




Original Article

Applying Bayesian model selection to determine ecological covariates for recruitment and natural mortality in stock assessment

John T. Trochta ^{1,*} and Trevor A. Branch²

¹Institute of Marine Research, P.O. Box 1870, Nordnes, 5817 Bergen, Norway

²School of Aquatic and Fishery Sciences, University of Washington, P.O. Box 355020, Seattle, WA 98195, USA

*Corresponding author: email: john.tyler.trochta@hi.no

Trochta, J. T., and Branch, T. A. Applying Bayesian model selection to determine ecological covariates for recruitment and natural mortality in stock assessment. – ICES Journal of Marine Science, 78: 2875–2894.

Received 3 March 2021; revised 4 August 2021; accepted 9 August 2021; advance access publication 9 September 2021.

Incorporating ecological covariates into fishery stock assessments may improve estimates, but most 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 used 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 were sensitive to prior specification and may provide potentially very different results. We recommend using multiple criteria and exploring different statistical assumptions about covariates for their use in stock assessment.

Keywords: Bayesian model selection, covariates, natural mortality, recruitment, stock assessment

Introduction

Population dynamics models, such as those used in fisheries management, are governed by biological parameters including growth, recruitment, and natural mortality (Hilborn and Walters, 1992). Explaining the variability in recruitment and natural mortality is perhaps the most challenging obstacle to conducting accurate fisheries stock assessments. Recruitment predictions that rely on a relationship with parental biomass are a key source of uncertainty in stock assessment (e.g. Needle, 2001; Maunder and Deriso, 2003; Maunder and Watters, 2003), in large part because of the high variance around estimated stock–recruitment relationships for many fish stocks (Gilbert, 1997; Lee *et al.*, 2012; Szuwalski *et al.*, 2015).

Natural mortality of young and old fish is also a key uncertainty (Vetter, 1988; Clark, 1999), proving difficult to estimate and causing biased estimates of stock status when mis-specified, especially when ignoring time-varying mortality (Deroba and 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 (Rice and Browman, 2014; Pepin, 2015), but efforts continue because improving the accuracy and precision of stock assessments could 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 (Maunder and Watters, 2003; Deriso *et al.*, 2008). 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 inferences made from 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 and 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 and Deroba, 2015). Additionally, the covariate itself may imperfectly represent the true forcing ecological factor and act in concert or interact with other unmodelled or unobserved factors. This unexplained variability should be treated as random effects (Maunder and Watters, 2003; Deriso *et al.*, 2008), with state-space formulations (Maunder *et al.*, 2015; Miller *et al.*, 2016), or by modeling covariates “as data” (Schirripa *et al.*, 2009; Crone *et al.*, 2019); 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).

Time series for covariates are usually observed or estimated outside of surveys conducted for single-species stock assessments, and thus have nonoverlapping time frames or missing years. Approaches to address missing data for covariates have been proposed and explored, 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 and Deriso, 2010). A state-space framework is the preferred approach (Maunder and Thorson, 2019), but simpler alternatives such as substituting the mean of the covariate data may also perform well under some circumstances (Maunder and Deriso, 2010).

Herring display large fluctuations in abundance (Hjort, 1914) as well as prolonged periods of low adult abundance and recruitment, even for decades (Trochta *et al.*, 2020). Consequently, including ecological covariates in herring (genus *Clupea*) stock assessments has long been a focus (Deriso *et al.*, 2008; Deroba *et al.*, 2018; Hultson *et al.*, 2018; Okamoto *et al.*, 2020). Pacific herring in Prince William Sound, Alaska offer an ideal case study, having 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 *et al.*, 2014). Various studies have investigated biological and ecological factors that may inhibit the recovery of Prince William Sound herring, each providing different answers (Williams and Quinn, 2000; Brown and Norcross, 2001; Deriso *et al.*, 2008; Pearson *et al.*, 2012; Sewall *et al.*, 2017; Ward *et al.*, 2017). Thus, there is a continued need to better understand the factors driving herring productivity in Prince William Sound.

Currently, a single-species Bayesian age-structured stock assessment model is used to estimate the stock status of Prince William Sound herring (Muradian *et al.*, 2017). A variety of model-selection

methods is available for Bayesian models on evaluating support for individual covariates, each with its benefits and limitations (Hooten and 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 and 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 (Piiironen and Vehtari, 2017; Vehtari *et al.*, 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; Piiironen and Vehtari, 2017). However, any one criterion is vulnerable to selecting the incorrect model, especially when models are overfitted or misspecified (Hooten and Hobbs, 2015; Piiironen and 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.

Material and methods

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.

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 (Williams and Quinn, 2000; Brown and Norcross, 2001; Deriso *et al.*, 2008; Pearson *et al.*, 2012; Sewall *et al.*, 2017; Ward *et al.*, 2017). 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 1.

- (i) *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; Quinn *et al.*, 2001; Marty *et al.*, 2003; Marty *et al.*, 2010). 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.
- (ii) *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); Bakun (1975) upwelling indices ($\text{m}^3 \text{s}^{-1} 100 \text{m}^{-1}$) 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>).
- (iii) *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.
- (iv) *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 (Polovina *et al.*, 1996; Mantua *et al.*, 1997; Mantua and Hare, 2002). 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.
- (v) *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).
- (vi) *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 *et al.*, 2017). We used stock assessment estimates of arrowtooth flounder total biomass (ages 1+) from the Gulf of Alaska (Spies *et al.*, 2017).
- (vii) *Gulf of Alaska walleye pollock spawning biomass (SB) and age-1 numbers (lagged 1-yr)*. While walleye pollock eat herring within Prince William Sound (Thorne, 2008; Gray *et al.*, 2019), 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 (Dorn *et al.*, 2017; Oken *et al.*, 2018; Barnes *et al.*, 2020) and Steller sea lions (Trites and 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.
- (viii) *Humpback whales*. Humpback whales are major predators of herring throughout the northeast Pacific and in Prince William Sound (Straley *et al.*, 2017; Moran *et al.*, 2018). 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 standardized sighting surveys and opportunistic efforts within Prince William Sound during the fall and winter (Moran and Straley, 2019).
- (ix) *Freshwater discharge*. Freshwater discharge into Prince William Sound impacts quality of nearshore nursery habitats for juvenile herring, changing zooplankton prey timing and quantity (Ware and Thomson, 2005) and altering salinity, which in turn cues changes in larval and juvenile fish behavior (Boehlert and 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).
- (x) *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

Table 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 by half-years 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 <i>et al.</i> (2003); Hulson <i>et al.</i> (2007)	Muradian <i>et al.</i> (2017); recent data from Hershberger (2007–2019)
Cause of endemic disease in herring that positively associates with mortality in older fish	<i>Ichthyophonus hoferi</i>	No	Yes, 3–8	1–2	1994–2006, 2007–2017	1994–2017, 2007–2017	Marty <i>et al.</i> (2003); Hulson <i>et al.</i> (2007)	Muradian <i>et al.</i> (2017); recent data from Hershberger (2007–2019)
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 <i>et al.</i> (2017)	NOAA Pacific Fisheries Environmental Laboratory (https://oceanwatch.pfeg.noaa.gov/products/PFELData/upwell/monthly/upindex.mon) http://www.o3d.org/npgo; Di Lorenzo <i>et al.</i> (2008)
Broad-scale summer climate associates (positively or negatively) with adult mortality	Summer NPGO	Yes	Yes, 3–9+	1	1980–2017	1980–2017, 1994–2017, 2007–2017, 1980–2009	NA	http://www.o3d.org/npgo; Di Lorenzo <i>et al.</i> (2008)
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 <i>et al.</i> (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 <i>et al.</i> (2008)	http://research.jisao.washington.edu/pdo/

Table 1. Continued

Hypothesis	Indicator	Used for recruitment?	Used for mortality and for which ages?	Index for half-year mortality (<i>b</i>)	Years of available data	Timeframes modeled	References for hypothesis	Source
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 <i>et al.</i> (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 <i>et al.</i> (2008); Pearson <i>et al.</i> (2012); Sewall <i>et al.</i> (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 <i>et al.</i> (2017)	Spies <i>et al.</i> (2017)
Pollock are alternative prey source for herring predators (Stellar sea lion and arrowtooth flounder) and negatively associate with mortality	Gulf of Alaska walleye pollock spawning biomass	No	Yes, 3–9+	1–2	1980–2017	1980–2017, 1994–2017, 2007–2017, 1980–2009	Pearson <i>et al.</i> (2012)	Dorn <i>et al.</i> (2017)
Humpbacks prey on herring and positively associate with mortality	Humpback whale abundance	No	Yes, 3–9+	1	1980, 1983–2009	1980–2009	Pearson <i>et al.</i> (2012); Moran <i>et al.</i> (2018)	Teerlink <i>et al.</i> (2015)
Humpbacks prey on herring and positively associate with mortality	Humpback whale counts	No	Yes, 3–9+	2	2006–2008, 2011–2017	2007–2017	Pearson <i>et al.</i> (2012); Moran <i>et al.</i> (2018)	Figure 5 from Moran and Straley (2019); raw data from Straley and Moran (2006–2020)
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	Ware and Thomson (2005); Ward <i>et al.</i> (2017)	Royer (1982)

