C. M., Fisheries & Oceans Canada, Christina.Bourne@dfo-mpo.gc.ca
NFFORB	Fortune Bay	NAFO 3KL and 3Ps	1966-2014	1982-2014	1988-2013	Survey	East Southeast Newfoundland	Fisheries and Oceans Canada (DFO)	Bourne, C. M., Fisheries & Oceans Canada, Christina.Bourne@dfo-mpo.gc.ca
WNFALL	West Newfoundland fall spawners	NAFO 4R	1965-2014	1965-2014	1965-2014	Virtual population analysis	West Newfoundland	Fisheries and Oceans Canada (DFO)	Doniol-Valcroze, T., Fisheries & Oceans Canada, Thomas.Doniol-Valcroze@dfo-mpo.gc.ca
WNSPRG	West Newfoundland spring spawners	NAFO 4R	1965-2014	1965-2014	1965-2014	Virtual population analysis	West Newfoundland	Fisheries and Oceans Canada (DFO)	Doniol-Valcroze, T., Fisheries & Oceans Canada, Thomas.Doniol-Valcroze@dfo-mpo.gc.ca
NSTLAW	Northern Gulf of St. Lawrence	NAFO 4S	1979-2014			NA	Northern Gulf of St. Lawrence	Fisheries and Oceans Canada (DFO)	
SSTLAS	Southern Gulf of St. Lawrence spring spawners	NAFO 4T	1981-2014	1978-2014	1978-2014	Virtual population analysis	Southern Gulf of St. Lawrence	Fisheries and Oceans Canada (DFO)	McDermid, J., Fisheries & Oceans Canada, Jenni.McDermid@dfo-mpo.gc.ca

SSTLAF	Southern Gulf of St. Lawrence spring spawners	NAFO 4T	1981-2014	1978-2014	1978-2014	Virtual population analysis	Southern Gulf of St. Lawrence	Fisheries and Oceans Canada (DFO)	McDermid, J., Fisheries & Oceans Canada, Jenni.McDermid@dfo-mpo.gc.ca
SCOFUN	Scotian Shelf and Bay of Fundy	4VWX	1963-2014	1965-2006	1964-2005	Virtual population analysis	Scotian Shelf	**Spawning biomass & Recruitment:** *Ricard, D., Minto, C., Jensen, O.P. and Baum, J.K. (2013) Evaluating the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish and Fisheries 13 (4) 380-398. DOI: 10.1111/j.1467-2979.2011.00435.x **Catches:** Fisheries and Oceans Canada (DFO)	
GOMGEB	Gulf of Maine/Georges Bank		1965-2011	1965-2011	1965-2011	Statistical catch-at-age model	Gulf of Maine	National Oceanic and Atmospheric Administration (NOAA), USA	Deroba, J., NOAA Northeast Fisheries Science Center, jonathan.deroba@noaa.gov
CALCRE	Crescent City		1973-2006			NA	California	Mello, J. (2006) Summary of 2005-2006 Pacific Herring Spawning-Ground Surveys and Commercial Catch in Humboldt Bay and Crescent City. California Department of Fish and Wildlife. https://wildlife.ca.gov/Fishing/Commercial/Herring/Season-Summaries.	
CALHUM	Humboldt Bay		1974-2007	1975-1976, 1991-1992, 2001-2007		Survey	California	Mello, J. (2007) Summary of 2006-2007 Pacific Herring Spawning-Ground Surveys and Commercial Catch in Humboldt Bay and Crescent City. California Department of Fish and Wildlife. https://wildlife.ca.gov/Fishing/Commercial/Herring/Season-Summaries.	
CALSFB	San Francisco Bay		1973-2014	1980-2002, 2004-2014	1983-2002, 2004-2014	Survey	California	Greiner, T., Bartling, R., and Weltz, A. (2014) Summary of the 2013-2014 Pacific Herring Spawning	Greiner, T. A., California Department

							Population and Commercial Fisheries in San Francisco Bay. California Department of Fish and Wildlife, Aquaculture and Bay Management Project, Herring Research and Management.	of Fish and Wildlife, USA, Tom.Greiner @wildlife.ca. gov
CALTOM	Tomales Bay	1973-2006	1974-1978, 1980-2006	1994-1997, 1999-2005	Survey	California	Watanabe, R. (2006) Summary of the 2005-06 Tomales Bay Herring Fishery Season. California Department of Fish and Wildlife. https://wildlife.ca.gov/Fishing/Commercial/Herring/Season-Summaries.	Douglas Hay, Fisheries & Oceans Canada, hayd@pac.df o-mpo.gc.ca
PSCHER	Cherry Point	1973-2014	1973-2014	1974-2008	Survey	Puget Sound	Stick, K. C., Lindquist, A., and Lowry, D. (2014). 2012 Washington State Herring Stock Status Report. Fish Program Technical Report No. FPA 14-09Washington Department of Fish and Wildlife, Fish Program, Fish Management Division.	Stick, K. C., Washington Department of Fish and Wildlife, USA, Kurt.Stick@ dfw.wa.gov
PSSQUA	Squaxin Pass		1975-1981, 1990-2014	1991-2009	Survey	Puget Sound	Stick, K. C., Lindquist, A., and Lowry, D. (2014). 2012 Washington State Herring Stock Status Report. Fish Program Technical Report No. FPA 14-09Washington Department of Fish and Wildlife, Fish Program, Fish Management Division. **Catch time series is not available for this stock*	Stick, K. C., Washington Department of Fish and Wildlife, USA, Kurt.Stick@ dfw.wa.gov
PSOTHR	Puget Sound (excluding Cherry Point and Squaxin)	1965-2014	1975-2014	1988-2009	Survey	Puget Sound	Stick, K. C., Lindquist, A., and Lowry, D. (2014). 2012 Washington State Herring Stock Status Report. Fish Program Technical Report No. FPA 14-09Washington Department of Fish and Wildlife, Fish Program, Fish Management Division.	Stick, K. C., Washington Department of Fish and Wildlife, USA, Kurt.Stick@ dfw.wa.gov
BCGEOR	Strait of Georgia	1951-2014	1951-2014	1953-2014	Integrated statistical catch-at-age model	British Columbia	DFO. 2015. Stock Assessment and Management Advice for BC Pacific Herring: 2015 Status and 2016 Forecast. DFO Can. Sci. Advis. Sec. Sci. Resp. 2015/038.	Cleary, J., Pacific Biological Station, Fisheries & Oceans Canada, Jaclyn.Cleary

								@dfo-mpo.gc.ca
BCWVAN	West Coast Vancouver Island	1951-2014	1951-2014	1953-2014	Integrated statistical catch-at-age model	British Columbia	DFO. 2015. Stock Assessment and Management Advice for BC Pacific Herring: 2015 Status and 2016 Forecast. DFO Can. Sci. Advis. Sec. Sci. Resp. 2015/038.	Cleary, J., Pacific Biological Station, Fisheries & Oceans Canada, Jaclyn.Cleary@dfo-mpo.gc.ca
BCCECO	Central Coast	1951-2014	1951-2014	1953-2014	Integrated statistical catch-at-age model	British Columbia	DFO. 2015. Stock Assessment and Management Advice for BC Pacific Herring: 2015 Status and 2016 Forecast. DFO Can. Sci. Advis. Sec. Sci. Resp. 2015/038.	Cleary, J., Pacific Biological Station, Fisheries & Oceans Canada, Jaclyn.Cleary@dfo-mpo.gc.ca
BCPRUP	Prince Rupert District	1951-2014	1951-2014	1953-2014	Integrated statistical catch-at-age model	British Columbia	DFO. 2015. Stock Assessment and Management Advice for BC Pacific Herring: 2015 Status and 2016 Forecast. DFO Can. Sci. Advis. Sec. Sci. Resp. 2015/038.	Cleary, J., Pacific Biological Station, Fisheries & Oceans Canada, Jaclyn.Cleary@dfo-mpo.gc.ca
BCHAGW	Haida Gwaii	1951-2014	1951-2014	1953-2014	Integrated statistical catch-at-age model	British Columbia	DFO. 2015. Stock Assessment and Management Advice for BC Pacific Herring: 2015 Status and 2016 Forecast. DFO Can. Sci. Advis. Sec. Sci. Resp. 2015/038.	Cleary, J., Pacific Biological Station, Fisheries & Oceans Canada, Jaclyn.Cleary@dfo-mpo.gc.ca

