ICES Journal of Marine Science, 2024, Vol. 81, Issue 2, 317–333
https://doi.org/10.1093/icesjms/fsad199
Received: 11 September 2023; **revised:** 1 November 2023; **accepted:** 7 November 2023
Advance access publication date: 5 January 2024
Original Article

Management strategy evaluation of harvest control rules for Pacific Herring in Prince William Sound, Alaska

Joshua A. Zahner * **and Trevor A. Branch**

School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195, United States
*Corresponding author. 1122 NE Boat St, Box 355020, Seattle, WA 98195-5020; e-mail: joshuazahner@gmail.com

Abstract

Management strategy evaluation (MSE) provides a mechanism to test the relative performance of alternative management strategies on a fishery. For Pacific herring in Prince William Sound, Alaska, no directed fisheries have occurred in over 30 years, providing an opportunity to evaluate potential management strategies before a fishery is opened. Here we evaluate and compare 10 harvest control rules (HCRs) ranging from simple threshold rules to rules accounting for population age structure, biomass trends, and weight distribution using an MSE integrated with a fully Bayesian stock assessment estimation model. We developed a utility function that shows simple threshold HCRs outperform the more complex rules, especially for catch stability. According to this utility function, the best rule had a lower limit threshold than the current default rule, while the worst rule had a higher limit threshold. Our simulations demonstrate that sufficient computing power exists for MSEs based on Bayesian estimation methods, thus opening a pathway for MSEs to simulation test probabilistic control rules, which provide a buffer against scientific uncertainty and should reduce the risk of overfishing.

Keywords: forage fish; management procedure; fisheries management; environmental modeling

Introduction

Fisheries management, at its core, involves deciding which of several management decisions will result in the most desirable outcome for the participants of the fishery (Hilborn and Walters 1992). This process often occurs in two distinct phases: (i) for each possible management action, potential outcomes, and their respective probabilities are identified; and (ii) given the potential outcomes, a "best" management action is chosen and implemented. Hilborn and Walters (1992) propose that most of a fisheries manager's time is spent in phase one, and a great many tools have been developed over the years for managers to use during this first phase. For the most data-rich fisheries, this first phase of management often takes the form of a full stock assessment model, in which data are collected about the current state of the population, a computational model of the biological dynamics of the population is built, and a statistical estimation framework is used to estimate the true state of the fishery from the available data. However, once the stock assessment model has estimated the true population state, the second phase of fisheries management is still required—choosing the best management action—for which there are far fewer tools available. The most prominent of these tools is management strategy evaluation (MSE).

MSE is a simulation framework for assessing the relative performance of potential management strategies that has been used in fisheries around the world (Rademeyer et al. 2007, Punt et al. 2016). The framework consists of two interacting components: an "operating model" (OM) that simulates the true state of the fishery (including stochasticity) and generates data (with observation error) mimicking the data available for the fishery; and a "management strategy" that assesses the state of the fishery based on the generated data (via an "estimation method," EM), and calculates an allowable catch for the following year (Fig. 1). The two models operate in a closed loop for many years (15–40 years is common), comprising a single simulation run. This process is then repeated many times (often 100–1000 times depending on computational resources) to account for uncertainty in population dynamics, monitoring data, and management implementation (Punt et al. 2016). Performance metrics are computed across the simulation runs and used to compare the various candidate management strategies based on how well they meet specific performance goals.

The MSE framework was first developed by the International Whaling Commission (IWC) in the 1980s to identify a catch-limit algorithm for the harvest of baleen whale species that would be robust to scientific uncertainty (Punt and Donovan 2007). The framework was subsequently applied in the 1990s in South Africa to identify robust harvest policies for multiple fish and invertebrate species (de Moor et al. 2022). In more recent years, the framework has expanded to many countries and regions (De Oliveira et al. 2008, Holland 2010), including Australia (e.g. Punt and Smith 1999, Dichmont et al. 2006), New Zealand (e.g. Bentley et al. 2003), Europe (e.g. Kell et al. 2005), Canada (e.g. Cox and Kronlund 2008), and the United States (e.g. Hicks et al. 2013), and has been used by multiple Tuna Regional Fisheries Management Organizations to manage highly migratory pelagic species (e.g. Polacheck et al. 1999, Kurota et al. 2010). The majority of operational MSEs across the world were developed specifically to evaluate potential harvest control rules (HCRs) or management procedures (MPs). However, the framework has found additional uses in a research context, including in value of information studies (Carruthers and Kell 2016, Muradian et al. 2019, Xia et al. 2021) and tests of the robustness of existing management strategies and estimation methods to

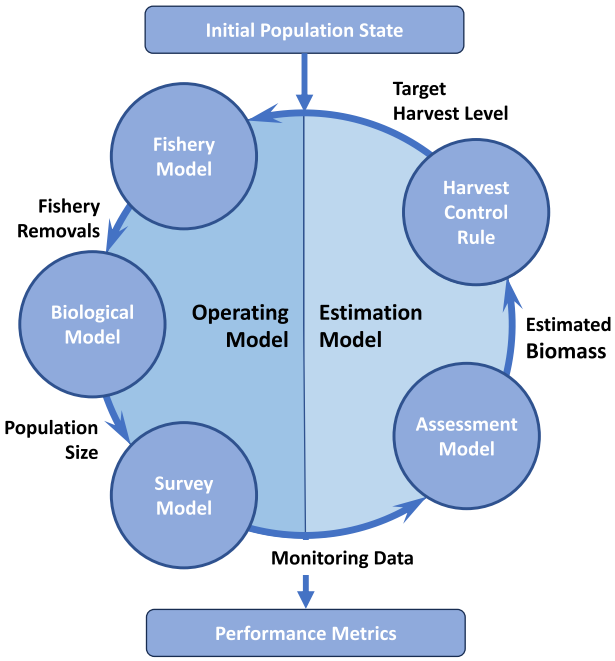

Figure 1. The MSE simulation framework (adapted from Punt et al. 2016).

environmental change (A'mar et al. 2009, Punt et al. 2013), spatial complexity (Jacobsen et al.), and ecosystem interactions (Fulton et al. 2014, Surma et al. 2021).

Because MSEs run thousands of computationally costly estimation routines, their adoption has been limited by computing power until relatively recently (Goethel et al. 2019), and even then has been largely restricted to frameworks using maximum-likelihood (or penalized maximum likelihood) estimation (MLE; Monnahan et al. 2019). Few, if any, MSEs include a fully Bayesian estimation method. Recent integration of advanced Markov-chain Monte Carlo (MCMC) sampling methods into major fisheries stock assessment frameworks, notably the No U-Turn Sampler (NUTS) algorithm (Monnahan et al. 2017, 2019, Monnahan and Kristensen 2018), has markedly reduced the computational requirements of Bayesian estimation methods, allowing for their use within an MSE. The MSE developed here represents one of the first to use a fully Bayesian, age-structured estimation method, though several other operational and research MSEs have implemented OMs conditioned on Bayesian parameter estimates (e.g. Jacobsen et al.). Use of such Bayesian assessment methods has the benefit of providing a more thorough picture of the uncertainty associated with a given MP.

We develop and apply an MSE framework to Pacific herring (*Clupea pallasii*) in Prince William Sound (PWS), Alaska, taking advantage of the NUTS algorithm to allow for convergence of the estimation method in a matter of minutes instead of many hours. This population sustained annual catches averaging over 5000 mt during the 1970s and 1980s but suffered from a severe population crash in 1992 that resulted in biomass declining by over 50% in less than a year and the closure of directed fisheries (Muradian et al. 2017). In the thirty years since the crash, the population has failed to rebound to pre-crash levels, despite continued fishery closures since 1999. However, the population has increased towards the lower limit threshold needed to reopen fisheries in recent years. The current HCR that is presently in place for the population was established by the Alaska Board of Fisheries in 1994 (Morstad et al. 1996, Botz et al. 2010) after the initial population crash, but has never been used for setting non-zero catches. We therefore have a unique opportunity to develop and evaluate a new HCR before herring fisheries are reopened. Here we use MSE to evaluate 10 diverse harvest control rules in combination with a Bayesian estimation method for PWS herring that range from threshold rules (including the current rule), to those accounting for population age structure diversity, recent trends in biomass, and average individual weight.

Methods

Operating model (OM)

The OM utilized here is based on the population dynamics equations developed and discussed in Muradian et al. (2017) and updated by Trochta and Branch (2021). Broadly, the OM consists of a model for fishery removals with four fisheries; an age-structured population dynamics model with 10 age classes, constant background mortality, and lognormally distributed recruitment; and an observation model that emulates the four scientific surveys currently used to monitor the PWS herring population.