Table 1. Continued

Hypothesis	Indicator	Used for recruitment?	Used for mortality and for which ages?	Index for half-year mortality (<i>b</i>)	Years of available data	Timeframes modeled	References for hypothesis	Source
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,	NA	Haught and Moffitt (2012-2019)
Pollock recruitment success correlates with herring recruitment	Gulf of Alaska walleye pollock age-1 (lagged 1-yr)	Yes	No	NA	1980–2015	1980–2009, 1980–2015, 1994–2015, 2007–2015,	Sewall <i>et al.</i> (2017)	Dorn <i>et al.</i> (2017)
Juvenile salmon have antagonistic interaction with herring and negatively associate with recruitment	Prince William Sound hatchery juvenile pink salmon	Yes	No	NA	1980–2017	1980–2009, 1980–2017, 1994–2017, 2007–2017	Deriso <i>et al.</i> (2008); Pearson <i>et al.</i> (2012); Sewall <i>et al.</i> (2017)	Rich Brenner (ADF&G)
Negative shift in mean recruitment, regardless of cause	1989 regime shift	Yes	No	NA	1980–2017	1980–2017, 1994–2017, 2007–2017	Ward <i>et al.</i> (2017)	NA

prey abundance and warmer summer temperatures (Batten *et al.*, 2016). We included a time series of scale increments from archived herring scale images collected from Prince William Sound (Haught and Moffitt, 2018).

- (xi) *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 and 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.
- (xii) *Null model.* The null model includes no covariates on natural mortality and recruitment.

Bayesian age-structured assessment (BASA) model

Each ecological covariate was incorporated into BASA (Muradian *et al.*, 2017), which is an updated version of the ADF&G assessment model used in previous modeling studies (Deriso *et al.*, 2007; Hulson *et al.*, 2007; Deriso *et al.*, 2008). A total of 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 (Muradian *et al.*, 2017), we provide a brief description and tables of the data types, model equations, parameters, and likelihood equations in the Supplementary Material. 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 further described in the Supplementary Material.

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 $\varepsilon_{Rec, y}$ is the estimated unexplained error in recruitment variation with log-normal bias-correction and \bar{R} is mean stationary recruitment across time:

$$R_y = \bar{R} e^{\beta I_y + \varepsilon_{Rec, y} - 0.5\sigma_{Rec}^2},$$

$$\varepsilon_y \sim N(0, \sigma_{Rec}^2),$$

$$\sigma_{Rec} \sim U(0.0001, 2).$$

There is a uniform prior that constrains σ_{Rec} (recruitment standard deviation) to a positive variance, and differs from BASA (Muradian *et al.*, 2017), which freely estimated annual recruitment. We fixed σ_{Rec} to different values as a sensitivity check on the results (Supplementary Figure S1).

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 ($S_{y, a, b}$) of adult herring

of age a , in year y , and half-year b (1 or 2) is:

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

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

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

in which \bar{M}_{9+} is the instantaneous mortality rate of the plus group in the first year, and was 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 at a time was included in either the recruitment or survival functions within BASA to provide a suite of independent models (Table 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 $\beta = 0$. 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 ($m_{1, 1992-1993}$ and $m_{2, 1992-1993}$) that were added to \bar{M} 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.

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 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 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 and Deriso, 2010).

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 *et al.*, 2017). We used the R package “adnuts” (Monnahan and 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 *et al.*, 2014a) of each parameter across chains and zero divergences. Models typically converged in 30–90 min.

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 1). We directly compared effect probabilities among models because all covariates were normalized and thus estimates for β are on the same scale.

For Bayesian model selection, we used the 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 (Figure 1). Details on how they are applied to the multiple data sets within BASA are provided in the Supplementary Material.

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 *et al.*, 2002) has been widely used with stock assessment (Wilberg and Bence, 2008; Punt *et al.*, 2014; Brooks *et al.*, 2019), 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 and Hobbs, 2015). WAIC is preferred to DIC because it integrates over the posterior densities (Figure 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 (Piiroinen and Vehtari, 2017; Vehtari *et al.*, 2017). PPL (Ibrahim and Laud, 1994, 1995; Gelfand and Ghosh, 1998) 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 (Piiroinen and 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 estimates (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 an integrated catch-at-age model such as BASA. Further details on PSIS-LOO diagnostics are provided in the Supplementary Material.

We included the null model in model selection. 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 and Yokoi, 2019). However, the computation of marginal likelihoods in MCMC sampling is computationally infeasible for the much more complex BASA model.

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.

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:

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

where $\varepsilon_{y,a,b}$ 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

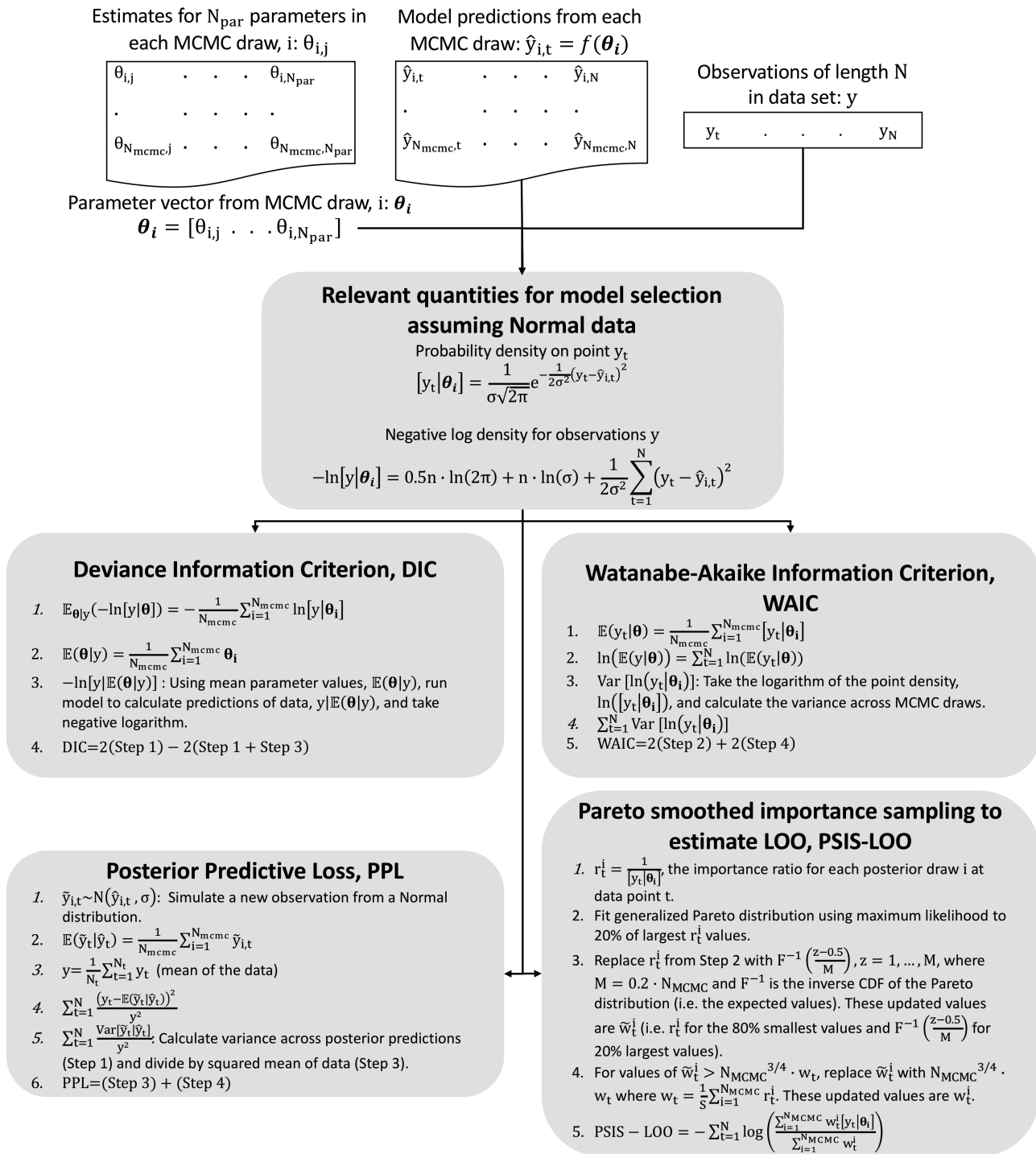


Figure 1. Schematic for how Bayesian model selection criteria were calculated in this analysis using a single example data set y with normally-distributed errors. This example data set has N total observations as indexed by t . Model estimates of the data $\hat{y}_{i,t}$ conditioned on parameter vector θ_i the i th 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 .