SEACRA	Craig	1980-2013	1980-2014	1988-2014	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
SEAERN	Ernest Sound	1980-1992, 1994-1995, 1998, 2000-2009, 2012-2013	1980-1986, 1989-1996, 1998, 2000-2014	1995-2014	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
SEAHOB	Hobart Bay-Port Houghton	1988-1993, 1995-2014	1988-2014	1995-2014	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
SEAHOO	Hoonah Sound	1981-2014	1981-2014	1995-2014	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
SEAKAH	Kah Shakes-Cat Island	1980-2001	1980-2001	1980-2001	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
SEASEY	Seymour Canal	1980-2010, 2012-2013	1980-2014	1980-2011	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
SEATEN	Tenakee Inlet	1980-2009, 2011-2013	1980-2010, 2012-2014	1997-2014	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and

								Game, USA, sherri.dressel@alaska.gov
SEASIT	Sitka	1980-2014	1980-2014	1980-2014	Statistical catch-at-age model	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel@alaska.gov
SEABEH	West Behm Canal	1980-2010, 2011-2014	1980-2014	1996-2014	Survey	Southeast Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel@alaska.gov
PRIWIL	Prince William Sound	1980-2014	1980-2014	1980-2014	Statistical catch-at-age model	Central Alaska	School of Aquatic & Fishery Science, University of Washington, Seattle, USA	Branch, T., University of Washington, USA, tbranch@uw.edu
KKKODI	Kamishak	1985-2014	1985-2014	1985-2014	Statistical catch-at-age model	Central Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel@alaska.gov
SWATOG	Togiak	1980-2000	1978-2014	1978-2014	Statistical catch-at-age model	Central Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel@alaska.gov
KKKODI	Kodiak	1980-2014		1980-2012	NA	Westward Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel@alaska.gov

KKAPEN	Alaska Peninsula & Chignik	1979-2014		NA	Westward Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
SWADUT	Dutch Harbor	1981-2014		NA	Westward Alaska	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
AYKGOO	Goodnews Bay	1981-2014	1981-1984, 1986, 1988-1993, 1996, 2000-2001, 2005, 2013-2014	Survey	Arctic-Yukon-Kuskokwim	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
AYKSEC	Security Cove	1981-2014	1981, 1983-1984, 1988-1993, 1996, 1999-2003, 2005, 2013-2014	Survey	Arctic-Yukon-Kuskokwim	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
AYKAVI	Cape Avinof	1988-2014	1988, 1991-1992, 2014	Survey	Arctic-Yukon-Kuskokwim	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
AYKNOR	Norton Sound	1979-2014	1979-2005. 2011	Survey	Arctic-Yukon-Kuskokwim	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
AYKNEL	Nelson Island	1985-2014	1986-1995, 1999, 2013-2014	Survey	Arctic-Yukon-Kuskokwim	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and

							Game, USA, sherri.dressel @alaska.gov
AYKROM	Cape Romanzof	1980-2014	1987, 2006	Survey	Arctic-Yukon-Kuskokwim	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
AYKNUN	Nunivak Island	1985-2014	1986, 1989-1994, 1998, 2000-2001	Survey	Arctic-Yukon-Kuskokwim	Alaska Department of Fish and Game (ADFG), USA	Dressel, S., Alaska Department of Fish and Game, USA, sherri.dressel @alaska.gov
KORKAR	Korf-Karaginsk	1939-2013		NA	Korf-Karaginsk	**1939-1977**: Hay, D. E., Toresen, R., Stephenson, R., et al. (2001). Taking stock: an inventory and review of world herring stocks in 2000. Herring: Expectations for a new millennium. Univ. of Alaska Sea Grant, AKSG-01-04, Fairbanks, 381-454. **1978-2013**: Radchenko, V. I., North Pacific Anadromous Fish Commission, vlrad@npafc.org	
GIZKAM	Gizigha-Kamchatka	1913-1999		NA	Sea of Okhotsk	Hay, D. E., Toresen, R., Stephenson, R., et al. (2001). Taking stock: an inventory and review of world herring stocks in 2000. Herring: Expectations for a new millennium. Univ. of Alaska Sea Grant, AKSG-01-04, Fairbanks, 381-454.	Douglas Hay, Fisheries & Oceans Canada, hayd@pac.dfo-mpo.gc.ca
OKHSEA	Okhotsk	1945-2013	1945-2000	Survey	Sea of Okhotsk	**1945-1964**: Hay, D. E., Toresen, R., Stephenson, R., et al. (2001). Taking stock: an inventory and review of world herring stocks in 2000. Herring: Expectations for a new millennium. Univ. of Alaska Sea Grant, AKSG-01-04, Fairbanks, 381-454. **1965-2013**: Radchenko, V. I., North Pacific Anadromous Fish Commission, vlrad@npafc.org	

							Spawning biomass: Magadan branch of TINRO (Naumenko 2001) (provided by Radchenko, V. I., North Pacific Anadromous Fish Commission, vlrad@npafc.org)	
DEK	Dekastri	1926-1999			NA	Sea of Japan	Hay, D. E., Toresen, R., Stephenson, R., et al. (2001). Taking stock: an inventory and review of world herring stocks in 2000. Herring: Expectations for a new millennium. Univ. of Alaska Sea Grant, AKSG-01-04, Fairbanks, 381-454.Hay et al. 2001	Douglas Hay, Fisheries & Oceans Canada, hayd@pac.dfo-mpo.gc.ca
HOKSAK	Hokkaido-Sakhalin	1885-2013	1907-1957*	1907-1956*	Survey	Sea of Japan	**1885-1974**: Hay, D. E., Toresen, R., Stephenson, R., et al. (2001). Taking stock: an inventory and review of world herring stocks in 2000. Herring: Expectations for a new millennium. Univ. of Alaska Sea Grant, AKSG-01-04, Fairbanks, 381-454. **Spawning biomass and recruitment time series are not presented nor analyzed in this study* **1975-2013**: Nishin, 2014 Hokkaido herring stock assessment (provided by Hiroshi Okamura).	Hiroshi Okamura, okamura@fra.affrc.go.jp
PETGRB	Peter the Great Bay	1910-1999			NA	Sea of Japan	Hay, D. E., Toresen, R., Stephenson, R., et al. (2001). Taking stock: an inventory and review of world herring stocks in 2000. Herring: Expectations for a new millennium. Univ. of Alaska Sea Grant, AKSG-01-04, Fairbanks, 381-454.	Douglas Hay, Fisheries & Oceans Canada, hayd@pac.dfo-mpo.gc.ca