Fisheries: Four historical fisheries have been present in PWS: purse-seine sac roe, gillnet sac roe, food-bait, and pound spawn-on-kelp. Harvest allocations to each fishery were based on the most recent PWS herring management plan (Botz et al. 2010):

- Purse-seine fishery: 63.1%
- Gillnet fishery: 3.7%
- Food-bait fishery: 17.7%
- Pound spawn-on-kelp fishery: 15.7%

A fifth fishery (wild spawn-on-kelp fishery) is excluded from the OM as it does not remove spawning fish from the population. Historically, it was allocated 8.0% of the allowable harvest. The allocations of the remaining four fisheries have been adjusted proportionately to reflect this. All four fisheries possess the same age-3 knife-edged selectivity curve (υ_a; Table 1). This assumption is likely violated in the actual fishery, but little information is available to parameterize unique selectivity curves for each fishery. Removals for the purse-seine, gillnet, and food-bait fishery are distributed proportionally across all selected age classes and occur prior to spawning, as they remove reproductively active adult fish. Removals for the pound spawn-on-kelp fishery are similarly distributed across the population but occur after reproduction, as it removes spawned eggs. Adult mortality in the pound spawn-on-kelp fishery is assumed to be 75% (Muradian et al. 2017), and the removal of spawned eggs does not modify recruitment. Catches for all simulated fisheries are assumed to be known without error (Table 2; Equation 2.2–2.5).

Biological model: The OM includes an age-structured population model that closely mirrors the population dynamics model used in the current PWS herring stock assessment model (Muradian et al. 2017, Trochta and Branch 2021, Zahner and Branch 2023; Table 2). This model includes ages 0–8 and a plus-group at age 9 (Table 2; Equation 2.8–2.10).

Survival by year and age class ($S_{y,a}$) is modeled as a function of natural mortality (M_a; Table 1) split into two half years to account for natural and fishery mortality

Table 1. Symbols and parameters used throughout the model equations.

Parameter name	Symbol
Allowable harvest rate	h
Selectivity-at-age	v_a
Natural mortality at age	M_a
Long-term average log-recruitment (millions of individuals)	$\bar{R}$
Log-recruitment deviation from average in year, y	δ_y
Recruitment variability in year, y	σ_y^2

acting on the population in different seasons. Natural mortality includes disease, predation, competition, and other ecological factors. It is assumed to be fixed for age-classes 0–8 at $M_{a_{0-8}} = 0.25\ yr^{-1}$, as this is considered the base value for M_a excluding additional mortality from disease (Quinn et al. 2001, Muradian et al. 2017) and matches with what is used by the estimation method. Meanwhile M_a for the plus group is taken to be $M_{a_{9+}} = 0.87\ yr^{-1}$, as was estimated by the most recent herring stock assessment (Zahner and Branch 2023). Age- and year-specific survival rates are calculated using Equation 2.7. Maturity, weight, and fecundity at age are all time-invariant and based on estimated mean values derived from historical surveys (Fig. 2; Table 3).

Recruitment (R_y) is assumed to be log-normally distributed about the long-term average recruitment level ($\bar{R}$; Table 1) of ∼127 million annual recruits (Equation 2.8). The historical recruitment time series for this population (Fig. 3) shows evidence of two recruitment regimes (Dias et al. 2022): a "high" regime during 1980–1992 and a "low" regime during 1993–2021. Means and standard deviations of the estimated recruitment deviations (δ_y; Table 1) for each regime period were used to parameterize regime-specific normal distributions from which future recruitment deviations were drawn [high: $\delta_y \sim N(\mu = 0.345,\ \sigma = 1.140)$; low: $\delta_y \sim N(\mu = -1.289,\ \sigma = 0.961)$]. Recruitment deviations are not required to sum to zero over the simulation period and were converted into numbers of individuals using Equation 2.8. Forecasted recruitment assumed alternating 15-year regimes, the approximate length of the previous high recruitment regime, starting with the "high" regime in the first simulation year (Fig. 3). The OM begins with a "high" regime to ensure that the population rebounds to a level where fishing is consistently occurring for all HCRs.

Observation model: three scientific survey indices were generated: an aerial mile-days-of-milt survey, an age-composition (ASL) survey, and an age-1 aerial school survey (see Table 4 for the parameterizations). Other data sources have been available in the past for use in monitoring the population, including an egg-deposition survey (which is assumed to measure absolute abundance), two hydroacoustic surveys, and age-composition information derived from purse-seine fisheries catch. None are simulated into the future as there is no guarantee that such data sources will be available in the future and age-composition information is already available from the age-composition survey. Seroprevalence data are also annually available and have been shown to be useful in estimating disease mortality and population abundance (Trochta et al. 2022), but these data are not included in the estimation method or simulated in future years.

The mile-days-of-milt survey estimate is assumed to be log-normally distributed about the total post-fishery spawning biomass, and proportional to true biomass as in the stock assessment (Muradian et al. 2017, Zahner and Branch 2023). Age-composition data from the survey is assumed to follow a multinomial distribution with a sample size of 1500 fish. Finally, the age-1 aerial school survey (Pegau 2022) is assumed to follow a negative-binomial distribution with an overdispersion parameter (k) set to the value (2.147) reported in the most recent stock assessment (Zahner and Branch 2023). In this parameterization of the negative binomial, a smaller value of k represents less overdispersion. Annual weight-at-age and fecundity-at-age are taken to be the average weight-at-age and fecundity-at-age in years they were collected since 1980 and are assumed to be known without error. The sex ratio of the population is assumed to follow a normal distribution with the mean and standard deviation of historical sex ratio data collected by the ASL survey.

Estimation model (EM)

The EM is the Bayesian Age Structured Assessment (BASA) model for Prince William Sound Pacific herring (Muradian et al. 2017). The model fits to historical catch data, historical egg

Table 2. Operating model (OM) equations (adapted from Muradian et al. 2017).

Equation number	Description	Equation
2.1	Total allowable catch	$TAC = h \cdot B_y$
2.2	Purse-seine catch	$C_{1,y} = TAC \cdot 0.631$
2.3	Gillnet catch-at-age	$C_{2,y,a} = h \cdot 0.03 \cdot N_{y,a} \cdot v_a$
2.4	Pound catch-at-age	$C_{3,y,a} = h \cdot 0.177 \cdot N_{y,a} \cdot v_a$
2.5	Food-bait catch-at-age	$C_{4,y,a} = h \cdot 0.157 \cdot N_{y,a} \cdot v_a$
2.6	Spring catch (millions of fish), total number of fish caught by purse-seine, gillnet, and pound fishery	$C_{S,y,a} = \hat{\theta}_{y,a} C_{1,y} + C_{2,y,a} + C_{3,y,a}$
2.7	Half-year survival	$S_{y,a} = exp(-0.5M_a)$
2.8	Pre-fishery total abundance, age 0	$N_{y+1,\ 0} = exp(\bar{R} + \delta_{y+1} - 0.5\sigma_{y+1}^2)$
2.9	Pre-fishery total abundance, ages 1–8	$N_{y+1,\ a+1} = [(N_{y,\ a} - C_{S,\ y,\ a})S_{y,\ a} - C_{4,\ y,\ a}]S_{y,\ a}$
2.10	Pre-fishery total abundance, age 9+	$N_{y+1,\ 9_+} = [(N_{y,\ 8} - C_{S,\ y,\ 8})S_{y,\ 8} - C_{4,\ y,\ 8}]S_{y,\ 8} + [(N_{y,\ 9_+} - C_{S,\ y,\ 9_+})S_{y,\ 9_+} - C_{4,\ y,\ 9_+}]S_{y,\ 9_+}$
2.11	Pre-fishery age composition	$\hat{\theta}_{y,a} = \frac{N_{y,\ a}}{\sum_a N_{y,\ a}}$
2.12	Pre-fishery total biomass	$B_y = \sum_{a=3}^{9+} N_{y,a} \cdot W_{y,a}$

Table 3. Natural mortality (M), survival (S), maturity, weight, fecundity, and proportion of the age class weighing >110 g for each of the 10 age classes used in the OM and EM.

Age	Natural mortality (M)	Survival (S)	Proportion mature	Average weight (g)	Average fecundity (eggs)	Proportion >110 g
0	0.25	0.778	0.00	–	–	–
1	0.25	0.778	0.00	–	–	–
2	0.25	0.778	0.00	–	–	–
3	0.25	0.778	1.00	70.34	10,289	0.01
4	0.25	0.778	1.00	90.24	14,084	0.28
5	0.25	0.778	1.00	116.16	17,308	0.65
6	0.25	0.778	1.00	131.81	20,284	0.85
7	0.25	0.778	1.00	146.67	22,775	0.94
8	0.25	0.778	1.00	161.97	25,397	0.99
9+	0.87	0.418	1.00	181.85	28,188	0.99

Age classes for which biological information is unavailable due to data constraints are indicated by –.