error distribution was specified for $\alpha \varepsilon_{y,a,b}$, where α is an estimated nuisance parameter that scales $\varepsilon_{y,a,b}$ to the normalized ecological time series, I_y :

$$\alpha \varepsilon_{y,a,b} \sim N(I_y, \sigma_{I_y,y}^2).$$

We fixed year-specific variance parameters, $\sigma_{I_y,y}^2$, to estimates of annual standard error or deviation values that are available for some time series. Most time series do not have accompanying standard errors. We assumed these had a constant standard deviation of $\sigma_{I_y,y}^2 = 0.3$ in all years. While this is arbitrary, it is a reasonable

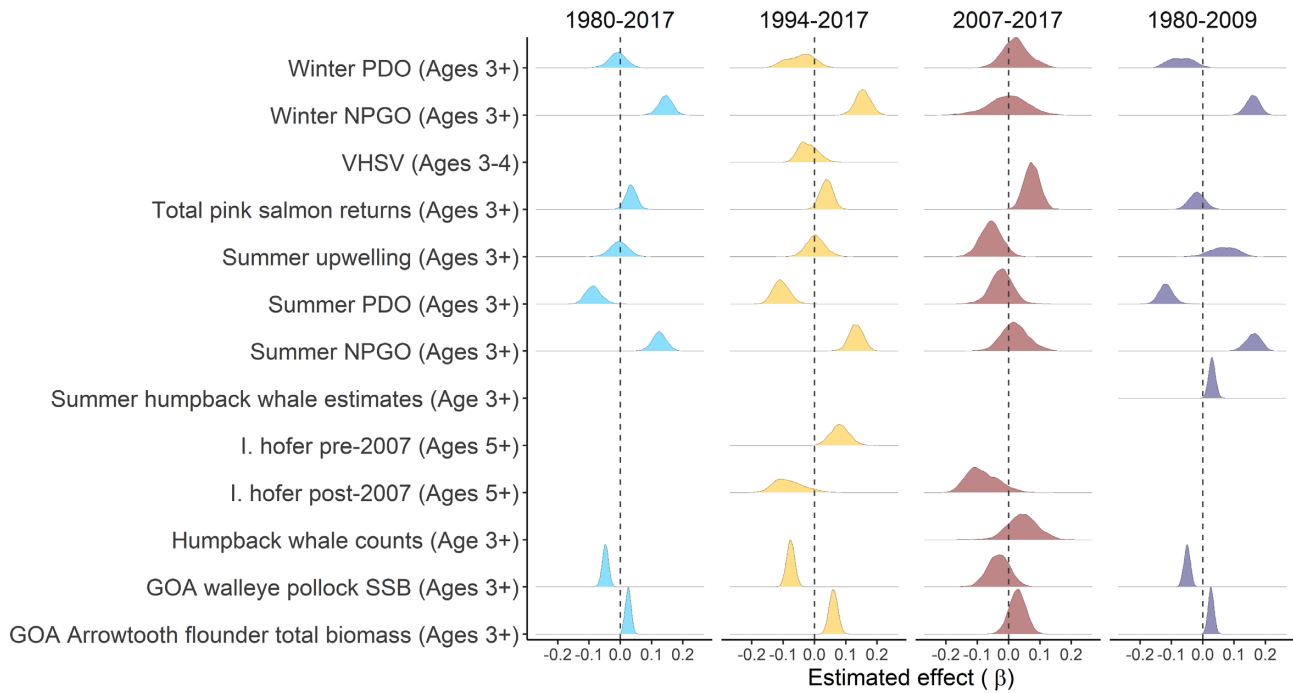


Figure 2. Posterior distributions of the estimated effects on natural mortality and for each time frame. A zero effect is denoted by a dashed vertical line. No posterior is shown for VHSV from 2007 to 2017 because indices were zero in most years except one year (0.0003).

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). We also conducted a sensitivity check by running all latent variable models with larger standard error values ($\sigma_{I_y, y}^2 = 0.7$; Supplementary Material). The resulting prior is:

$$\sum_y \left[\ln(\sigma_{I_y, y}) + \frac{(\alpha \varepsilon_{y, a, b} - I_y)^2}{2\sigma_{I_y, y}^2} \right].$$

Running BASA with uniform priors on the scalar for the annual errors (α) resulted in various models failing to meet convergence criteria. To overcome this issue, we placed informative priors on α using a Normal distribution centered around 0:

$$\alpha \sim N(0, 1^2).$$

We also refit models with a larger standard deviation in the above normal prior ($\sigma = 5$) as a sensitivity check (Supplementary Figure S2). However, we retained a Uniform prior $\alpha \sim U(-10, 10)$ in models with the recruitment covariates as latent variables because these models passed the convergence criteria.

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:

$$R_y = \bar{R} e^{\varepsilon_{Rec, y} - 0.5\sigma_{Rec}^2},$$

$$\sum_y \left[\ln(\sigma_{I_y, y}) + \frac{(\ln(\alpha e^{\varepsilon_{Rec, y}}) - \ln(e^{I_y}))^2}{2\sigma_{I_y, y}^2} \right].$$

We also include in the total likelihood (Supplementary Table S4) the shrinkage distribution for estimating the recruitment devia-

tions, $\varepsilon_{Rec, y}$:

$$38 \ln(\sigma_{Rec}) + \frac{1}{2\sigma_{Rec}^2} \sum_{y \in Y} \varepsilon_{Rec, y}^2.$$

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.

Results

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% of posterior draws in the direction of the hypothesized effect for that covariate, be it positive or negative), 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 (Figure 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 to 2017 (Figure 3). The median proportions of variance explained in $\log(R_y)$ from 1980 to 2017 is substantial for hatchery-released juvenile pink salmon and the 1989 regime shift, both at 0.37 (95%

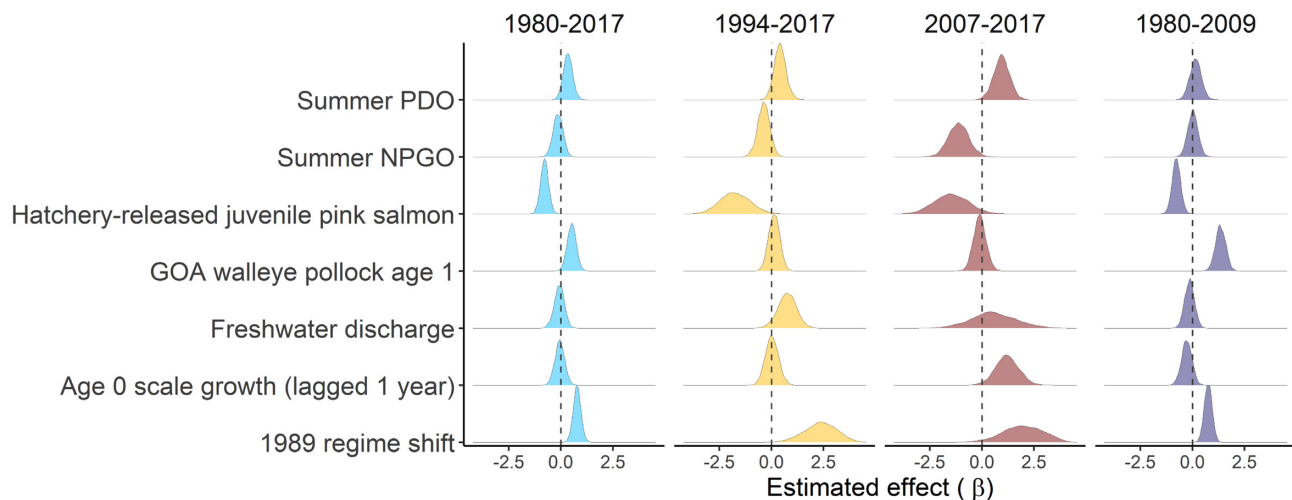


Figure 3. Posterior distributions of the estimated effects on recruitment and for each time frame. A zero effect is denoted by a dashed vertical line.

uncertainty from 0.10 to 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.