Table A2: Spawning and other biological characteristics of select herring stocks

Herring stock	*Latitude (°N)*	*Longitude (°W if <0, °E if >0)*	*Month of peak spawn*	*First age-at-maturity (year)*	*Age-at-recruitment (year)*	*Mean age 5 body weight (g)*	*Log(Max historical catch) (metric tons)*
Baltic Sea	57.06	19.48	May	2	1	61	12.82
Bothnian Bay	64.77	22.84	May	2	1	61	9.18
Bothnian Sea	61.91	19.21	May	2	1	61	11.61
Gulf of Riga	57.70	23.60	May	2	1	61	10.62
Central Coast	52.07	-128.36	Mar	2	2	146	10.69
Strait of Georgia	50.00	-124.00	Mar	2	2	146	11.25
Haida Gwaii	52.50	-130.87	Mar	2	2	150	11.26
Prince Rupert District	54.37	-130.54	Mar	2	2	146	10.70
West Coast Vancouver Island	49.50	-127.00	Mar	2	2	146	11.15
San Francisco Bay	37.70	-122.27	Jan	2	3	110	9.25
Irish Sea	53.56	-5.16	Nov	2	1	185	10.56
Northwest Ireland	55.74	-9.47	Nov	2	1	185	10.79
West of Ireland	52.61	-11.23	Nov	2	1	185	10.70
Gulf of Maine/Georges Bank	42.80	-70.40	Oct	3	1	180	13.08
Icelandic summer spawners	66.62	-19.25	Jul	4	3	245	11.97
Kamishak	59.21	-153.92	Jun	3	3	150	8.62
St Mary's & Placentia Bays	46.56	-54.59	May	2	4	220	1.38
Norwegian spring spawners	63.00	5.85	Mar	5	0	206	14.49
North Sea autumn spawners	56.04	3.15	Oct	2	0	220	13.97
Prince William Sound	60.00	-148.00	Apr	3	3	125	9.91
Cherry Point	48.96	-122.87	May	2	2	160	8.30
Scotian Shelf and Bay of Fundy	43.00	-65.40	Sep	3	1	260	12.21
Seymour Canal	59.23	-134.21	Mar	3	3	150	7.23
Sitka Sound	57.17	-135.33	Mar	3	3	150	9.78
S. Gulf of St. Lawrence fall spawners	47.19	-64.70	Sep	5	4	290	11.11
S. Gulf of St. Lawrence spring spawners	47.04	-64.70	May	5	4	230	10.18
Togiak	59.05	-160.37	Jun	4	3	200	10.22

Herring stock	*Latitude (˚N)*	*Longitude (˚W if <0, ˚E if >0)*	*Month of peak spawn*	*First age-at-maturity (year)*	*Age-at-recruitment (year)*	*Mean age 5 body weight (g)*	*Log(Max historical catch) (metric tons)*
West of Scotland	57.12	-6.86	Oct	2	1	195	12.26
W. Newfoundland fall spawners	50.80	-57.40	Sep	5	3	200	9.82
W. Newfoundland spring spawners	48.40	-58.80	May	4	3	230	9.86

Table A3: Environmental characteristics near herring locations during spawning

Herring stock	*SST (˚C)*				*SSHA (cm)*				*Month*
	Mean	*St. Dev.*	*Slope of trend*	*Years*	*Mean*	*St. Dev.*	*Slope of trend*	*Years*	
Baltic Sea	7.40	0.94	0.05	1986-2014	6.20	5.97	0.48	1986-2009	May
Bothnian Bay	3.90	0.89	0.06	1985-2013	5.30	5.66	0.43	1985-2009	May
Bothnian Sea	5.20	0.81	0.06	1986-2014	5.60	5.61	0.39	1986-2009	May
Gulf of Riga	6.70	0.94	0.07	1986-2014	6.10	6.23	0.46	1986-2009	May
Central Coast	7.80	0.62	-0.02	1986-2014	2.90	4.07	0.12	1986-2009	Mar
Strait of Georgia	8.60	0.80	-0.03	1986-2014	2.80	3.25	0.11	1986-2009	Mar
Haida Gwaii	7.40	0.55	-0.02	1986-2014	2.60	4.20	0.10	1986-2009	Mar
Prince Rupert District	6.90	0.51	-0.01	1986-2014	3.10	4.21	0.14	1986-2009	Mar
West Coast Vancouver Island	8.20	0.70	-0.03	1986-2014	2.30	3.01	0.10	1986-2009	Mar
San Francisco Bay	12.10	0.84	-0.02	1986-2014	1.80	4.73	0.03	1986-2009	Jan
Irish Sea	11.80	0.58	0.04	1985-2013	3.20	2.74	0.24	1985-2008	Nov
Northwest Ireland	11.40	0.47	0.04	1985-2013	3.40	2.62	0.28	1985-2008	Nov
West of Ireland	12.60	0.53	0.03	1985-2013	3.40	2.33	0.29	1985-2008	Nov
Gulf of Maine/Georges Bank	14.30	0.77	0.04	1983-2011	2.70	2.64	0.19	1983-2008	Oct
Icelandic summer spawners	8.20	0.72	0.04	1986-2014	3.20	2.33	0.26	1986-2008	Jul
Kamishak	8.60	0.57	0.00	1986-2014	4.80	3.16	0.30	1986-2009	Jun
St Mary's & Placentia Bays	3.70	1.14	0.02	1984-2012	1.90	2.82	0.26	1984-2009	May
Norwegian spring spawners	7.10	0.41	0.03	1986-2014	5.10	3.89	0.28	1986-2009	Mar
North Sea autumn spawners	12.70	0.78	0.05	1986-2014	2.80	2.95	0.29	1986-2008	Oct
Prince William Sound	5.20	0.37	-0.01	1986-2014	2.80	2.80	0.14	1986-2009	Apr

Herring stock	*SST (˚C)*				*SSHA (cm)*				*Month*
	Mean	*St. Dev.*	*Slope of trend*	*Years*	*Mean*	*St. Dev.*	*Slope of trend*	*Years*	
Cherry Point	11.90	0.79	0.02	1980-2008	2.40	2.89	0.16	1980-2008	May
Scotian Shelf and Bay of Fundy	17.40	0.86	0.03	1975-2003	0.20	1.81	0.18	1975-2003	Sep
Seymour Canal	5.80	0.67	0.01	1983-2011	2.90	3.49	0.13	1983-2009	Mar
Sitka Sound	6.00	0.57	0.00	1986-2014	3.50	3.38	0.15	1986-2009	Mar
S. Gulf of St. Lawrence fall spawners	13.20	0.61	0.06	1987-2015	1.80	1.89	0.23	1987-2008	Sep
S. Gulf of St. Lawrence spring spawners	4.50	0.57	0.04	1987-2015	1.80	2.38	0.24	1987-2009	May
Togiak	6.50	1.03	0.01	1986-2014	3.40	2.77	0.23	1986-2009	Jun
West of Scotland	12.20	0.53	0.04	1985-2013	2.90	2.57	0.28	1985-2008	Oct
W. Newfoundland fall spawners	11.90	0.69	0.05	1986-2014	2.10	2.18	0.27	1986-2008	Sep
W. Newfoundland spring spawners	3.20	0.77	0.03	1986-2014	1.30	2.82	0.23	1986-2009	May

Table A4: Environmental characteristics near herring locations after spawning

Herring stock	*SST (˚C)*				*SSHA (cm)*				*Month*
	Mean	*St. Dev.*	*Slope of trend*	*Years*	*Mean*	*St. Dev.*	*Slope of trend*	*Years*	
Baltic Sea	15.60	0.92	0.06	1986-2014	5.80	4.88	0.55	1986-2009	6-8
Bothnian Bay	12.00	0.56	0.04	1985-2013	5.40	4.61	0.45	1985-2009	6-8
Bothnian Sea	13.40	0.85	0.07	1986-2014	5.40	4.55	0.44	1986-2009	6-8
Gulf of Riga	15.20	0.86	0.07	1986-2014	6.10	5.42	0.59	1986-2009	6-8
Central Coast	10.50	0.59	-0.01	1986-2014	3.20	3.19	0.21	1986-2009	4-6
Strait of Georgia	12.00	0.62	0.00	1986-2014	2.70	2.62	0.17	1986-2009	4-6
Haida Gwaii	9.90	0.58	-0.01	1986-2014	3.10	3.56	0.21	1986-2009	4-6
Prince Rupert District	9.60	0.56	-0.00	1986-2014	3.20	3.37	0.20	1986-2009	4-6
West Coast Vancouver Island	10.80	0.60	-0.01	1986-2014	2.60	2.66	0.20	1986-2009	4-6
San Francisco Bay	12.00	0.79	-0.04	1986-2014	2.40	3.82	0.09	1986-2009	2-4
Irish Sea	9.00	0.47	0.02	1985-2013	3.80	2.40	0.21	1985-2008	12-2
Northwest Ireland	10.00	0.34	0.02	1985-2013	3.90	2.33	0.22	1985-2008	12-2
West of Ireland	11.00	0.34	0.02	1985-2013	3.20	1.77	0.23	1985-2008	12-2
Gulf of Maine/Georges Bank	8.50	0.61	0.02	1983-2011	2.30	2.69	0.16	1983-2008	11-1
Icelandic summer spawners	7.30	0.66	0.05	1986-2014	3.50	2.28	0.28	1986-2008	8-10
Kamishak	11.40	0.67	0.00	1986-2014	3.60	2.47	0.19	1986-2009	7-9
St Mary's & Placentia Bays	11.90	1.15	0.04	1984-2012	1.90	2.44	0.28	1984-2009	6-8
Norwegian spring spawners	8.70	0.41	0.03	1986-2014	4.30	2.91	0.31	1986-2009	4-6
North Sea autumn spawners	8.50	0.66	0.03	1986-2014	4.80	3.77	0.22	1986-2008	11-1
Prince William Sound	9.70	0.56	0.01	1986-2014	3.00	2.59	0.22	1986-2009	5-7