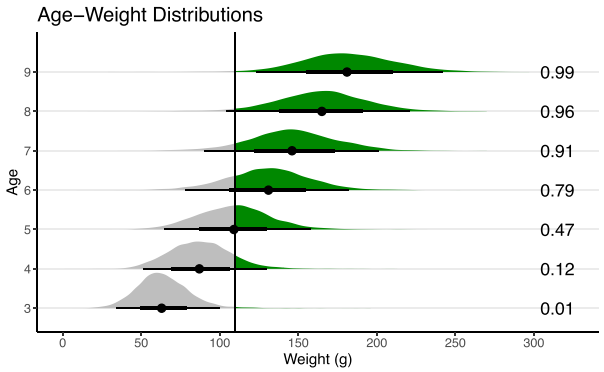

Figure 2. Weight distribution for age 3+ based on data from 1980 to 2021 provided by the Alaska Department of Fish and Game. The vertical line denotes the 110 g threshold used for the Big Fish rule, with the proportion to the right of this line listed on the right for each age class. Points and lines show the median, 50 and 95% CI.

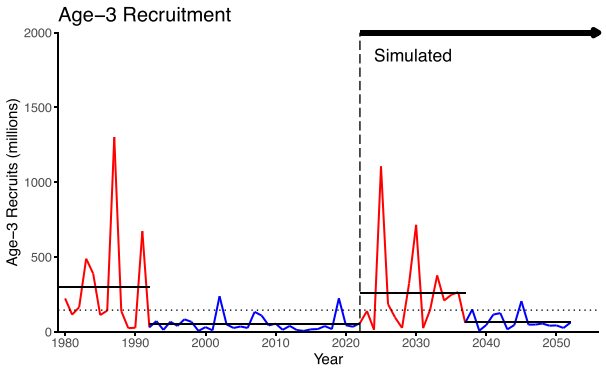

Figure 3. One randomly chosen future recruitment time series. The historical recruitment time series (1980–2021) shows evidence of two distinct regimes, a "high" regime (red) and a "low" regime (blue). Future recruitment is simulated from the "high" regime for 15 years and then from the "low" regime for 15 years. Solid horizontal lines indicate the mean recruitment level for each regime, and the horizontal dotted line indicates the average recruitment level ($\bar{R}$) from 1980 to 2022.

deposition data, and more recent milt, hydroacoustic biomass, age-1 aerial, and age-composition survey data. The EM estimates recruitment separately for each year and does not assume that there are low and high recruitment regimes. The EM runs in AD Model Builder (ADMB) version 13.1 (Fournier et al. 2012), and MCMC sampling is performed using the NUTS algorithm through the "adnuts" R package (Monnahan et al. 2019).

The EM used here requires ~2 min to obtain 2000 MCMC samples across four MCMC chains using the NUTS algorithm for 42 years of data (Zahner and Branch 2023). The addition of 30 years of simulated data from the MSE (a total of 72 years of data by the end of the simulations) would bring the average runtime of the model to ~9 min (again for 2000 MCMC samples and four chains). While this runtime is quite low compared to other Bayesian assessments (Monnahan et al. 2019), it would have required >50 days on the available computational resources to finish the 1500 MSE simulations reported here. Instead, we ran these simulations using a single MCMC chain for 1000 iterations. The small number of iterations and single chain is often considered insufficient to guarantee model convergence or an accurate characterization of the posterior

Table 4. Survey model sampling distributions and parameterizations.

Survey name	Sampling distribution	Parameters
Mile-days-of-milt	Log-normal: $\sim LN(\mu, \sigma)$	$\mu = 0$ $\sigma = 0.32$
Spawner age-composition	Multinomial: $\sim MN(s, p)$	$s = 1500$ $p = N_{y,a} / \sum N_{y,a}$
Age-1 aerial juvenile	Negative binomial $\sim NB(k, \mu)$	$\mu = N_{y,2} e^{q_j}$ $q_j = 71.16$ $k = 2.17$

Standard deviations for the Mile-days-of-Milt and hydroacoustic surveys follow from Muradian et al. (2019). $N_{y,a}$ represents the vector of number of individuals in each age class age-3 and greater during year y, s is the sample size of the multinomial distribution, p is the proportion of the population in each age class, q_j represents the catchability of the aerial school survey, and k is the overdispersion parameter for the negative binomial distribution.

Table 5. Harvest control rule parameterizations based on annual harvest rates (h).

HCR	Name	Type	h_{min}	h_{max}	Limit (mt)	Target (mt)	Notes
1	Default	Threshold	0.0	0.2	19,958	38,555	
2	Low harvest	Threshold	0.0	0.1	19,958	38,555	
3	High harvest	Threshold	0.0	0.4	19,958	38,555	
4	Low threshold	Threshold	0.0	0.2	10,000	38,555	
5	High threshold	Threshold	0.0	0.2	30,000	38,555	
6	Three-step threshold	Threshold	0.0	0.6	19,958	38,555	Default rule with additional step to $h = 0.6$ occurring at a biomass of 60,000 mt
7	Evenness	Threshold	0.0	0.2	19,958	38,555	Default rule scaled by age structure evenness [Equation (6)]
8	Biomass gradient	Gradient	0.0	1.0	19,958	n/a	Default rule scaled by 3-year biomass change [Equation (7)]
9	Big fish	Threshold	0.0	0.2	19,958	38,555	Uses biomass of fish >110 g instead of total biomass
10	No fishing	Constant F	0.0	0.0	0	0.0	

There is no "Target" for the Gradient rule, as there is no biomass level beyond which the harvest rate reaches a plateau.

distributions of the parameters or derived quantities (Gelman et al. 2014). While neither ideal, nor best practice in Bayesian modeling, this MCMC sampling protocol was chosen to minimize total runtime, while still ensuring accurate calculations of the following years harvest rate by the HCRs. Analysis of the behavior of the BASA model under a variable number of MCMC chains and variable MCMC chain length (available in Appendix 1) showed that the model converged (as defined by standard Bayesian diagnostics) and accurately estimated stock biomass (to within 1% of the biomass reported by the official 2022 stock assessment), while reducing runtime eightfold. Under this MCMC sampling protocol, total runtime of all simulations in this paper decreased to 7 days.

Each run of the EM was considered converged if the number of divergent transitions was <0.001 and Gelman-Rubin diagnostic ($\hat{R}$) based on a single chain was <1.05 (Monnahan, C., Alaska Fisheries Science Center). Only simulations in which the estimation method converged for all 30 years and for all tested control rules were used for computing performance metrics.

Harvest control rules

Ten harvest control rules (HCRs) were evaluated (Table 5, Fig. 4). Seven of these rules are threshold-based rules, described in this paragraph, and the remaining three (the Evenness rule, the Gradient rule, and the Big Fish rule) are discussed thereafter. The Default rule is the HCR that is currently used to support management of the PWS herring population (Morstad et al. 1996, Botz et al. 2010). It linearly scales the harvest rate from 0.0 to 0.2 when biomass is between 19,958 and 38,555 mt, with a harvest rate of 0.0 below the lower threshold (h_{min}) and 0.2 above the upper threshold (h_{max}) [Equation (5); Fig. 4a.1]. The Low Harvest (Fig. 4a.2) and High Harvest (Fig. 4a.3) rules follow the same shape as the Default rule but with lower ($h_{max} = 0.10$) and higher ($h_{max} = 0.40$) maximum allowable harvest rates. Similarly, the Low Threshold (Fig. 4a.4) and High Threshold (Fig. 4a.5) rules, respectively, have lower ($l = 10,000$ mt) and higher ($l = 30,000$ mt) limit thresholds than the Default rule. Additionally, there is a Three-Step Threshold Rule that follows the same shape and parameterization as the Default rule but with an added step at 60,000 mt, beyond which the allowable harvest rate is 0.6 (Fig. 4a.6), and a No Fishing (Fig. 4a.8) rule that sets $h = 0.0$.

$$h = \begin{cases} h_{min} & B < l \\ \frac{B-l}{u-l} \times h_{max} & l \le B < u \\ h_{max} & B \ge u \end{cases} \quad (5)$$

The Evenness rule (Fig. 4b) reduces the allowable harvest rate when the population biomass is dominated by a single cohort. This HCR is implemented by reducing the harvest rate from the Default rule using the Shannon-Wiener evenness index (J′; Equation (6); Shannon and Weaver 1949). Specifically, harvest rates from the Default rule are rescaled by a multiplier between 0.5 and 1.0 when the evenness index (J') is between 0.4 and 0.8 (i.e. when J' < 0.4, harvest rates are multiplied by 0.5, when J' > 0.8, harvest rates are multiplied by 1.0, and when 0.4 < J' < 0.8 harvest rates are scaled proportionally from 0.5 to 1.0). These evenness threshold values were chosen from empirical estimates of age structure evenness, which ranged from 0.32 to 0.95 during 1980–2021 (Appendix 2).

$$h_{evenness} = h_{default} \cdot J$$

$$J = \begin{cases} 0.5 & J' < 0.4 \\ \frac{J'-0.4}{0.8-0.4} \times 0.5 + 0.5 & 0.4 \le J' \le 0.8 \\ 1.0 & J' > 0.8 \end{cases} \quad (6)$$

$$J' = \frac{-\sum_{a=3}^{10} p_a \ln(p_a)}{7}$$

The Gradient rule (Fig. 4c) rescales the harvest rate from the Default rule by the relative change in biomass in the most recent 3-year period [Equation (7)]. The net effect is that allowable harvest rates increase when biomass is increasing and decline when biomass is declining. Fishing is not allowed when biomass is below the limit threshold of 19 958 mt, though there is no specific maximum harvest rate above the target threshold because the gradient component of the control rule acts as a multiplicative scalar on the default rule. This new "gradient-based" harvest rate (h') is averaged together with the harvest rate recommended by the "default" rule to reduce the annual variation in allowable harvest rate [Equation (7)].