Bayesian model selection

For natural mortality effects, model selection most consistently supported the model with total pink salmon returns (Figure 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 (Figure 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 selection 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 (Figure 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 estimates. This result did not change when σ_{Rec} was fixed to a high value (2.0), but when σ_{Rec} was set low (0.3), hatchery-released pink salmon and the 1989 regime shift performed much better than the null model in two of four criteria (Supplementary Figure S1). Still, 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 (Figure 5). Assuming a standard deviation of 0.3 ($\sigma_{I_y, y}$) for the latent variables resulted in humpback whale counts vaulting to the top position in two of four criteria (PSIS-LOO and WAIC). Additionally, the models with disease indices and total pink salmon returns minimized PPL and DIC, respectively. These rankings changed with a higher assumed standard deviation ($\sigma_{I_y, y} = 0.7$) or weaker prior on the scaling parameter in the natural mortality models ($\alpha \sim N(0, 5.0^2)$; Supplementary Figure S2). With a larger $\sigma_{I_y, y}$, the best natural mortality models also included humpback whale counts (PPL) and walleye pollock SSB (whose DIC nearly tied that of total pink salmon returns), as well as age-0 scale growth in the recruitment function (PSIS-LOO and WAIC). With a weaker prior on α for mortality errors, winter PDO was also favored by PSIS-LOO and WAIC, while total pink salmon returns, humpback whale counts and disease still produced lower values amongst the four criteria.

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 (Figure 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.

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 (Figures 7 and 8). Top natural mortality covariates (as fixed effects and as latent variables) tended to produce more pronounced differences in trends

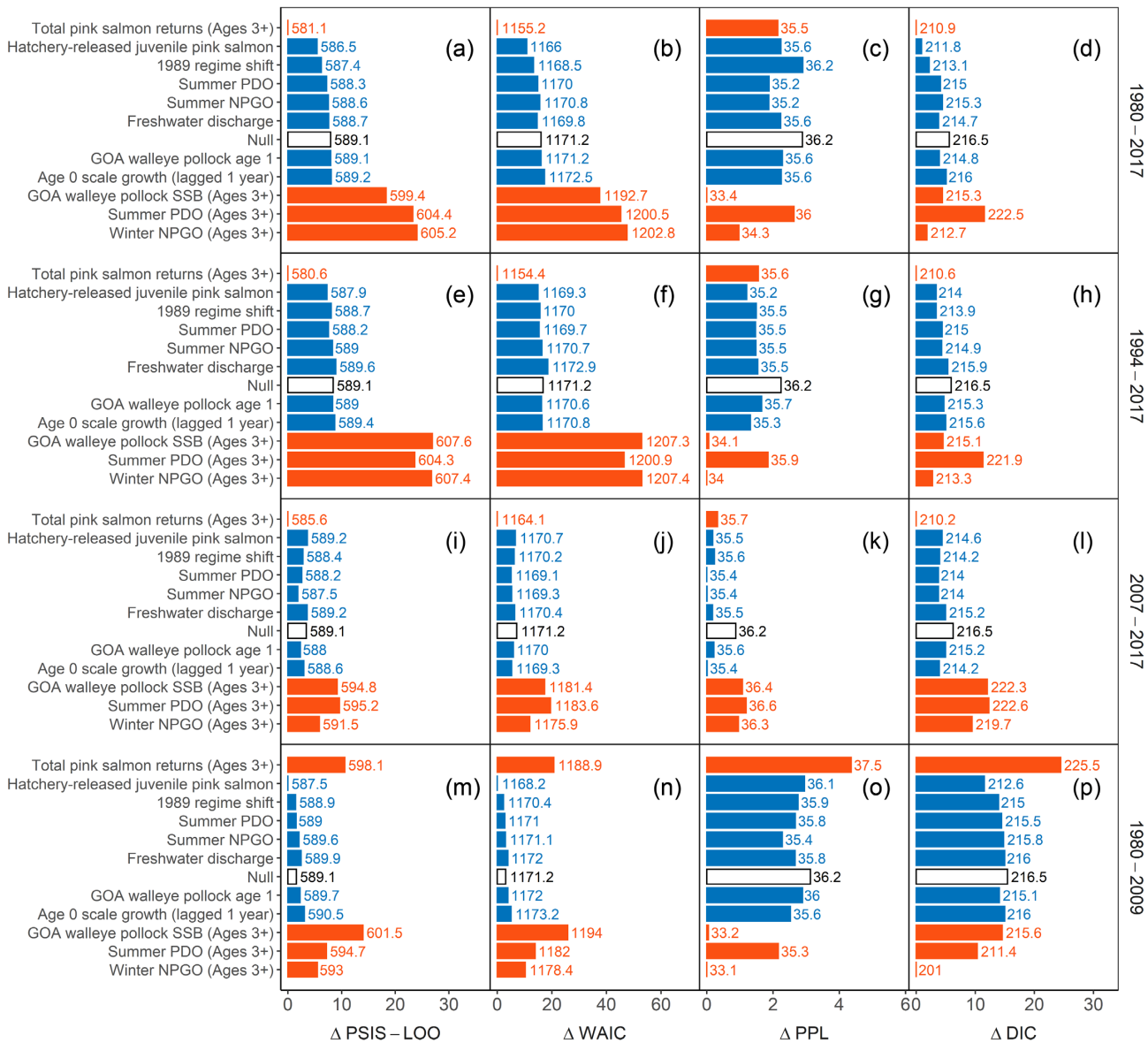


Figure 4. Bar charts of model selection values across select covariates as fixed effects model variants of BASA with at least two criteria better than the value of the Null model (empty black bar). Colors indicate the process affected, either natural mortality (red) or recruitment (blue). Each row represents 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 Δ PSIS – LOO in plot a).

or scale of spawning biomass estimates in recent years. The most consistently supported covariate, total pink salmon returns, estimated different spawning biomass and recruitment levels depending on how the covariate was incorporated; as a fixed effect, estimates differed little from the null model, while as a latent variable, very different trends resulted especially in biomass. Hatchery-released juvenile pink salmon, one of the top covariates affecting recruitment, had no impact on spawning biomass and recruitment estimates (Figure 7); in fact, all recruitment covariates, when implemented as fixed effects, had little impact on recruitment estimates (not shown). However, including age-0 scale growth as a latent variable increased biomass estimates in the second half of the time series while reducing uncertainty of the most recent recruitment esti-

mates (Figure 8). Furthermore, all recruitment covariates as latent variables produced different estimates of spawning biomass and recruitment, as with the natural mortality covariates, but most were unsupported by selection criteria.

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). The impact of pink salmon on population estimates differed by how it was incorporated into BASA. Evidence for other covariates was more ambiguous: many covariates had a high