Herring stock	*SST (˚C)*				*SSHA (cm)*				*Month*
	Mean	*St. Dev.*	*Slope of trend*	*Years*	*Mean*	*St. Dev.*	*Slope of trend*	*Years*	
Cherry Point	14.40	0.76	0.03	1980-2008	2.50	2.31	0.18	1980-2008	6-8
Scotian Shelf and Bay of Fundy	12.00	0.70	0.01	1975-2003	1.00	2.16	0.18	1975-2003	10-12
Seymour Canal	8.20	0.65	0.02	1983-2011	2.80	2.86	0.10	1983-2009	4-6
Sitka Sound	8.40	0.58	0.01	1986-2014	3.30	2.67	0.20	1986-2009	4-6
S. Gulf of St. Lawrence fall spawners	5.50	0.56	0.06	1987-2015	2.10	2.03	0.22	1987-2008	10-12
S. Gulf of St. Lawrence spring spawners	13.10	0.60	0.05	1987-2015	1.70	2.11	0.26	1987-2009	6-8
Togiak	11.10	1.12	0.02	1986-2014	3.00	2.15	0.20	1986-2009	7-9
West of Scotland	9.90	0.41	0.03	1985-2013	3.90	2.37	0.24	1985-2008	11-1
W. Newfoundland fall spawners	4.70	0.59	0.05	1986-2014	2.00	2.05	0.25	1986-2008	10-12
W. Newfoundland spring spawners	11.60	0.81	0.06	1986-2014	1.50	2.30	0.26	1986-2009	6-8

Table A5: Names and trends of herring predator populations used in our analysis

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
Arrowtooth flounder Bering Sea and Aleutian Islands	Fish	0.993	0.716	1.269	Cape Avinof
					Cape Romanzof
					Goodnews Bay
					Nelson Island
					Norton Sound
					Nunivak Island
					Security Cove
					Alaska Peninsula & Chignik
					Dutch Harbor
					Togiak
Arrowtooth flounder Gulf of Alaska	Fish	0.966	0.699	1.233	Kamishak
					Prince William Sound
					Togiak
					Craig
					Hobart Bay-Port Houghton
					Hoonah Sound
					Kah Shakes-Cat Island
					Seymour Canal
					Ernest Sound
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
Atlantic cod Baltic Areas 22 and 24	Fish	0.338	-0.038	0.713	Western Baltic spring spawners
Atlantic cod Baltic Areas 25-32	Fish	-0.230	-0.438	-0.022	Baltic Sea
Atlantic Cod Celtic Sea	Fish	-0.333	-1.057	0.391	West of Ireland
Atlantic cod Georges Bank	Fish	-0.898	-1.147	-0.650	Gulf of Maine/Georges Bank
Atlantic cod Gulf of Maine	Fish	-0.360	-0.663	-0.056	Gulf of Maine/Georges Bank
Atlantic cod Iceland	Fish	-0.245	-0.627	0.136	Icelandic summer spawners

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
Atlantic cod Irish Sea	Fish	-0.879	-1.027	-0.731	Irish Sea
Atlantic cod Kattegat	Fish	-0.778	-0.968	-0.587	Western Baltic spring spawners
Atlantic cod NAFO 2J3KL	Fish	0.147	-0.011	0.305	Bonavista & Trinity Bays
					Conception Bay & Southern Shore
					White & Notre Dame Bays
Atlantic cod NAFO 3NO	Fish	-0.202	-0.348	-0.055	St Mary's & Placentia Bays
Atlantic cod NAFO 3Pn4RS	Fish	-0.054	-0.126	0.019	W. Newfoundland fall spawners
					W. Newfoundland spring spawners
					Northern Gulf of St. Lawrence
Atlantic cod NAFO 3Ps	Fish	-0.437	-0.886	0.012	Fortune Bay
					St Mary's & Placentia Bays
Atlantic cod NAFO 4TVn	Fish	-0.507	-0.606	-0.407	S. Gulf of St. Lawrence fall spawners
					S. Gulf of St. Lawrence spring spawners
Atlantic cod NAFO 4X	Fish	-0.839	-1.002	-0.677	Scotian Shelf and Bay of Fundy
Atlantic cod North Sea	Fish	-0.867	-1.044	-0.690	North Sea autumn spawners
Atlantic cod Northeast Arctic	Fish	0.416	0.217	0.615	Norwegian spring spawners
Atlantic cod West of Scotland	Fish	-0.950	-1.261	-0.640	Northwest Ireland
					West of Scotland
Atlantic Halibut NAFO-5YZ	Fish	0.363	-0.886	1.613	Gulf of Maine/Georges Bank
					Scotian Shelf and Bay of Fundy
Beluga Whale Cook Inlet, Alaska	Cetacean	-0.727	-1.332	-0.122	Kamishak
California Sea Lion NE Pacific (California)	Pinniped	0.040	-1.101	1.181	San Francisco Bay
					Tomales Bay
					Crescent City
					Humboldt Bay

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
Fin_whale N Atlantic	Cetacean	0.856	0.080	1.633	Icelandic summer spawners
					Norwegian spring spawners
Grey Seal Baltic Sea	Pinniped	1.109	0.165	2.054	Baltic Sea
					Bothnian Bay
					Bothnian Sea
					Gulf of Riga
Grey Seal Eastern Shore, Nova Scotia	Pinniped	0.631	0.503	0.758	Fortune Bay
					St Mary's & Placentia Bays
					Scotian Shelf and Bay of Fundy
Grey Seal Great Britain (Combined)	Pinniped	1.196	1.109	1.282	Irish Sea
					Northwest Ireland
					West of Ireland
					West of Scotland
Grey Seal Gulf of St. Lawrence, Canada	Pinniped	1.928	1.801	2.055	Northern Gulf of St. Lawrence
					S. Gulf of St. Lawrence fall spawners
					S. Gulf of St. Lawrence spring spawners
Grey Seal Inner Hebrides, Scotland	Pinniped	0.177	-0.033	0.386	Irish Sea
Grey Seal Netherlands	Pinniped	0.399	0.202	0.595	North Sea autumn spawners
Grey Seal North Sea	Pinniped	1.014	0.981	1.047	Norwegian spring spawners
Grey Seal Orkney Islands, Scotland	Pinniped	1.038	1.000	1.075	West of Scotland
					Irish Sea
Grey Seal Outer Hebrides, Scotland	Pinniped	-0.344	-0.576	-0.112	West of Scotland
					Northwest Ireland
Grey Seal Sable Island, Nova Scotia	Pinniped	0.836	0.707	0.966	Northern Gulf of St. Lawrence
					S. Gulf of St. Lawrence fall spawners
					S. Gulf of St. Lawrence spring spawners