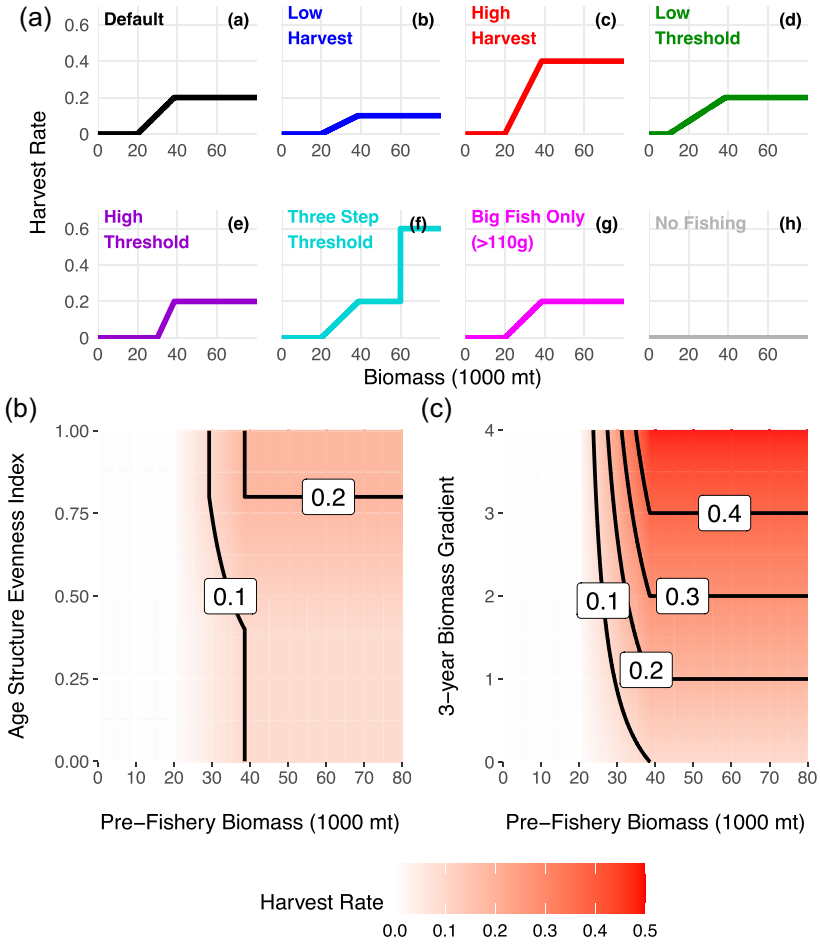

Figure 4. Relation between biomass and harvest rate for the 10 harvest control rules (HCRs). The harvest rate for the Evenness and Gradient rules is indicated by the red shading and contour lines in the last two plots.

$$h' = h_{default} \times \left(\frac{B_y}{B_{y-3}}\right)$$
$$h_{gradient} = \frac{h_{default} + h'}{2} \quad (7)$$

Finally, the Big Fish (Fig. 4a.7) rule has the same shape and parameterization as the Default rule but sets the allowable harvest rate based on the biomass of fish weighing >110 g rather than total stock biomass, as for other threshold rules. This accounts for fishery preferences for large fish, which are more valuable per unit weight than small fish. The weight of 110 g was selected because ~40% of sampled fish from 1980 to 2021 were larger than this threshold. Weights are not tracked for individual fish, thus the proportion of biomass of fish weighing >110 g in each age class is estimated from age-specific weight distributions (Fig. 2), based on age-sex-length data for the fishery (Morella 2022). For this rule, the harvest rate is set based on biomass of fish weighing >110 g but applied to all fish above the age of selectivity.

Performance metrics

The HCRs were assessed using eight metrics related to catch level, catch stability, and biomass level (Table 6). Average annual catch and average realized harvest rate metrics are reflective of fishery catch levels. Average annual catch variation and the proportion of years the fishery is closed are reflective of catch stability. Final year, average, and lowest annual relative biomass, as well as the proportion of years population biomass is <20 000 mt, are all reflective of biomass objectives. Biomass-related performance metrics are reported as relative to the biomass in each year that would have resulted in the absence of fishing ("dynamic" unfished biomass) to control for random variation in population dynamics. Performance metrics are computed over a 30-year period and presented as the median and 80% confidence interval across 150 simulations, each with a different recruitment time series.

HCRs were evaluated using a utility function based on one performance metric from each of the three key axes of performance: average annual catch, average annual catch variation, and average relative depletion. The utility function we chose is based on Bentley et al. (2003) and defines two thresholds for each performance metric, a minimum acceptable value (m), below which utility is 0, and a maximum utility value (k), above which utility is 1, with linear scaling between these thresholds [Table 7; Equation (8)]. The choice of values for m and k are subjective but are loosely derived from historical fishery performance and existing management objectives. The total utility of each HCR was computed by taking the geometric mean of the utility function value $U_{i,j}$ for average annual catch, average annual catch variation, and final year

Table 6. Performance metrics used to compare and evaluate the relative performance of harvest control rules.

Metric Name	Symbol	Definition	Equation
Annual catch	C_{ann}	Average annual fishery catch (mt)	$C_{ann} = \frac{\sum_{i=1}^{30} C_y}{30}$
Average annual catch variation	AAV	Average variation in catch between successive years	$AAV = \frac{1}{29}\sum_{y=2}^{30} \frac{\lvert C_y - C_{y-1}\rvert}{\bar{C}}$
Final year biomass relative to unfished conditions	B_{fin}	Biomass in the final simulation year relative to unfished biomass	$B_{fin} = median(\frac{B_{30}}{B_{30}^{unfished}})$
Average biomass relative to unfished conditions	B_{ann}	Average biomass across the simulation relative to unfished biomass	$B_{ann} = \frac{\sum_{y=1}^{30}(\frac{B_y}{B_y^{unfished}})}{30}$
Lowest biomass relative to unfished conditions	B_{min}	Lowest single-year biomass relative to unfished biomass	$B_{min} = min(\frac{B_y}{B_y^{unfished}})$
Proportion of years below threshold	P_B	Proportion of simulation years median biomass (B_y) is <19 958 mt	$P_B = \frac{\sum_{y=1}^{30}\{0\ B_y < 19958;\ 1\ B_y \geq 19958}{30}$
Proportion of years with an open fishery	P_O	Proportion of years that median biomass (B_y) is below the HCR specific limit threshold	$P_O = \frac{\sum_{y=1}^{30}\{0\ B_y < B_{lim};\ 1\ B_y \geq B_{lim}}{30}$
Average realized harvest rate	H_R	Median harvest rate across all simulation years	$H_R = \frac{median(\frac{C_y}{B_y})}{30}$

Metrics were computed for each simulation and are reported as the median and 80% CI across simulations. B_y is the biomass in year y and C_y is the catch in year y.

Table 7. Threshold values for each performance metric used for calculation of total utility.

Performance metric	Minimum acceptable utility (m)	Maximum utility value (k)
Annual catch	3,000 mt	20,000 mt
Average annual catch variation	1.0	0.0
Average depletion level	0.5	2.0

depletion.

$$U_{i,j} = \begin{cases} 0 & I < m \\ \frac{I-m}{k-m} & m \leq I \leq k \\ 1 & I > k \end{cases}. \quad (8)$$

Results

Biomass and catch trajectories

Pre-fishery biomass levels rebound to, or exceed, the target biomass level of 40,000 mt within the initial 15-year high-recruitment regime regardless of the harvest control rule that was applied (Fig. 5a). Once the low-recruitment regime begins in year 16 of the simulation, biomass begins to decline toward the current limit biomass threshold of 19,958 mt before stabilizing for all control rules. By the end of each 30-year simulation, biomass under each control rule was very similar, with only a 10% difference between the control rule resulting in the highest biomass level (no fishing) and the control rule resulting in the lowest biomass level (low threshold; Fig. 5b). The median biomass at the end of the 30-year simulation under the No Fishing Control Rule was 17,38 5 mt.