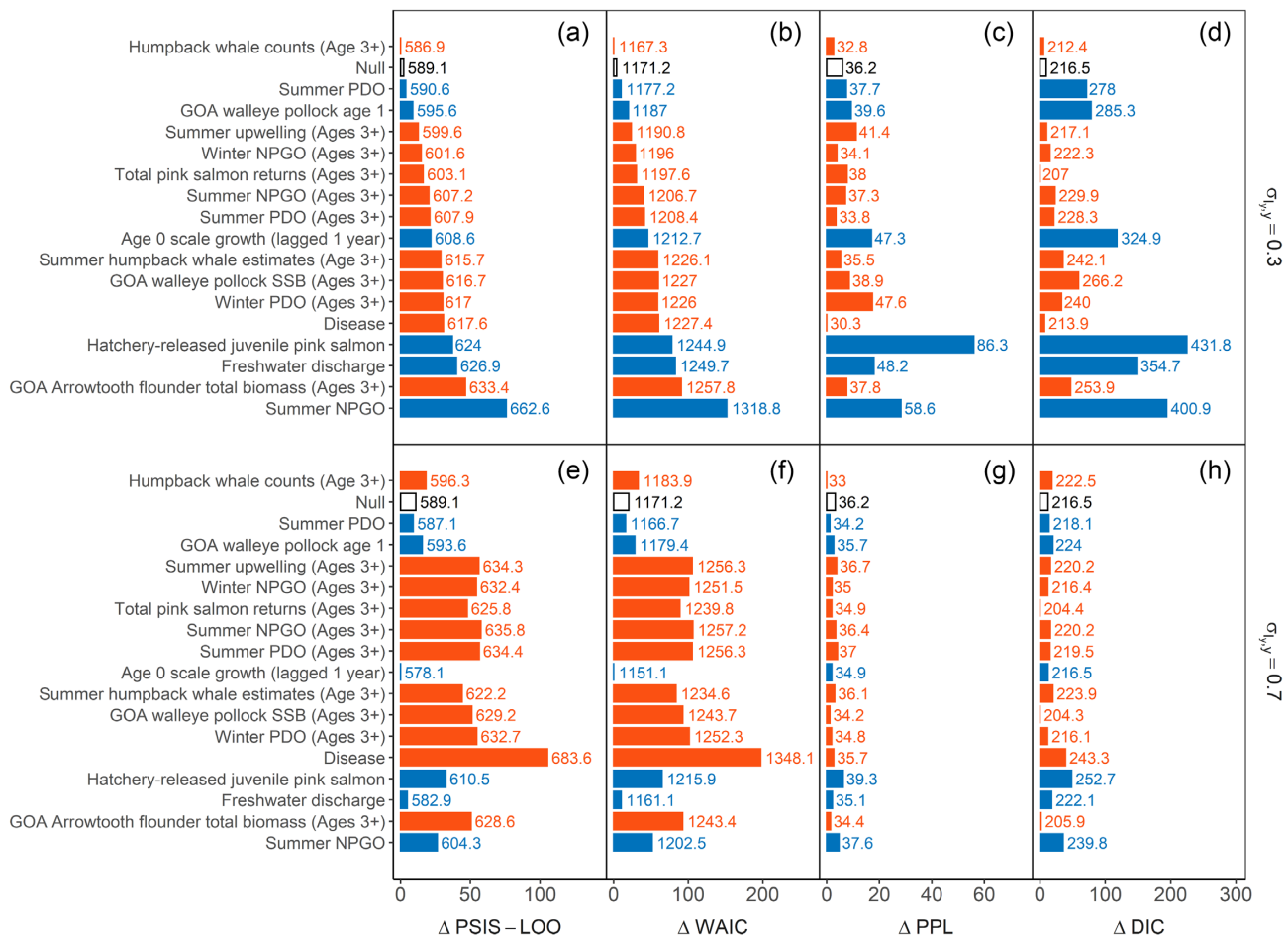


Figure 5. Bar charts of model selection values across all covariates as latent variable model variants of BASA. The format is identical to Figure 2.4 (red = natural mortality effect, blue = recruitment effect) and is only shown for one time frame (1980–2017). Results are presented from two different assumed values for $\sigma_{l,y}$: $\sigma_{l,y} = 0.3$ (a–d) and 0.7 (e–h). The ecological covariates are ordered from the smallest to largest values of Δ PSIS – LOO in plot a).

probability of an effect, fewer had support from model selection in general, and none had support from all criteria or for all time periods. 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 several covariates may partially inform variability in herring population dynamics.

Supported covariates of natural mortality

Our results support an antagonistic interaction between adult herring mortality and Prince William Sound pink salmon. However, the most recent period of pink salmon returns (2007–2017) appeared influential to our results because pink salmon were not selected as a covariate when 1980–2009 was considered. This suggests a risk of spurious correlation, especially considering the negative autocorrelation in even-year and odd-year pink salmon returns due to their two-year life cycle. Furthermore, 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 and Pauly, 1999; Pearson *et al.*, 2012), but there has been recent evidence for irregular localized predation impacts on Prince William Sound herring (Stur-

devant *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 *et al.*, 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 and 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 ambiguous support for a pink salmon effect suggests 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

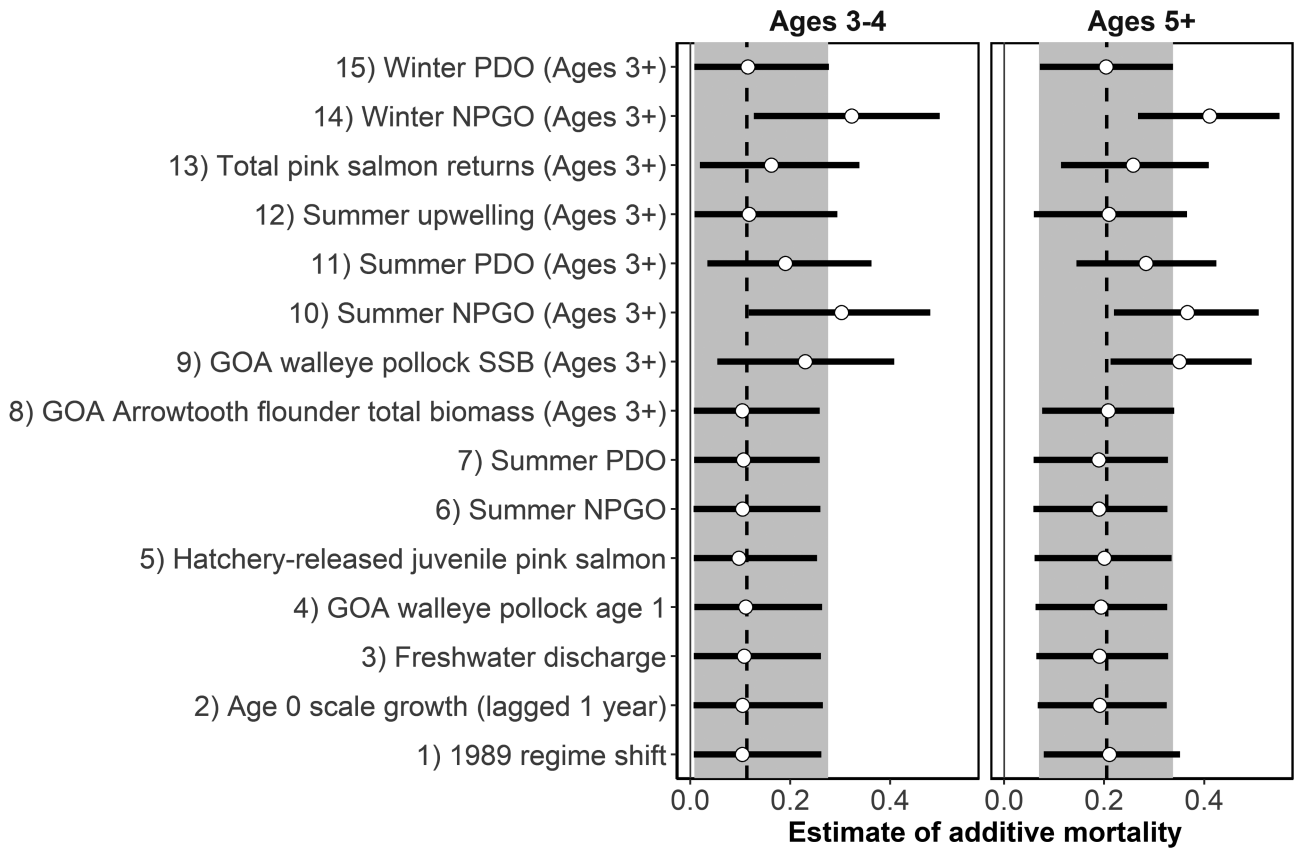


Figure 6. Median (empty circle) and 95% credibility intervals (blue lines) of additional mortality in 1993 for two different age groups (Ages 3–4 and Ages 5+). Recruitment (1–7) and natural mortality (8–15) specific effects are shown together with estimates of the null model denoted by the shaded regions (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).