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					St Mary's & Placentia Bays
					White & Notre Dame Bays
					Bonavista & Trinity Bays
					Conception Bay & Southern Shore
					Fortune Bay
					W. Newfoundland fall spawners
					W. Newfoundland spring spawners
Grey Seal Schleswig-Holstein, Germany	Pinniped	1.268	-0.185	2.721	Western Baltic spring spawners
Haddock Georges Bank	Fish	0.434	0.140	0.728	Gulf of Maine/Georges Bank
Haddock Iceland	Fish	0.175	-0.340	0.691	Icelandic summer spawners
Haddock ICES IIIa and North Sea	Fish	-0.349	-0.520	-0.178	North Sea autumn spawners
Haddock ICES VIIb-k	Fish	0.529	0.322	0.736	West of Ireland
Haddock Irish Sea	Fish	0.625	0.235	1.014	Irish Sea
Haddock NAFO-4X5Y	Fish	-0.212	-0.592	0.168	Scotian Shelf and Bay of Fundy
Haddock Northeast Arctic	Fish	0.383	0.218	0.547	Norwegian spring spawners
Haddock Rockall Bank	Fish	-0.917	-1.233	-0.602	Northwest Ireland
Haddock West of Scotland	Fish	-0.510	-0.959	-0.060	West of Scotland
Hake Northeast Atlantic North	Fish	-0.616	-1.043	-0.189	Irish Sea
					West of Ireland
Harbour Porpoise San Francisco - Russian River Stock	Cetacean	0.733	-1.676	3.143	San Francisco Bay
					Tomales Bay
Harbour Seal E Scotland	Pinniped	-0.055	-1.394	1.285	North Sea autumn spawners
Harbour Seal England (Combined)	Pinniped	1.437	1.133	1.741	Irish Sea
Harbour Seal Grand Manan Island, New Brunswick	Pinniped	0.997	0.854	1.140	Gulf of Maine/Georges Bank
					Scotian Shelf and Bay of Fundy

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
Harbour Seal N Scotland	Pinniped	-0.915	-1.474	-0.356	Northwest Ireland
					West of Ireland
					West of Scotland
Harbour Seal Wadden Sea	Pinniped	0.402	0.241	0.562	North Sea autumn spawners
Harbour Seal Washington (Inland)	Pinniped	0.934	0.722	1.147	Cherry Point
					Puget Sound
					Squaxin Pass
Harp Seal NW Atlantic	Pinniped	0.966	0.732	1.199	Northern Gulf of St. Lawrence
					S. Gulf of St. Lawrence fall spawners
					S. Gulf of St. Lawrence spring spawners
					St Mary's & Placentia Bays
					White & Notre Dame Bays
					Bonavista & Trinity Bays
					Conception Bay & Southern Shore
					Fortune Bay
					W. Newfoundland fall spawners
					W. Newfoundland spring spawners
Hooded Seal NW Atlantic Stock	Pinniped	0.967	0.812	1.122	Northern Gulf of St. Lawrence
					S. Gulf of St. Lawrence fall spawners
					S. Gulf of St. Lawrence spring spawners
					St Mary's & Placentia Bays
					White & Notre Dame Bays
					Bonavista & Trinity Bays
					Conception Bay & Southern Shore
					Fortune Bay

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					W. Newfoundland fall spawners
					W. Newfoundland spring spawners
					Icelandic summer spawners
Humpback whale N Atlantic	Cetacean	0.625	0.368	0.883	Scotian Shelf and Bay of Fundy
					Gulf of Maine/Georges Bank
					Northern Gulf of St. Lawrence
					S. Gulf of St. Lawrence fall spawners
					S. Gulf of St. Lawrence spring spawners
					St Mary's & Placentia Bays
					White & Notre Dame Bays
					Bonavista & Trinity Bays
					Conception Bay & Southern Shore
					Fortune Bay
					W. Newfoundland fall spawners
					W. Newfoundland spring spawners
					Icelandic summer spawners
					Norwegian spring spawners
Humpback whale N Pacific	Cetacean	0.424	0.080	0.768	Kamishak
					Prince William Sound
					Togiak
					Craig
					Ernest Sound
					Hobart Bay-Port Houghton
					Hoonah Sound
					Kah Shakes-Cat Island

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					Seymour Canal
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
					Central Coast
					Haida Gwaii
					Prince Rupert District
					Strait of Georgia
					West Coast Vancouver Island
Humpback whale U.S. West Coast	Cetacean	0.734	0.238	1.229	Cherry Point
					Puget Sound
					Squaxin Pass
					San Francisco Bay
					Tomales Bay
					Crescent City
					Humboldt Bay
Killer Whale NE Pacific - N Residents	Cetacean	1.021	0.949	1.094	Central Coast
					Haida Gwaii
					Prince Rupert District
					Craig
					Ernest Sound
					Hobart Bay-Port Houghton
					Hoonah Sound
					Kah Shakes-Cat Island
					Seymour Canal
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
Killer Whale NE Pacific - S Residents	Cetacean	0.619	0.110	1.129	Strait of Georgia
					West Coast Vancouver Island
					Cherry Point

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					Puget Sound
					Squaxin Pass
Mackerel ICES Northeast Atlantic	Fish	-0.503	-1.425	0.420	Northwest Ireland
					West of Ireland
					West of Scotland
					North Sea autumn spawners
					Norwegian spring spawners
					Icelandic summer spawners
Minke whale NE Atlantic	Cetacean	0.787	0.043	1.530	Icelandic summer spawners
					Norwegian spring spawners
Northern Fur Seal Pribilof Islands, Alaska	Pinniped	-0.941	-1.211	-0.671	Cape Avinof
					Cape Romanzof
					Goodnews Bay
					Nelson Island
					Norton Sound
					Nunivak Island
					Security Cove
					Togiak
Pacific cod Bering Sea and Aleutian Islands	Fish	-0.581	-0.856	-0.305	Cape Avinof
					Cape Romanzof
					Goodnews Bay
					Nelson Island
					Norton Sound
					Nunivak Island
					Security Cove
					Alaska Peninsula & Chignik
					Dutch Harbor
					Togiak
Pacific cod Gulf of Alaska	Fish	-1.134	-1.256	-1.012	Kamishak
					Prince William Sound

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					Kodiak
					Craig
					Ernest Sound
					Hobart Bay-Port Houghton
					Hoonah Sound
					Kah Shakes-Cat Island
					Seymour Canal
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
Pacific cod Hecate Strait	Fish	-0.419	-0.586	-0.252	Central Coast
					Haida Gwaii
					Prince Rupert District
Pacific cod West Coast of Vancouver Island	Fish	-0.224	-0.484	0.036	West Coast Vancouver Island
Pacific hake Pacific Coast	Fish	-0.966	-1.186	-0.745	Central Coast
					Haida Gwaii
					Prince Rupert District
					Strait of Georgia
					West Coast Vancouver Island
Pacific halibut North Pacific	Fish	-1.086	-1.338	-0.834	Craig
					Ernest Sound
					Hobart Bay-Port Houghton
					Hoonah Sound
					Kah Shakes-Cat Island
					Seymour Canal
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
					Kamishak
					Prince William Sound
					Togiak

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					Alaska Peninsula & Chignik
					Dutch Harbor
					Kodiak
Pollock ICES IIIa, VI and North Sea	Fish	0.258	-0.208	0.723	North Sea autumn spawners
					West of Scotland
					Northwest Ireland
Pollock NAFO-4VWX5	Fish	-0.283	-0.613	0.048	Scotian Shelf and Bay of Fundy
Pollock NAFO-5YZ	Fish	1.115	0.807	1.423	Gulf of Maine/Georges Bank
Pollock Northeast Arctic	Fish	0.849	0.399	1.299	Norwegian spring spawners
Pollock or saithe Iceland Grounds	Fish	-0.411	-0.808	-0.014	Icelandic summer spawners
Ribbon seal Sea of Okhotsk	Pinniped	1.067	0.707	1.427	Gizigha-Kamchatka
					Okhotsk
Sablefish Eastern Bering Sea / Aleutian Islands / Gulf of Alaska	Fish	-0.784	-1.159	-0.409	Cape Avinof
					Cape Romanzof
					Goodnews Bay
					Nelson Island
					Norton Sound
					Nunivak Island
					Security Cove
					Kamishak
					Prince William Sound
					Togiak
					Alaska Peninsula & Chignik
					Dutch Harbor
					Kodiak
					Craig
					Ernest Sound
					Hobart Bay-Port Houghton
					Hoonah Sound