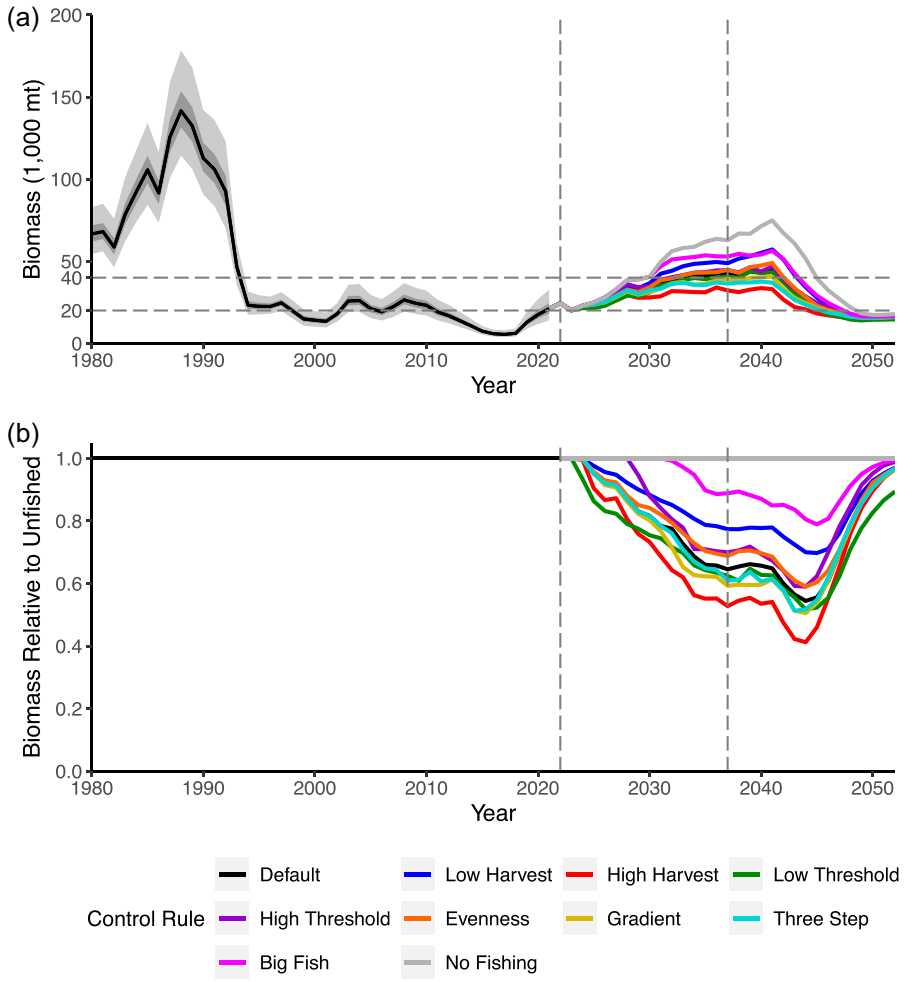

Figure 5. Biomass trajectories under the different HCRs. (a) The unscaled biomass trajectories in thousands of metric tons. (b) Biomass trajectories scaled relative to dynamic unfished biomass. Vertical dashed lines indicate the start of new recruitment regimes.

Peak biomass was largely a function of the maximum harvest rate. Control rules with lower maximum harvest rates (the No Fishing and Low Harvest rules) yielded larger maximum biomass levels. Stock biomass also spent more time above the management target threshold when maximum harvest rates were low (Fig. 5a). In addition, lower maximum harvest rates resulted in more variable biomass trajectories and less variable catch trajectories (Fig. 6). The position of the lower limit threshold (Low Threshold vs. High Threshold) had little effect on the biomass trajectory, though there were substantial differences in the resulting catch time series between control rules with different limit thresholds. The limit threshold position did not have any obvious effect on biomass variation (Fig. 6).

Control rule performance

There were substantial differences in median annual catch (2,173–6,629 mt; excluding the No Fishing and Big Fish rules for which median annual catch was 0 mt; Fig. 7a), average annual catch variation (0.375–0.858; excluding the No Fishing and Big Fish rules; Fig. 7b), and the proportion of years of fishery closure (0–0.6; Fig. 7h) across the control rules, while there were minimal differences in final year relative biomass (0.894–0.997; Fig. 7d). Control rules with lower maximum harvest rates and larger biomass levels over which fishing was allowed (the Low Harvest and Low Threshold rules) tended to have lower annual catch variation than control rules with the opposite properties.

The objective function identifies the Low Threshold, Default, and Evenness HCRs as having the highest total utility (Table 8). These three control rules all possess similar levels of annual catch, low annual catch variation (compared to the other HCRs), and similar levels of final year biomass. This indicates that the objective function is optimizing for low variation in annual catches, perhaps not surprising given the lack of contrast between rules in annual catches and final-year biomass. In the remaining seven control rules, the objective function also assigns the highest utilities to the rules with the lowest annual catch variation (Table 8).

Management objective tradeoffs

There is an inherent tradeoff between annual catch, catch variation, and biomass. As annual catch increases, relative biomass declines (Fig. 8d) and annual catch variation increases (Fig. 8g), while as biomass increases, annual catch variation declines (Fig. 8h). The three rules that the objective function identified as having the highest utility (the Low Threshold, Evenness, and Default rules) performed better than

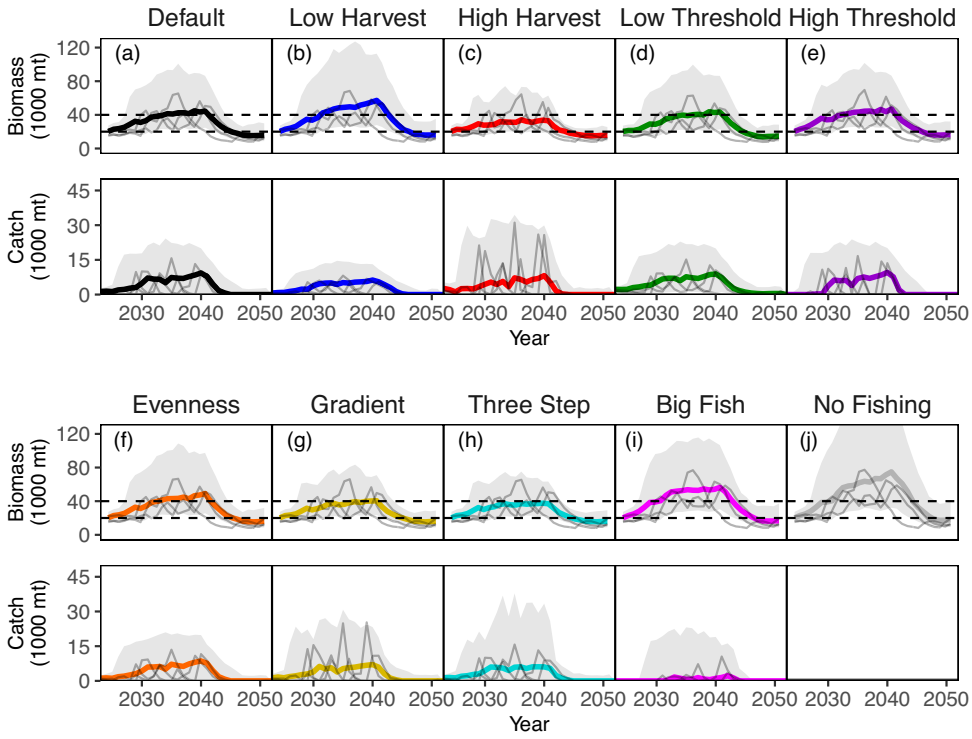

Figure 6. Biomass and annual catch trajectories over the 30-year simulation period. Thick colored lines are the median biomass and catch across 50 simulations, while thinner gray lines are five randomly selected biomass and catch trajectories. The gray-shaded regions indicate the inner-80th percentiles.

expected given the expected tradeoffs between performance metrics (Fig. 8).

How these tradeoff relationships are transformed by the objective function into a utility value demonstrates which tradeoffs are most important for improving the utility of a control rule. For the catch-biomass tradeoff, all 10 control rules occur in a narrow band of similar utility (Fig. 8b). A similar pattern can be observed for the catch-catch variation tradeoff (Fig. 8c). The biomass-catch variation tradeoff, however, occurs in a band that spans a wide range of utilities (Fig. 8f), indicating that, by moving along this tradeoff, significant gains in utility can be achieved. The three control rules with the highest calculated utility all occur near the upper end of this tradeoff plot, highlighting that the biomass-catch variation tradeoff is the most important tradeoff to manipulate to improve overall HCR utility.