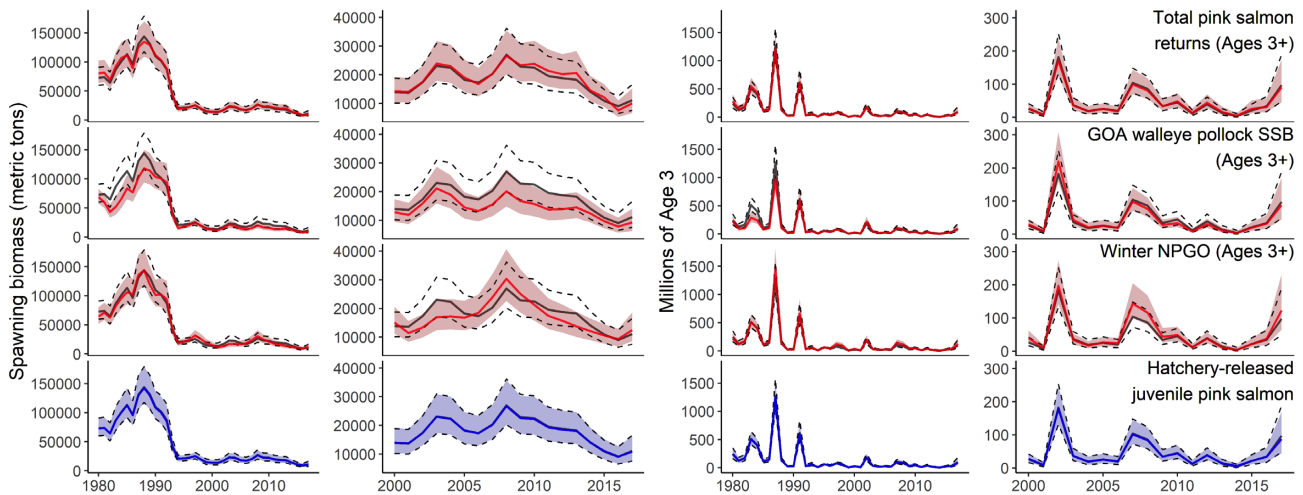


Figure 7. Estimates of spawning biomass and recruitment (in millions of age 3 fish) from select models with covariates as fixed effects from 1980 to 2017 that were the best model in at least one criterion compared to the Null model (dark grey lines). These include effects on natural mortality from total pink salmon returns, GOA walleye pollock SSB, and Winter NPGO, and an effect on recruitment from hatchery-released juvenile pink salmon. 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.

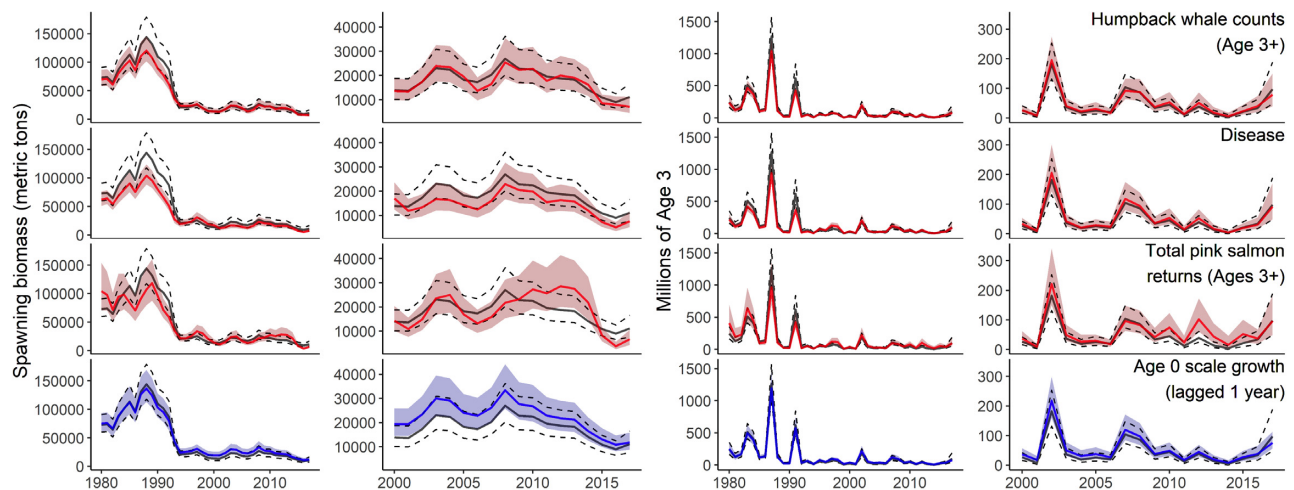


Figure 8. Estimates of spawning biomass and recruitment (in millions of age 3 fish) from select models with covariates as latent variables that were the best model in at least one criterion compared to the Null model (dark grey lines). These include errors in natural mortality informed by winter humpback whale counts, disease, and total pink salmon returns, and errors in recruitment informed by age-0 scale growth. 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. 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.

of Alaska, including Steller sea lions (Trites and Donnelly, 2003; Womble and 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 (Yang, 1993; Spies *et al.*, 2017). Prey switching by predators could occur depending on the relative availability of their prey, as has been implied for Steller sea lions (Trites and 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 *et al.*, 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, 2020; Puerta *et al.*, 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, 2020; Puerta *et al.*, 2019). Following 1988/1989, NPGO explained more climate variance (Di Lorenzo *et al.*, 2010; Yeh *et al.*, 2011) and associated with fish population dynamics such as salmon survival in the North Pacific (Kilduff *et al.*, 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 peri-

ods (Figure 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, 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.

When included as latent variables, some of these same covariates were also selected (total pink salmon returns and walleye pollock) in addition to humpback whales and disease. Humpback whales (summer estimates and overwinter counts) are also likely to increase mortality (Figure 2). Humpback whales are frequently recorded targeting herring aggregations (Pearson *et al.*, 2012; Straley *et al.*, 2017; Moran *et al.*, 2018). Importantly, humpback whale consumption within Prince William Sound in the late 2000s was estimated at 21–77% of herring spawning biomass (Moran *et al.*, 2018). 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.

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, 2003, 2010). 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 (Figure 2). Data on the exposure history to these pathogens, such as from neutralizing antibody tests (Hart *et al.*, 2017), may better allow for a more accurate assessment of the impact of past infections on herring.

The remaining covariates with negligible support, in particular 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 *et al.*, 2002) despite being linked to herring recruitment elsewhere in the northeast Pacific (Williams and Quinn, 2000; Reum *et al.*, 2011). Arrowtooth flounder in the Gulf of Alaska have a diverse diet where herring are a minor prey item compared to other species (Yang, 1993; Spies *et al.*, 2017) and predation on herring is mitigated by the abundance of other prey species as noted earlier.

Ambiguous support for recruitment covariates

We did not find any convincing covariates that consistently explained Prince William Sound herring recruitment. While effects of hatchery-released juvenile pink salmon and a 1989 regime shift had high probabilities across time periods, explained a moderate amount of variance, and were favored in model selection with low assumed recruitment variance (Supplementary Figure S1), 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 mixed support for scale growth and summer NPGO and PDO in 2007–2017 (by posterior probabilities, PPL when included as fixed effects, and PSIS-LOO and WAIC when included as latent variables) 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 and 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).

As fixed effect versus as latent variable models

For our analysis, most covariates as latent variables were sensitive to parameter specification and often exhibited worse model selection values. 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 Schirripa *et al.*, 2009; Crone *et al.*, 2019). Our results suggest estimated recruitment or mortality deviations “fit” to latent variables depend on the assumed value for $\sigma_{I_y, y}$ and the standard deviation of the Normal prior for α . Latent variables act as informative priors that alter the posterior geometry sampled in MCMC to produce different model selection results. With more informative priors ($\sigma_{I_y, y} = 0.3$ or $\alpha \sim N(0, 1.0^2)$), covariates as latent variables led to worse model estimates (i.e. large criteria values) than covariates as fixed effects, especially compared to the null model. When the error distributions were loosened ($\sigma_{I_y, y} = 0.7$ or $\alpha \sim N(0, 5.0^2)$), different covariates provided even lower criteria values compared to the best fixed effect models (e.g. age-0 scale growth and winter PDO). This suggests latent variable errors on mortality or recruitment may improve estimates beyond fixed effects models with the same covariate. A more in-depth analysis is needed to understand more precisely how and why these differences result before determining if these models offer better estimation performance.

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 with 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 estimates.

Mixed messaging in Bayesian model selection

Model selection criteria may fail under certain conditions and with certain models (Vehtari and Ojanen, 2012; Gelman *et al.*, 2014b; Pironen and Vehtari, 2017). 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 and Yokoi, 2019), while in another study, criteria were able to favor models that produce reliable estimates, despite failing to select the true model (Dey *et al.*, 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 (Figures 2 and 3), differences in criteria

were small amongst various models (including the top and null models; Figures 4 and 5), and population estimates differed little between the best models (Figures 7 and 8), 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).

That we found some consensus amongst criteria for at least 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.

Conclusions

Our study demonstrates the continual difficulty in discerning modeling evidence for any single ecological effect despite more data, newer models and analytical techniques, and more robust assumptions. Still, using single covariate models with model selection to determine evidence for certain ecological factors 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.