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					Kah Shakes-Cat Island
					Seymour Canal
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
Sablefish Pacific Coast of Canada	Fish	-0.946	-1.079	-0.814	Central Coast
					Haida Gwaii
					Prince Rupert District
					Strait of Georgia
					West Coast Vancouver Island
Spotted spiny dogfish Pacific Coast	Fish	-0.993	-1.070	-0.915	Cherry Point
					Puget Sound
					Squaxin Pass
					San Francisco Bay
					Tomales Bay
					Crescent City
					Humboldt Bay
					Central Coast
					Haida Gwaii
					Prince Rupert District
					Strait of Georgia
					West Coast Vancouver Island
Steller Sea Lion British Columbia	Pinniped	0.411	0.145	0.678	Central Coast
					Haida Gwaii
					Prince Rupert District
					Strait of Georgia
					West Coast Vancouver Island
Steller Sea Lion E Alaska (incl. SE AK, BC, WA, OR, CA)	Pinniped	1.245	1.092	1.398	Central Coast
					Haida Gwaii
					Prince Rupert District
					Strait of Georgia

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
					West Coast Vancouver Island
					Cherry Point
					Puget Sound
					Squaxin Pass
					San Francisco Bay
					Tomales Bay
					Crescent City
					Humboldt Bay
					Craig
					Ernest Sound
					Hoonah Sound
					Kah Shakes-Cat Island
					Seymour Canal
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
Steller Sea Lion W Alaska	Pinniped	-0.071	-0.142	0.001	Kamishak
					Prince William Sound
					Togiak
					Cape Avinof
					Cape Romanzof
					Goodnews Bay
					Nelson Island
					Norton Sound
					Nunivak Island
					Security Cove
					Alaska Peninsula & Chignik
					Dutch Harbor
					Kodiak
Walleye pollock Aleutian Islands	Fish	-0.866	-1.085	-0.646	Alaska Peninsula & Chignik
					Dutch Harbor

Name & area	*Type*	*Coefficient (slope)*	*Lower 95th conf. int.*	*Upper 95th conf. int.*	*Potential prey populations*
Walleye pollock Eastern Bering Sea	Fish	-0.483	-0.854	-0.112	Cape Avinof
					Cape Romanzof
					Goodnews Bay
					Nelson Island
					Norton Sound
					Nunivak Island
					Security Cove
					Togiak
Walleye pollock Gulf of Alaska	Fish	-0.921	-1.140	-0.702	Kamishak
					Prince William Sound
					Kodiak
					Craig
					Ernest Sound
					Hobart Bay-Port Houghton
					Hoonah Sound
					Kah Shakes-Cat Island
					Seymour Canal
					Sitka Sound
					Tenakee Inlet
					W. Behm Canal
White hake Georges Bank / Gulf of Maine	Fish	-0.981	-1.170	-0.793	Gulf of Maine/Georges Bank
Whiting ICES VIa	Fish	-0.769	-0.904	-0.634	West of Scotland
Whiting ICES VIIe-k	Fish	0.398	-0.224	1.020	Irish Sea
					Northwest Ireland
					West of Ireland
Whiting NS-VIId	Fish	-0.611	-0.938	-0.284	North Sea autumn spawners

Appendix B

DIC (Spiegelhalter et al., 2002) is calculated using the following equation:

$$\mathrm{DIC} = 2\mathbb{E}_{\theta|y}\left(-\ln(\mathrm{y}|\boldsymbol{\theta})\right) + \left[-2\ln\left(\mathrm{y}\middle|\mathbb{E}(\boldsymbol{\theta}|\mathrm{y})\right) + 2\mathbb{E}_{\boldsymbol{\theta}|y}\left(-\ln(\mathrm{y}|\boldsymbol{\theta})\right)\right]$$

The first term contains the log-likelihood of the data y given the parameter vector $\boldsymbol{\theta}$ ($\ln(\mathrm{y}|\boldsymbol{\theta})$) from each posterior draw $\boldsymbol{i}$. The second term is a penalty expressed as the difference between two alternative realizations of deviance and is analogous to the effective number of parameters. $\mathbb{E}$ denotes the expectation of the quantity within the parentheses which is taken as the mean (for all other criteria presented here as well). To compute DIC, we use the conditional likelihood (i.e. random or process error parameters, such as recruitment, are sampled along with parameters during MCMC) and exclude all prior densities. The best model minimizes DIC and the difference in DIC between each model and the best model (ΔDIC) can be interpreted in the same way as ΔAIC (ΔDIC <2 deserves serious consideration, 2<ΔDIC<7 suggests potential support, ΔDIC>7 is unlikely).

WAIC (Watanabe, 2013) incorporates the posterior distributions of the observations and sums across K = 6 data sets fit within BASA as follows:

$$\mathrm{WAIC} = -2\sum_{\mathrm{k}=1}^{K}\sum_{\mathrm{t}=1}^{N_K}\ln\left(\mathbb{E}\left[\mathrm{y}_{\mathrm{t,k}}|\boldsymbol{\theta}\right]\right) + 2\sum_{\mathrm{k}=1}^{K}\sum_{\mathrm{t}=1}^{N_K}\mathrm{Var}\left(\ln\left[\mathrm{y}_{\mathrm{t,k}}|\boldsymbol{\theta}\right]\right)$$

in which N_K is the number of observations within each data set. The first term is the ln point-wise predictive score which is the logged expectation of posterior point densities of the data (see Figure 2.1 for calculation steps). The second term is a correction factor for the biased estimate of the true posterior predictive score and follows the form recommended by Gelman et al. (2014b) the sum of the variances of the ln posterior densities for each data point. The best performing model minimizes WAIC.

PPL (Gelfand & Ghosh, 1998) is the sum of a loss and risk function using the posterior predictive distributions $\tilde{\mathrm{y}}_{\mathrm{t,k}}|\hat{\mathrm{y}}_{\mathrm{t,k}}$ of the data:

$$\mathrm{PPL} = \sum_{\mathrm{k}=1}^{K}\sum_{\mathrm{t}=1}^{N_K}\frac{\left(\mathrm{y}_{\mathrm{t,k}} - \mathbb{E}\left(\tilde{\mathrm{y}}_{\mathrm{t,k}}|\hat{\mathrm{y}}_{\mathrm{t,k}}\right)\right)^2}{\bar{\mathrm{y}}_{\mathrm{k}}{}^2} + \sum_{\mathrm{k}=1}^{K}\sum_{\mathrm{t}=1}^{N_K}\frac{\mathrm{Var}\left(\tilde{\mathrm{y}}_{\mathrm{t,k}}|\hat{\mathrm{y}}_{\mathrm{t,k}}\right)}{\bar{\mathrm{y}}_{\mathrm{k}}{}^2}$$

The first term is loss and is the squared error between the observed data and the expectation of the posterior predictive distribution (i.e. the posterior predictive mean). The second term is risk and is the variance of the posterior predictive distributions that acts as a penalty for model complexity (i.e. more parameters induce less precision). As with WAIC, PPL sums over K =6 within BASA to provide a single PPL for each model (see Figure 2.1). To normalize the differences in scale across data sets (e.g. egg deposition versus hydroacoustic survey data), but maintain differences in variance in their observations over time, we define a relative loss and risk function where each term is divided by the squared mean of each observed data set $\bar{\mathrm{y}}_{\mathrm{k}}{}^2$. For the age-composition data, we multiply the observations and posterior predictions of the proportions at each age by the sample size in year t for $\mathrm{y}_{\mathrm{t,k}}$ and $\tilde{\mathrm{y}}_{\mathrm{t,k}}|\mathrm{y}_{\mathrm{t,k}}$ respectively, and then use the mean of the annual sample size for $\bar{\mathrm{y}}_{\mathrm{k}}$. The best performing models should minimize both the loss and risk functions across the sum of all data sets, with the minimum PPL leading to the selected best model.