Discussion

Harvest control rules

Threshold-based harvest control rules, such as those investigated here, are generally considered to be among the best for managing heavily exploited populations in general (Quinn et al. 1990) and forage fish in particular (Pikitch et al. 2012, Surma et al. 2021), and have been implemented for many data-rich fisheries in the United States alone (Free et al. 2023). Our results support simple threshold HCRs, since three of the four top-performing control rules all follow a simple threshold functional form (albeit the evenness rules include an extra scaling parameter). The more complex rules (the Gradient and Big Fish Rules), in attempting to account for more complicated dynamics, yield little increase in catch or biomass in return for much higher variation in annual catches. These two HCRs try to account for short-term trends in biomass or weight distribution, making them more susceptible to greatly changing harvest rates in response to survey variability or other short-term productivity declines that the less sensitive threshold rules may ignore. The Evenness and Three-Step Threshold rules, while not following the typical threshold functional form, have higher utility than the other complex rules for the inverse reason: their catch-setting algorithms are not as susceptible to short-term variations.

HCRs that account for gradients in either biomass estimates, survey indices, or other data have been developed for many other fisheries (e.g. Polacheck et al. 1999, Dichmont et al. 2006, Roel and De Oliveira 2007, Plaganyi et al. 2018, CCSBT 2019), but rarely take the form of the Gradient rule developed here. More often, these "gradient-based" rules compute the following year's catch as a function of the current year's catch, multiplied by the recent gradient in biomass or survey index, and are applied to data-limited fisheries, particularly throughout Europe (ICES 2022a). The functional form of the Gradient rule used in this study was developed to make full use of the available data for the stock, while also considering recent changes in biomass. While "gradient-based" rules have seen substantial use throughout the world's fisheries, it is acknowledged that they carry greater risk of overfishing and subsequent biomass decline than constant *F*-style rules (Polacheck et al. 1999, Roel and De Oliveira 2007). The results of our simulations using the Gradient HCR were not substantially worse than many of the threshold HCRs, though this could be because of its inherent reliance on a threshold HCR to set the base catch level that is subsequently rescaled by the biomass gradient. A rule more similar to the data-limited rules

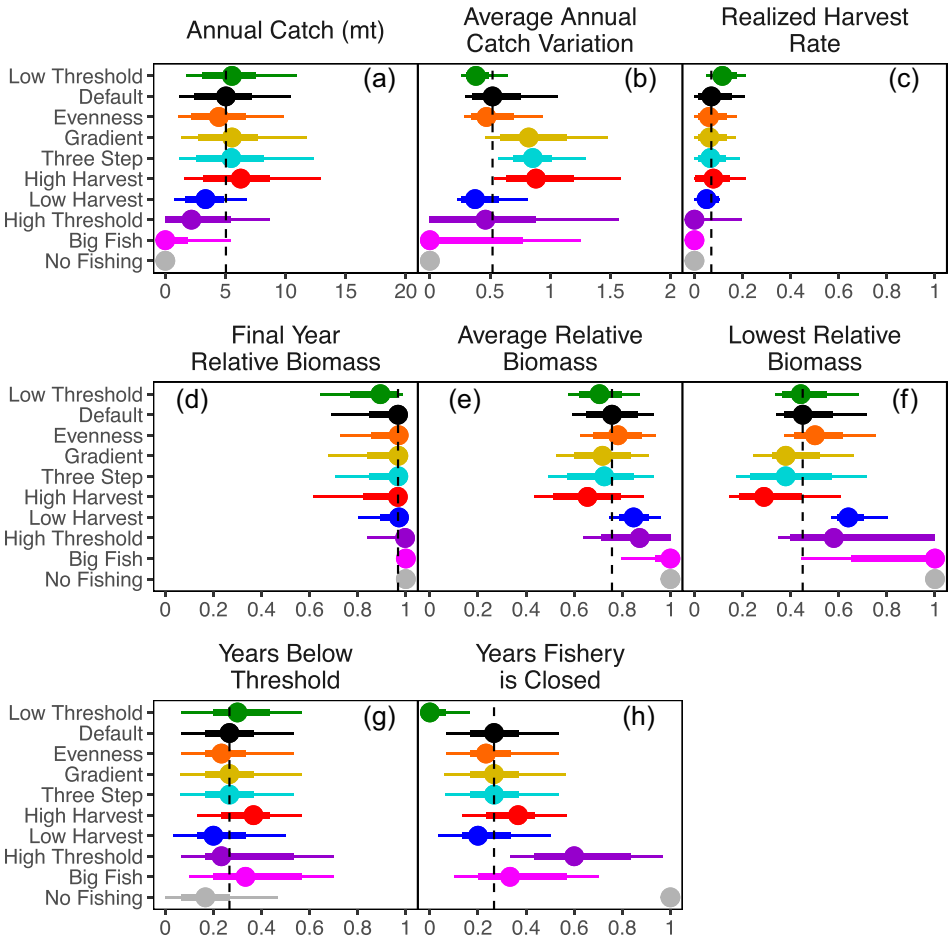

Figure 7. Performance metric summaries of tested harvest control rules. Large points indicate the median value of each performance metric across 150 simulations. Thick inner bars indicate the inner-50th percentile while the thin outer bars indicate the inner-80th percentile. The black, vertical lines indicate the level of each performance metric achieved by the Default rule. Refer to Table 6 for descriptions of the performance metrics.

Table 8. Total and relative utility values for each HCR as computed by Equation (8).

Control rule	Median total utility	Relative utility	Failed metric
Low threshold	**0.418**	**1.000**	–
Default	0.377	0.904	–
Evenness	0.349	0.835	–
Gradient	0.293	0.701	–
High harvest	0.277	0.663	–
Three-step threshold	0.271	0.648	–
Low harvest	0.228	0.547	–
Big fish only	0.000	0.000	Annual catch
High threshold	0.000	0.000	Annual catch
No fishing	0.000	0.000	Annual catch

Relative utilities are expressed relative to the HCR with the highest median total utility. The "Failed Metric" indicates which of the included metrics (annual catch, average annual catch variation, and average relative depletion) failed to meet the required utility threshold.

used by ICES for Tier-3 stocks in Europe (ICES 2022a), may show results more in line with those of other studies. Alternatively, this discrepancy between the results of our simulations and those performed by Polacheck et al. (1999) and Roel and De Oliveira (2007) could be due to substantial differences in the life histories of the species being examined (herring vs. tuna and mackerel, respectively). It is possible that the "boom-bust" dynamics of herring populations make including recent biomass gradients useful by allowing the fishery to take advantage of recent large-year classes at the expense of higher annual catch variation.

To our knowledge, no HCRs have previously been developed to account directly for population age structure, like the Evenness rule presented here, even though age structure is considered important for sustaining long-term healthy fish populations (Berkeley et al. 2004). By explicitly accounting for cohort dominance, the Evenness HCR should be more resilient in fisheries where the population often is dominated by a