Data availability statement

The basic code for running BASA and Bayesian model selection is located at https://github.com/johnt23/pwsher_baymodsel. Code for running the entire analysis is available on request. The data underlying this article are available from the various third parties or public repositories listed in Table 1. For those data provided by permission from third parties, data will be shared on request of the corresponding author with permission of these third parties.

Supplementary material

Supplementary material is available at the ICESJMS online version of the manuscript.

Acknowledgements

Funding for this work was provided by the Exxon Valdez Oil Spill 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 thank Rich Brenner and Stormy Haight (ADF&G) for sharing pink salmon data, and John Moran and Jan Straley (NOAA) for sharing whale data for use in this analysis. Data provided by ADF&G are the property of ADF&G, who retains the intellectual rights to these data. Any dissemination of the data must credit ADF&G as the source, with a disclaimer that exonerates the department for errors or deficiencies in reproduction, subsequent analysis, or interpretation.

We also thank Scott Pegau and Cole Monnahan for valuable discussions and feedback throughout the development of this study, and Andre Punt, Tim Essington, Ole Shelton, and two anonymous reviewers for comments that improved the quality and clarity of the manuscript.

Disclaimer

The Exxon Valdez Oil Spill Trustee Council encourages publications but did not have any role in design, data, writing, or the decision to submit this particular article for publication. Furthermore, the findings and conclusions presented by the authors are their own and do not necessarily reflect the views or position of the Trustee Council.

References

- Andrews, A. G., Strasburger, W. W., Jr, Farley, E., V., Murphy, J. M., and 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.
- 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.
- Barnes, C. L., Beaudreau, A. H., Dorn, M. W., Holsman, K. H., and Mueter, F. J. 2020. Development of a predation index to assess trophic stability in the Gulf of Alaska. Ecological Applications, 30:e02141. doi: 10.1002/eap.2141.
- Batten, S., Moffitt, S., Pegau, W., and Campbell, R. 2016. Plankton indices explain interannual variability in Prince William Sound herring first year growth. Fisheries Oceanography, 25: 420–432.
- Bishop, M. A., and 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.
- Boehlert, G. W., and 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.
- Brooks, E. N., and 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., Klibansky, N. *et al.* 2019. Paulik revisited: statistical framework and estimation performance of multistage recruitment functions. Fisheries Research, 217: 58–70.
- Brown, E. D., and 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.
- 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.
- Crone, P. R., Maunder, M. N., Lee, H., and 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.
- Deriso, R. B., Maunder, M. N., and Pearson, W. H. 2008. Incorporating covariates into fisheries stock assessment models with application to Pacific herring. Ecological Applications, 18: 1270–1286.
- Deriso, R. B., Maunder, M. N., and Skalski, J. R. 2007. Variance estimation in integrated assessment models and its importance for hypothesis testing. Canadian Journal of Fisheries and Aquatic Sciences, 64: 187–197.