The equation for PSIS-LOO (Vehtari et al., 2017) is the sum of the expected negative log pointwise predictive densities for all observations ($\widehat{\mathrm{elpd}}_{\mathrm{psis-loo}}$, which is the notation used by Vehtari, Gelman, & Gabry, 2017):

$$\text{PSIS} - \text{LOO} = \widehat{\text{elpd}}_{\text{psis}-\text{loo}} = -\sum_{k=1}^{K}\sum_{t=1}^{N_K} \log\left(\frac{\sum_{i=1}^{N_{MCMC}} w_{t,k}^{i}\left[y_{t,k}|\boldsymbol{\theta_i}\right]}{\sum_{i=1}^{N_{MCMC}} w_{t,k}^{i}}\right)$$

where $w_{t,k}^{i}$ are weights calculated for each individual observation. These weights are derived from a smoothing procedure that uses fits of the generalized Pareto distribution applied to the upper tail (>80th percentile) of raw importance ratios that are calculated from the posterior densities. This procedure is detailed in Fig. 2.1. Generally, when the difference in $\widehat{\text{elpd}}_{\text{psis}-\text{loo}}$ between models is less than four, than predictive performance is similar; however, performances can be similar even when differences are greater than four if the standard error of $\widehat{\text{elpd}}_{\text{psis}-\text{loo}}$ is large. Estimates of $\widehat{\text{elpd}}_{\text{psis}-\text{loo}}$ and their standard errors are computed for each model using the *loo* function from the 'loo' package in R (Vehtari et al., 2020).

Several outputs from *loo* should be considered in the interpretation of PSIS-LOO. First, shape parameters of the generalized Pareto distributions, $\hat{k}$, are estimated for each posterior predictive distribution (each observation) within the model; in other words, there are $K \cdot N_K$ estimates. For $\text{PSIS} - \text{LOO}$ to have moderate to high accuracy, all (or at least most) points should have $\hat{k} < 0.7$. Many instances of $\hat{k} > 0.7$ may indicate poorly fit outliers, model misspecificationo generally flexible models. Another diagnostic is the estimated effective number of parameters (P_{eff}) which can be compared to the actual number of estimated parameters (P) to indicate the predictive ability of an individual model; P_{eff} should approximate or be slightly less than P, if not then this too may indicate model misspecification, weak priors, or overdispersed data (e.g. when $P_{eff} \ll P$, but especially $P_{eff} > P$). For further details on interpreting PSIS-LOO and cross-validation more generally, we recommend readers refer to Vehtari et al. (2017), the 'loo' R package documentation (Vehtari et al., 2020), and their accompanying github page (https://avehtari.github.io/modelselection/).

Appendix C

Derivation of Immune proportions equations in the Estimation model

The proportion of each age within the population that is immune to infection in year y is $\tilde{\Omega}_{y,a}$, so the numbers immune at each age are $N_{immune,y,a} = \tilde{\Omega}_{y,a} N_{y,a}$, To project the numbers immune in the next year $y+1$ ($N_{immune,y+1,a+1}$), we must account for background survival, the proportion immune in the current year $\tilde{\Omega}_{y,a}$, and the proportion that are susceptible and vulnerable, but became infected and recover ($v_a E_{y,a} \omega_y \lambda$):

$$N_{immune,y+1,a+1} = \exp[-M_y](\tilde{\Omega}_{y,a} + v_a E_{y,a} \omega_y \lambda) N_{y,a}$$

The new immune proportions, $\widetilde{\Omega}_{y+1,a+1}$, are then:

$$\tilde{\Omega}_{y+1,a+1} = \frac{\exp[-M_y](\tilde{\Omega}_{y,a} + v_a E_{y,a} \omega_y \lambda) N_{y,a}}{N_{y+1,a+1}}$$

Where $N_{y+1,a+1}$ is projected from Eq. 26:

$$N_{y+1,a+1} = \exp[-M_y][1 - v_a^{trans} E_{y,a} \omega_y (1-\lambda)] N_{y,a}$$

Substituting $N_{y+1,a+1}$ back into the equation for $\tilde{\Omega}_{y+1,a+1}$ above gives:

$$\tilde{\Omega}_{y+1,a+1} = \frac{\exp[-M_y](\tilde{\Omega}_{y,a} + v_a E_{y,a} \omega_y \lambda) N_{y,a}}{\exp[-M_y][1 - v_a E_{y,a} \omega_y (1-\lambda)] N_{y,a}}$$

Where s_a and $N_{y,a}$ cancel out to provide Eq. 27:

$$\tilde{\Omega}_{y+1,a+1} = \frac{\tilde{\Omega}_{y,a} + v_a E_{y,a} \omega_y \lambda}{1 - v_a E_{y,a} \omega_y (1-\lambda)}$$

For the plus age group, both numbers from one year younger than the starting age of the plus group as well as the plus group in the current year are projected into the next year following:

$$N_{y+1,n_a} = \exp[-M_y][1 - v_{n_a-1} E_{y,n_a-1} \omega_y (1-\lambda)] N_{y,n_a-1} + \exp[-M_y][1 - v_{n_a} E_{y,n_a} \omega_y (1-\lambda)] N_{y,n_a}$$

Calculating the immune numbers in the plus-group is a similar equation, but accounts for current immune proportions:

$$N_{immune,y+1,n_a} = \exp[-M_y](\widetilde{\Omega}_{y,n_a-1} + v_{n_a-1} E_{y,n_a-1} \omega_y \lambda) N_{y,n_a-1} + \exp[-M_y](\widetilde{\Omega}_{y,n_a} + v_{n_a} E_{y,n_a} \omega_y \lambda) N_{y,n_a}$$

The next year's immune proportion of the plus group is then:

$$\widetilde{\Omega}_{y+1,n_a} = \frac{\exp[-M_y](\widetilde{\Omega}_{y,n_a-1} + v_{n_a-1} E_{y,n_a-1} \omega_y \lambda) N_{y,n_a-1} + \exp[-M_y](\widetilde{\Omega}_{y,n_a} + v_{n_a} E_{y,n_a} \omega_y \lambda) N_{y,n_a}}{N_{y+1,n_a}}$$

Table C1: Descriptions of indices, derived quantities, and parameters used in the operating and estimation models. Values are grouped by their usage in both the operating and estimation models, in only the operating model, or only the estimation model. Default values (where specified by used) and deriving equations are provided. Estimation model parameters are indicated as fixed or estimated (Est). Default values (or equations in a couple instances) provided for estimation model parameters are the true values specified by the user or computed from variables in the operating model.

Notation	Description	Type	Default values/Deriving equations
OM/EM shared			
d	Day	Index	--
a	Age	Index	0-7
y	Year	Index	--
j	Age transmitting infection	Index	0-7
p	Index for combined positive and negative antibody prevalence by age	Index	1-14
d_{end}	Index of last day in season of transmission	Index	120
n_a	Number of ages in age-structure	Index	8
n_y	Number of years fit in estimation model	Index	50
$N_{y,a}$	Number of fish of age a during year y	Derived quantity	OM: $\sum\left(T_{y,d=1,a}, I_{y,d=1,a}, C_{y,d=1,a}\right)$ EM: Eq. 30-31
B_y	Survey (spawning) biomass of fish in year y	Derived quantity	Eq. 20
v_a	Proportion of fish mixing with reservoir population as a function of age a	Derived quantity	Eq. 2
s_a	Fishery-independent selectivity as a function of age a	Derived quantity	Eq. 20
$\tilde{\Omega}_{y,a}$	Proportion of age a fish that are immune in year y	Derived quantity	Eq. 23 & 28
$\tilde{\Omega}_y$	Total proportion of population that are immune	Derived quantity	$\sum_{a=1}^{n_a} \tilde{\Omega}_{y,a}$
w_a	Weight-at-age (grams)	Parameter	70, 94, 115, 134, 150, 160, 165, 168
M_y	Annual background natural mortality rate (yr^{-1})	Parameter	0.25
$\bar{R}$	Average age-0 recruitment (millions)	Parameter	181.27 (Est)
ε_y	Annual deviations in recruitment about constant average recruitment	Parameter	--
σ_R	Standard deviation of annual recruitment	Parameter	1.2 (Est)