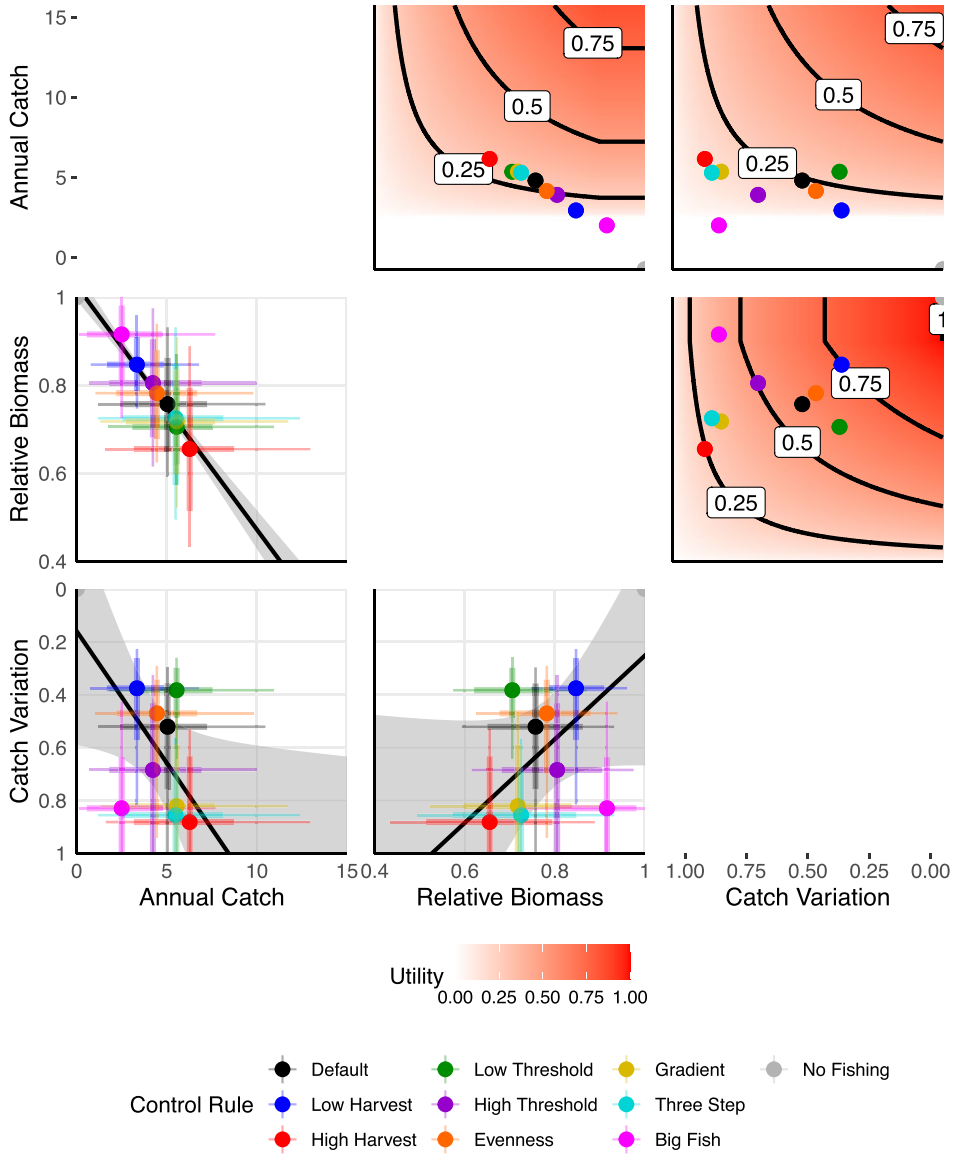

Figure 8. Performance tradeoff and total utility of harvest control rules. Colored points indicate the median metric value of each harvest control rule, while bars indicate the inner-50th and 80th percentiles. Solid black lines on lower three plots show the expected tradeoff between the two metrics. Background color on the upper three plots shows the combined utility [Equations (8)] for each combination of two of the three performance metrics.

single age class, as predicted by the portfolio effect (Schindler et al. 2010). Our results demonstrate that the Evenness rule outperforms most of the pure threshold rules with respects to average biomass levels and average annual catch variation, indicating that additional rules of this form could be worth exploring in other fisheries. It is likely that such HCRs would only be of particular use when applied to species and stocks that are reliant on a few large year-classes to sustain and grow the population (as is the case for most forage fish species), but further research into the appropriate form of the evenness scaling function [e.g. Equation (6)] is needed to determine the conditions under which their use is appropriate. The Evenness rule developed here actively changes the allowable harvest rate based on age structure, but alternative control rules could also manipulate the selectivity curve of a fishery as is recommended by Brunel and Piet (2013) (though this would be difficult for a purse-seine fishery, such as the one that predominantly operates on PWS herring) or use slot limits to target particular ages (Barnett et al. 2017), a strategy that was found to positively impact catch, biomass, and catch variability in Norwegian spring spawning herring (*Clupea harengus*; Enberg 2005). Neither slot limits nor modifications to the fishery selectivity curve were used for the Evenness HCR due to difficulties posed by annual modifications to fishing gear that would be required for implementation.

The Big Fish rule tested here is one of the first HCRs to set catch levels directly based on preferred weight of the individuals in the population, although weight distribution has been previously used as a performance metric for evaluating the performance of HCRs and has even been shown to be a better indicator of biomass decline than catches or CPUE (Punt et al. 2001). While the Big Fish control rule did not perform particularly well when compared to the other threshold rules in this study, such a class of HCRs could be seen as analogous to

length- or size-based control rules that have been implemented for many marine invertebrate species. In addition, if the performance metrics were weighted by economic value, the Big Fish rule may be ranked higher, since catches of big fish are worth more to the fishing industry.

Utility functions are frequently used in the field of economics to identify optimal actions (Tversky and Kahneman 1981), though their use in fisheries management, particularly in MSE, is limited. Where they are used in fisheries management, utility functions are often reduced to simple linear combinations of relevant management metrics (Keeney 1977, Lane and Stephenson 1998). This technique, however, makes specification of reasonable tradeoffs between competing metrics difficult to define for stakeholders (Bentley et al. 2003). We use the utility function described by Bentley et al. (2003) instead, because it allows for simple specification of minimum and maximum levels of utility for each metric, which is often easier for stakeholders to define. This type of utility function also has the benefit of quickly identifying HCRs that fail to meet pre-defined management objectives.

The utility function used here accounts for performance in catch, catch variation, and biomass, three of the most common fishery performance metrics. The other reported metrics, including realized harvest rate, final year biomass, and the proportion of years the fishery is closed, are typically highly correlated with one or more of the selected metrics, but can be more difficult to convey to stakeholders, and thus were not included when calculating total utility. Only a single set of threshold values for each metric were considered, based loosely on historical fishery performance, though exploration of a range of such values would represent a valuable extension to this work, as the ranking of the HCRs will depend on the selection of the threshold performance levels. Ideally, both the performance metrics themselves and the threshold levels used to define the utility functions would be selected through an elicitation process with stakeholders. From there, various methods could be employed to find metrics and threshold values that accurately reflect stakeholder values.

Model complexity

One of the strengths of the MSE framework is its ability to evaluate how well HCRs perform when the underlying assumptions of the EM are incorrect by changing the OM. This study does not investigate the performance of its HCRs under different OMs and instead uses an OM that nearly exactly matches the EM, except for recruitment. The OM used here is a relatively simple age-structured population dynamics model that includes a single spatial area, no time varying vital rates, no multispecies or ecosystem interactions, and integrates no environmental drivers. It is often recommended that MSEs performed on small pelagic fish, including most forage fish species, should account for one or more of these things to be compliant with ecosystem-based fisheries management practices (Siple et al. 2021). However, including many of the recommendations made by Siple et al. (2021) was infeasible given the available data and current understanding of the biology, ecology, and dynamics of the PWS herring population. Furthermore, the objectives of this MSE were to specifically evaluate the efficacy of HCRs for managing the herring population, and not to test their robustness to model misspecification or added model complexity.

The one place in the OM where additional complexity was integrated was recruitment, which was modeled as a series of regimes (distinct from the EM, which assumed recruitment deviates from a single long-term average level). Forage fish recruitment is characterized by "boom-bust" dynamics, and, thus, the functional form of recruitment is exceptionally difficult to estimate (Subbey et al. 2014). Environmental drivers are frequently considered to be the cause of such variability, and can lead to distinct high and low recruitment regimes (Szuwalski et al. 2019) as was assumed by the OM. There is substantial literature regarding how to incorporate environmental drivers into MSEs (Brunel et al. 2010, Punt et al. 2013, Haltuch et al. 2019), and there are multiple examples of MSEs using environmentally informed recruitment (e.g. Hurtado-Ferro et al. 2010, Tommasi et al. 2017). Much work has gone into trying to characterize drivers of recruitment for PWS herring, though no conclusive links have been discovered (Trochta and Branch 2021), and there remains concern about using environmental indices directly for recruitment predictions where no relationships truly exist (De Oliveira and Butterworth 2005). While environmental drivers were not directly included in the OM, modeling recruitment as occurring in distinct regimes is a first step towards accounting for environmental factors within the PWS ecosystem.

We believe these results to be robust to many of the recommendations made by Siple et al. (2021). However, if age 0–8 mortality was higher than the value of $M = 0.25$ that was used here (selected specifically to match the values used by the EM), there could be substantial impacts on the performance of the gradient and evenness HCRs, particularly. If M was substantially higher, the population age structure would be uneven more frequently than the EM would estimate, resulting in higher exploitation rates being recommended than should have been under the evenness rule. This would likely degrade the performance of the evenness rule with respect to conservation-based performance metrics (such as average biomass). Meanwhile, higher mortality could result in larger, and more frequent, declines in biomass, which would result in the gradient rule decreasing allowable catch levels. This could further amplify the degree of annual catch variation for the gradient rule, which would further increase the already high annual catch variation for this rule.