- Deroba, J. J., Gaichas, S. K., Lee, M.-Y., Feeney, R. G., Boelke, D., and 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.
- Deroba, J. J., and Schueller, A. M. 2013. Performance of stock assessments with misspecified age- and time-varying natural mortality. *Fisheries Research*, 146: 27–40.
- Dey, S., Delampady, M., and Gopalaswamy, A. M. 2019. Bayesian model selection for spatial capture–recapture models. *Ecology and Evolution*, 9: 11569–11583.
- Di Lorenzo, E., Cobb, K. M., Furtado, J. C., Schneider, N., Anderson, B. T., Bracco, A., Alexander, M. A. *et al.* 2010. Central Pacific El Niño and decadal climate change in the North Pacific Ocean. *Nature Geoscience*, 3: 762–765.
- Di Lorenzo, E., Schneider, N., Cobb, K. M., Franks, P., Chhak, K., Miller, A. J., McWilliams, J. C. *et al.* 2008. North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophysical Research Letters*, 35: L08607.
- Dorn, M., Aydin, K., Fissel, B., Jones, D., McCarthy, A., Palsson, W., and Spalinger, K. 2017. Assessment of the walleye pollock stock in the Gulf of Alaska. NOAA.
- Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A. *et al.* 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 27: 233–249.
- Gelfand, A. E., and Ghosh, S. K. 1998. Model choice: a minimum posterior predictive loss approach. *Biometrika*, 85: 1–11.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. 2014a. *Bayesian Data Analysis*. Chapman: Boca Raton, FL.
- Gelman, A., Hwang, J., and Vehtari, A. 2014b. Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24: 997–1016.
- Gilbert, D. 1997. Towards a new recruitment paradigm for fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, 54: 969–977.
- Gray, B. P., Bishop, M. A., and 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.
- Hare, S. R., and Mantua, N. J. 2000. Empirical evidence for North Pacific regime shifts in 1977 and 1989. *Progress in Oceanography*, 47: 103–145.
- Hart, L. M., MacKenzie, A., Purcell, M. K., Powers, R. L., and 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.
- Haught, S., and Moffitt, S. 2012–2019. Scales as growth history records for Pacific herring in Prince William Sound. *In* ADFG Surveys: Aerial Survey Route, Biomass, Age Sex Length, and Spawn. Gulf of Alaska Data Portal, ska Ocean Observing System (AOOS). https://gulf-of-alaska.portal.aos.org/#metadata/35fd35d8-f6f1-4762-9cf0-8e2e970755c4/project/folder_metadata/3115471.
- Hershberger, P. K. 2007–2019. Herring Infection Prevalence Data, EVOS Herring Program, *In* Herring disease program. Gulf of Alaska Data Portal. Alaska Ocean Observing System (AOOS). https://gulf-of-alaska.portal.aos.org/#metadata/61b4ec5a-f15c-4347-b0ba-8a25ad763675/project/folder_metadata/282909.
- Hershberger, P. K., Gregg, J. L., Hart, L. M., Moffitt, S., Brenner, R., Stick, K., Coonradt, E. *et al.* 2016. The parasite *Ichthyophonus* sp. in Pacific herring from the coastal NE Pacific. *Journal of Fish Diseases*, 39: 395–410.
- Hilborn, R., Hively, D. J., Jensen, O. P., and 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., and Walters, C. J. 1992. *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer: Norwell, MA.
- 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.
- Hooten, M. B., and Hobbs, N. 2015. A guide to Bayesian model selection for ecologists. *Ecological monographs*, 85: 3–28.
- Hulson, P.-J. F., Miller, S. E., Dressel, S. C., Quinn, T. J. II, and 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* Impacts of a Changing Environment on the Dynamics of High-latitude Fish and Fisheries. Mueter, F. J., Baker, M. R., Dressel, S. C., and Hollowed, A. B. (Eds.), University of Alaska Fairbanks, Alaska Sea Grant.
- Hulson, P.-J. F., Miller, S. E., Quinn, T. J., Marty, G. D., Moffitt, S. D., and 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.
- Ibrahim, J. G., and Laud, P. W. 1994. A predictive approach to the analysis of designed experiments. *Journal of the American Statistical Association*, 89: 309–319.
- Johnson, K. F., Monnahan, C. C., McGilliard, C. R., Vert-pre, K. A., Anderson, S. C., Cunningham, C. J., Hurtado-Ferro, F. *et al.* 2014. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. *ICES Journal of Marine Science*, 72: 137–150.
- Kaeriyama, M., Nakamura, M., Yamaguchi, M., Ueda, H., Anma, G., Takagi, S., Aydin, K. Y. *et al.* 2000. Feeding ecology of sockeye and pink salmon in the Gulf of Alaska. *North Pacific Anadromous Fish Commission Bulletin*, 2: 55–63.
- Kai, M., and Yokoi, H. 2019. Performance evaluation of information criteria for estimating a shape parameter in a Bayesian state-space biomass dynamics model. *Fisheries Research*, 219: 105326.
- Kemp, I. M., Beauchamp, D. A., Sweeting, R., and 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., Lorenzo, Di, E., Botsford, L., W., and 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.
- Laud, P. W., and Ibrahim, J. G. 1995. Predictive model selection. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57: 247–262.
- Lee, H.-H., Maunder, M. N., Piner, K. R., and Methot, R. D. 2012. Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? *Fisheries Research*, 125–126: 254–261.
- Litzow, M. A., Ciannelli, L., Puerta, P., Wettstein, J. J., Rykaczewski, R. R., and Opiekun, M. 2018. Non-stationary climate–salmon relationships in the Gulf of Alaska. *Proceedings of the Royal Society B: Biological Sciences*, 285: 20181855.
- Litzow, M. A., Ciannelli, L., Puerta, P., Wettstein, J. J., Rykaczewski, R. R., and Opiekun, M. 2019. Nonstationary environmental and community relationships in the North Pacific Ocean. *Ecology*, 100: e02760.
- Litzow, M. A., Hunsicker, M. E., Bond, N. A., Burke, B. J., Cunningham, C. J., Gosselin, J. L., Norton, E. L. *et al.* 2020. The changing physical and ecological meanings of North Pacific Ocean climate indices. *Proceedings of the National Academy of Sciences*, 117: 7665–7671.
- Mantua, N. J., and 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., and Francis, R. C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society*, 78: 1069–1079.
- Marty, G. D., Freiberg, E. F., Meyers, T. R., Wilcock, J., Farver, T. B., and Hinton, D. E. 1998. Viral hemorrhagic septicemia virus, *Ichthyophonus hoferi*, and other causes of morbidity in Pacific herring *Clupea pallasii* 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, T. J. II, Moffitt, S. D., and 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., and Willits, N. H. 2003. Role of disease in abundance of a Pacific herring (*Clupea pallasii*) population. *Canadian Journal of Fisheries and Aquatic Sciences*, 60: 1258–1265.
- Maunder, M. N., and 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., and 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., and 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., and Thorson, J. T. 2019. Modeling temporal variation in recruitment in fisheries stock assessment: a review of theory and practice. *Fisheries Research*, 217: 71–86.
- Maunder, M. N., and 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.
- Miller, T. J., Hare, J. A., and 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.
- Monnahan, C. C., and Kristensen, K. 2018. No-U-turn sampling for fast Bayesian inference in ADMB and TMB: introducing the admtn and tmbstan R packages. *Plos ONE*, 13: e0197954.
- Monnahan, C. C., Thorson, J. T., and Branch, T. A. 2017. Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods in Ecology and Evolution*, 8: 339–348.
- Moran, J. R., Heintz, R. A., Straley, J. M., and Vollenweider, J. J. 2018. 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.
- Moran, J. R., and Straley, J. M. 2019. Long-term monitoring of humpback whale predation on Pacific herring in Prince William Sound 19120114-O. In Exxon Valdez Oil Spill Restoration Project Annual Report. Exxon Valdez Oil Spill Trustee Council: Anchorage, AK. <https://evostc.state.ak.us/restoration-projects/project-search/ltm-program-long-term-monitoring-of-humpback-whale-predation-on-pacific-herring-in-prince-william-sound-19120114-o/>.
- Mueter, F. J., Ware, D. M., and 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.
- Muradian, M. L., Branch, T. A., Moffitt, S. D., and Hulson, P.-J. F. 2017. Bayesian stock assessment of Pacific herring in Prince William Sound, Alaska. *Plos ONE*, 12: e0172153.
- Myers, R. A. 1998. When do environment-recruitment correlations work? *Reviews in Fish Biology and Fisheries*, 8: 285–305.
- Needle, C. L. 2001. Recruitment models: diagnosis and prognosis. *Reviews in Fish Biology and Fisheries*, 11: 95–111.
- Okamoto, D. K., Poe, M. R., Francis, T. B., Punt, A. E., Levin, P. S., Shelton, A. O., Armitage, D. R. *et al.* 2020. Attending to spatial social-ecological sensitivities to improve trade-off analysis in natural resource management. *Fish and Fisheries*, 21: 1–12.
- Oken, K. L., Essington, T. E., and 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.
- Okey, T. A., and 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. Final Report, Restoration Project 99330-1. Exxon Valdez Oil Spill Trustee Council: Anchorage, AK.
- Pearson, W. H., Deriso, R. B., Elston, R. A., Hook, S. E., Parker, K. R., and 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.
- 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.
- Piironen, J., and Vehtari, A. 2017. Comparison of Bayesian predictive methods for model selection. *Statistics and Computing*, 27: 711–735.
- Polovina, J., Mitchum, G., and 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.
- Puerta, P., Ciannelli, L., Rykaczewski, R. R., Opiekun, M., and 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., Hurtado-Ferro, E., and Whitten, A. R. 2014. Model selection for selectivity in fisheries stock assessments. *Fisheries Research*, 158: 124–134.
- Quinn, T., Marty, G. D., Wilcock, J., and Willette, M. 2001. Disease and population assessment of Pacific herring in Prince William Sound, Alaska. In Lowell Wakefield Symposium, pp. 363–379. University of Alaska Sea Grant College Program, Anchorage, AK.
- R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Reum, J. C., Essington, T. E., Greene, C. M., Rice, C. A., and 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.
- Rice, J., and Browman, H. I. 2014. Where has all the recruitment research gone, long time passing? *ICES Journal of Marine Science*, 71: 2293–2299.
- 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., and 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.
- Sewall, E., Norcross, B., Mueter, F., and 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.
- Sharma, R., Porch, C., Babcock, E., Maunder, M., and Punt, A. 2019. Recruitment: Theory, Estimation, and Application in Fishery Stock Assessment Models. *Fisheries Research*, 217: 1–4.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and 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., and Palsson, W. 2017. Assessment of the arrowtooth flounder stock in the Gulf of Alaska. NOAA.
- Straley, J. M., and Moran, J. R. 2006–2020. Significance of Whale Predation On Natural Mortality Rate of Pacific Herring in Prince William Sound, Alaska. In Pelagic: Humpback whale predation on herring. Gulf of Alaska Data Portal. Alaska Ocean Observing System (AOOS). https://gulf-of-alaska.portal.aocos.org/#metadata/54adceab-74cb-4419-b02c-bacb6d2acb8b/project/folder_metadata/2660555t.
- Straley, J. M., Moran, J. R., Boswell, K. M., Vollenweider, J. J., Heintz, R. A., Quinn, T. J. II, Witteveen, B. H. *et al.* 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.

- Sturdevant, M. V., Brenner, R., Fergusson, E. A., Orsi, J. A., and Heard, B. 2013. Does predation by returning adult pink salmon regulate pink salmon or herring abundance. North Pacific Anadromous Fish Commission Technical Report. 9: 153–164.
- Szuwalski, C. S., Vert-Pre, K. A., Punt, A. E., Branch, T. A., and Hilborn, R. 2015. Examining common assumptions about recruitment: a meta-analysis of recruitment dynamics for worldwide marine fisheries. *Fish and Fisheries*, 16: 633–648.
- Teerlink, S. F., von Ziegesar, O., Straley, J. M., Quinn, T. J., Matkin, C. O., and 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.
- 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 (16p.).
- Trites, A., and 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., Shelton, A. O., and 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.
- Vehtari, A., Gelman, A., and Gabry, J. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27: 1413–1432.
- Vehtari, A., and Ojanen, J. 2012. A survey of Bayesian predictive methods for model assessment, selection and comparison. *Statistics Surveys*, 6: 142–228.
- Vetter, E. F. 1988. Estimation of natural mortality in fish stocks: a review. *Fisheries Bulletin*, 86: 25–43.
- Walters, C. J., and Ludwig, D. 1981. Effects of measurement errors on the assessment of stock–recruitment relationships. *Canadian Journal of Fisheries and Aquatic Sciences*, 38: 704–710.
- Ward, E. J., Adkison, M., Couture, J., Dressel, S. C., Litzow, M. A., Moffitt, S., Neher, T. H. *et al.* 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., and Thomson, R. E. 2005. Bottom-up ecosystem trophic dynamics determine fish production in the Northeast Pacific. *Science*, 308: 1280–1284.
- Watanabe, S. 2013. A widely applicable Bayesian information criterion. *Journal of Machine Learning Research*, 14: 867–897.
- Wilberg, M. J., and Bence, J. R. 2008. Performance of deviance information criterion model selection in statistical catch-at-age analysis. *Fisheries Research*, 93: 212–221.
- Williams, E. H., and Quinn, T. 2000. Pacific herring, *Clupea pallasii*, recruitment in the Bering Sea and north-east Pacific Ocean, II: relationships to environmental variables and implications for forecasting. *Fisheries Oceanography*, 9: 300–315.
- Womble, J. N., and 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.
- Yang, M.-S. 1993. Food Habits of the Commercially Important Groundfishes in the Gulf of Alaska in 1990. NOAA.
- Yeh, S.-W., Kang, Y.-J., Noh, Y., and Miller, A. J. 2011. The North Pacific climate transitions of the winters of 1976/77 and 1988/89. *Journal of Climate*, 24: 1170–1183.

Handling Editor: Shijie Zhou