ρ	Extent of autocorrelation in recruitment deviations	Parameter	0.6
a_{50}^{mix}	Age at which 50% are mixing with reservoir population	Parameter	1 or 3 (Est)
a_{95}^{mix}	Age at which 95% are mixing with reservoir population	Parameter	2 or 4 (Est)
a_{50}^{survey}	Age at which 50% are selected for fishery-independent sampling	Parameter	3 (Est)
a_{95}^{survey}	Age at which 95% are selected for fishery-independent sampling	Parameter	4 (Est)
q	Survey scalar	Parameter	-0.5
σ_B	Standard deviation for log-normal errors on biomass survey indices	Parameter	0.3
OM specific			
$\mathbf{T}_{d,a}$	Disease stage transition matrix of age a in day d	Derived quantity	Eq. 5
$T_{y,d,a}$	Number of total susceptible individuals of age a in year y on day d	Derived quantity	Eq. 10
$S_{y,d,a}$	Number of reservoir susceptible individuals of age a in year y on day d	Derived quantity	Eq. 1
$I_{y,d,a}$	Number of active infected individuals of age a in year y on day d	Derived quantity	Eq. 11
$C_{y,d,a}$	Number of carrier/immune individuals of age a in year y on day d	Derived quantity	Eq. 12
$\boldsymbol{n}(d)$	Vector of all numbers of age a fish in each disease stage in year y on day d	Derived quantity	Eq. 3
$\beta_{I,d,a}$	Probability of becoming infected due to contacts with infected individuals	Derived quantity	Eq. 6
$\beta_{C,d,a}$	Probability of becoming infected due to contacts with carrier individuals	Derived quantity	Eq. 7
$\Theta_{y,a}$	Simulated age-composition data in fishery-independent sample in year y	Derived quantity	Eq. 22
$A_{y,a}^{+}$	Simulated positive antibody numbers at age a in year y	Derived quantity	Eq. 24
$A_{y,a}^{-}$	Simulated negative antibody numbers at age a in year y	Derived quantity	Eq. 25
$\boldsymbol{A}_{\boldsymbol{y},\boldsymbol{p}}$	Combined simulated numbers with positive and negative antibodies across ages	Derived quantity	Eq. 26
$\Psi_{y,d}$	Infection prevalence within total population on day d of year y	Derived quantity	Eq. 33
$\bar{\Psi}_y$	Mean infection prevalence from simulated samples of true infection prevalence in year y	Derived quantity	$\frac{\sum_1^3 \Psi_{y,d}}{3}$
M_d	Daily background natural mortality rate (d^{-1})	Parameter	6.85×10^{4}
α	Daily probability of death due to disease given infection (based on instantaneous rate 11 yr^{-1} under ambient temperatures; unpublished data provided by P. Hershberger)	Parameter	0.0297

γ	Daily probability of recovery from disease given infection (based on instantaneous rate 26 yr^{-1} under ambient temperatures; unpublished data provided by P. Hershberger)	Parameter	0.0688
μ_I	Mean transmission rate of disease from infected individuals to susceptible (d^{-1})	Parameter	0.01
μ_C	Mean transmission rate of disease from carrier individuals to susceptible (d^{-1})	Parameter	0.000001
α_y	Time-varying probability of death due to disease (based on rates 9 - 21 yr^{-1})	Parameter	0.025 – 0.056
γ_y	Time-varying probability of recovery from disease (based on rates 20 - 70 yr^{-1})	Parameter	0.053 – 0.175
δ_y	Multiplicative lognormal observation error	Parameter	Random
$Z_{\Theta,y}$	Sample size for fishery-independent age-composition	Parameter	20 or 200
$Z_{A,y}$	Sample size for antibody prevalence survey	Parameter	20 or 200
EM specific			
$\hat{B}_y$	Predicted biomass in fishery-independent survey	Derived quantity	Eq. 19
$\hat{\Theta}_{y,a}$	Predicted age-composition in fishery-independent sample in year y	Derived quantity	Eq. 22
$\hat{A}_{+,y,a}$	Predicted numbers of fish with antibodies at age a in year y	Derived quantity	Eq. 24
$\hat{A}_{-,y,a}$	Predicted numbers of fish without antibodies at age a in year y	Derived quantity	Eq. 25
$\hat{A}_{y,p}$	Predicted proportions of combined positive and negative antibodies across ages	Derived quantity	Eq. 26
$E_{y,a}$	Proportion of age a fish that are susceptible/exposed in year y	Derived quantity	Eq. 29
$\overline{N}$	Initial average numbers across ages	Parameter	Est
$\tilde{\Omega}_a^*$	Initial proportions of immune fish in each age	Parameter	0 (Fixed)
ω_y	Infection rate of susceptible fish in year y	Parameter	$\frac{\sum_{d=1}^{n_{day}}(\beta_{I,d,a}+\beta_{C,d,a})v_a S_{y,d,a}}{v_a S_{y,d=1,a}}$ (Est)
λ	Annual recovery rate of infected fish	Parameter	$\gamma/(\alpha+\gamma)$ (Est)
κ	Scaling coefficient for the effect of annual infection prevalence	Parameter	Est
$k_{y,a}$	Observed number of seropositive fish in samples of age a and in year y	Data	--

$\hat{p}_{y,a}$	Predicted proportion of seropositive fish in age a and year y	Derived quantity	$\frac{\hat{A}^{+}_{y,a}}{\hat{A}^{+}_{y,a} + \hat{A}^{-}_{y,a}}$
$N^{sero}_{y,a}$	Total number of fish sampled for seroprevalence in age a and year y	Data	--

Table C2: Equations for the plus-age group in the operating and estimation models. Table 3.1 provides variable definitions.

Equations	Description
Operating model	
$T_{y+1,d=1,,n_a} = \exp\left[-M_d \cdot 365\right]\left[\left(1-v_{n_a-1}\right)T_{y,d=1,n_a-1} + \left(1-v_{n_a}\right)T_{y,d=1,n_a}\right] + \exp\left[-M_d \cdot \Delta t\right]\left[v_{n_a-1}T_{y,n_{day},n_a-1} + v_{n_a}T_{y,n_{day},n_a}\right]$	Total susceptible numbers in the plus-group at the start of the next season
$S_{y+1,d=1,,n_a} = v_{n_a}T_{y+1,d=1,,n_a}$	Reservoir susceptible numbers in the plus-group at the start of the next season
$I_{y+1,d=1,,n_a} = 0$	Infected numbers in the plus-group at the start of the next season
$C_{y+1,d=1,,n_a} = \exp\left[-M_d \cdot \Delta t\right]\left(C_{y,n_{day},n_a-1} + C_{y,n_{day},n_a}\right)$	Carrier numbers in the plus-group at the start of the next season
Estimation model	
$\mathrm{N}_{y+1,n_a} = \exp\left[-M_y\right]\left(\left(1-v_{n_a-1}E_{y,n_a-1}\omega_y(1-\lambda)\right)\mathrm{N}_{y,n_a-1} + \left(1-v_{n_a}E_{y,n_a}\omega_y(1-\lambda)\right)\mathrm{N}_{y,n_a}\right)$	Numbers in plus-group
$\widetilde{\Omega}_{y+1,n_a} = \frac{\exp\left[-M_y\right]\left(\widetilde{\Omega}_{y,n_a-1} + v_{n_a-1}E_{y,n_a-1}\omega_y\lambda\right)\mathrm{N}_{y,n_a-1} + \exp\left[-M_y\right]\left(\widetilde{\Omega}_{y,n_a} + v_{n_a}E_{y,n_a}\omega_y\lambda\right)\mathrm{N}_{y,n_a}}{\mathrm{N}_{y+1,n_a}}$	Proportion of immune in plus-group