Extensions and future work

By using a fully Bayesian estimation method, this MSE was not explicitly constrained to rely solely on the mean or median of the estimated biomass distribution as would be the case for MSEs based on maximum likelihood methods. While not tested here, control rules based on other distributional percentiles could be easily investigated. Such probabilistic control rules (then referred to as CI-HCRs) make uncertainty propagation more transparent and easier to understand for decision-makers (Dankel et al. 2016). The International Council for the Exploration of the Sea (ICES) uses a variation on this technique for its tier-2 stocks, where catch recommendations are set as a percentile of the catch distribution under an F_{MSY} harvest policy (ICES 2022b). Similarly, on the US West Coast, the National Marine Fisheries Service uses the $P*$ (p-star) approach, whereby the allowable catch is reduced based on the uncertainty in biomass estimates for many stocks (Prager et al. 2003, Shertzer et al. 2010, Privitera-Johnson and Punt 2019). Such probability-based HCRs have also been shown to lead

to less variable fishing mortality rates and, ultimately, larger median stock biomass, than their deterministic counterparts (Mildenberger et al. 2020). While beyond the scope of this study, probabilistic control rules represent a potentially novel management technique that could now be computationally feasible to thoroughly study and test with MSE.

Another HCR form that was considered but ultimately discarded for this study are rules that inherently account for regimes. Such regime-based HCRs have been developed and tested for snow crab in Alaska and were found to be risky unless the dynamics of the stock are "truly regime-like" (Szuwalski and Punt 2012). In contrast, Mohn and Chouinard (2007) found that stocks displaying significant changes in productivity should be managed using HCRs that account for those changes in productivity. Among others, one of the major concerns regarding regime-based HCRs is how they detect a change in regime and how quickly they react to such a change. The Sequential T-test Analysis of Regime Shifts (STARS) algorithm (Rodionov 2004) has been shown to be able to reliably detect changes in regimes and has even been used within an MSE (Szuwalski and Punt 2012), but it is an open question how managers should react to a detected change in regime. One option would be to select a distinct HCR to apply when each regime is active, though this could lead to overfishing if a shift from a high-productivity to a low-productivity regime occurs and is not promptly detected. Such situations could be common depending on how different the regimes are and how much uncertainty is present in the data being used to identify the regimes. A more advanced option may be to have distinct regime-specific HCRs, as well as some linear combination of HCRs to apply when the ecosystem is moving between regimes, or when the regime cannot be positively identified. Either of these options could be useful in fisheries displaying long-running regimes, where shifts would be infrequent, but would likely lead to high annual catch variations if regimes switched frequently.

Future work on the PWS herring stock should seek to better integrate the recommendations of Siple et al. (2021), or test the robustness of a subset of the HCRs evaluated here to additional complexity in the OM. For instance, Pacific herring are key prey species for many marine predators in PWS (Pearson et al. 2012, Trochta and Branch 2021), and, thus, a MICE (Model of Intermediate Complexity for Ecosystem Assessments) could be a valuable way to evaluate which control rules are robust to wider ecosystem needs. Additionally, integration of additional spatial complexity could be important for evaluating future survey design, as clear evidence of range contraction and expansion for PWS herring as the total population size fluctuated has been identified (McGowan et al. 2021).

Conclusion

With this MSE developed and validated, future work can begin to use it to investigate many other questions related to PWS herring. A value-of-information study would be useful for determining economically optimal survey coverage in future years, especially if the fishery was re-opened. There are also outstanding questions related to spatial stock structure and ecosystem interactions that could be investigated through the development of more complex OMs Evaluating the robustness of the BASA model to model misspecification would also lend valuable insight into how best to manage this recovering population in coming years.

Acknowledgements

We thank Jennifer Morella (ADF&G) and Scott Pegau (ADF&G) for providing data input that were used for the stock assessment as well as parameterizing the OM. Data provided by ADF&G are the property of ADF&G, which 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 Sherri Dressel (ADF&G) for organizing a remote workshop with Alaskan herring managers to select reasonable HCRs to test within the MSE framework. Finally, we thank Andre Punt (UW), Kelli Johnson (NOAA), Margaret Siple (NOAA), Cole Monnahan (NOAA), and Sherri Dressel (ADF&G) for valuable discussion, feedback, and comments throughout the course of the study. Comments from three anonymous reviewers are also appreciated.

Data availability

The computer code to run these analyses, as well as the catch and population trajectory data output from each simulation, is available in a GitHub repo at: https://github.com/Ovec8hkin/pws-herring-mse. The Bayesian Age Structured Assessment model that serves as the estimation method is similarly available on GitHub at: https://github.com/Ovec8hkin/pws-herring-basa. The full suite of data is available on request.

Author contributions

TAB conceptualized the study and obtained funding. JAZ developed the computational framework, ran simulations, and analyzed the resulting data. JAZ wrote the initial draft, while TAB revised the text.

Conflict of interest: The authors certify that they have no conflicts of interest with regards to this work. 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.

Funding

Funding for this work was provided by the Exxon Valdez Oil Spill Trustee Council (EVOSTC). TAB was additionally supported in part by the Richard C. and Lois M. Worthington Endowed Professor in Fisheries Management.

References

A'mar ZT, Punt AE, Dorn MW. The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. *ICES J Mar Sci* 2009;**66**:1614–32. https://doi.org/10.1093/icesjms/fsp044

Barnett LAK, Branch TA, Ranasinghe RA *et al*. Old-growth fishes become scarce under fishing. *Curr Biol* 2017;**27**:2843–8.e2.https://doi.org/10.1016/j.cub.2017.07.069

Bentley N, Breen PA, PJ Starr, Sykes DR. Development and evaluation of decision rules for management of New Zealand rock lobster fisheries. New Zealand: Ministry of Fisheries, 2003, 14. https://fs.fish.govt.nz/Doc/17347/2003%20FARs/03_29_FAR.pdf.ashx

Berkeley SA, Hixon MA, Larson RJ *et al.* Fisheries sustainability via protection of age structure and spatial distribution of fish populations. *Fisheries* 2004;**29**:23–32. https://doi.org/10.1577/1548-8446(2004)29%5b23:FSVPOA%5d2.0.CO;2

Botz J, Hollowell G, Bell J *et al.* 2009 Prince William Sound area finfish management report. In: *Fishery Management Report No. 10–55. Alaska Department of Fish and Game*. Alaska: Anchorage, 2010.

Brunel T, Piet GJ, van Hal R *et al.* Performance of harvest control rules in a variable environment. *ICES J Mar Sci* 2010;**67**:1051–62. https://doi.org/10.1093/icesjms/fsp297

Brunel T, Piet GJ. Is age structure a relevant criterion for the health of fish stocks? *ICES J Mar Sci* 2013;**70**:270–83. https://doi.org/10.1093/icesjms/fss184

Carruthers TR, Kell LT. Beyond MSE: opportunities in the application of Atlantic bluefin tuna operating models. *Coll Vol Scien Pap ICCAT* 2016;**73**:2543–51.

CCSBT. Report of the twenty fifth meeting of the scientific committee. Comission for the Conservation of Southern Bluefin Tuna, 2019. https://www.ccsbt.org/sites/default/files/userfiles/file/docs_english/meetings/meeting_reports/ccsbt_27/report_of_SC25.pdf

Cox SP, Kronlund AR. Practical stakeholder-driven harvest policies for groundfish fisheries in British Columbia. *Fish Res* 2008;**94**:224–37. https://doi.org/10.1016/j.fishres.2008.05.006

Dankel DJ, Vōlstad JH, Aanes S. Communicating uncertainty in quota advice: a case for confidence interval harvest control rules (CI-HCRs) for fisheries. *Can J Fish AquatSci* 2016;**73**:309–17. https://doi.org/10.1139/cjfas-2015-0078

de Moor CL, Butterworth DS, Johnston S. Learning from three decades of management strategy evaluation in South Africa. *ICES J Mar Sci* 2022;**79**:1843–52. https://doi.org/10.1093/icesjms/fsac114

De Oliveira JAA, Butterworth DS. Limits to the use of environmental indices to reduce risk and/or increase yield in the South African anchovy fishery. *Afr J Mar Sci* 2005;**27**:191–203. https://doi.org/10.2989/18142320509504078

De Oliveira JAA, Kell LT, Punt AE *et al.* Managing without best predictions: the management strategy evaluation framework. In: Payne A, Cotter J, Potter T (eds.), *Advances in Fisheries Science. 50 Years on from Beverton and Holt*. Oxford: Blackwell Publishing Ltd, 2008, 104–34.

Dias BS, McGowan DW, Campbell R *et al.* Influence of environmental and population factors on Prince William Sound herring spawning phenology. *Marine Ecology Progress Series* 2022;**696**:103–117.

Dichmont CM, Deng A, Punt AE *et al.* Management strategies for short lived species: the case of Australia's Northern Prawn Fishery: 2. Choosing appropriate management strategies using input controls. *Fish Res* 2006;**82**:221–34. https://doi.org/10.1016/j.fishres.2006.06.009

Enberg K. Benefits of threshold strategies and age-selective harvesting in a fluctuating fish stock of Norwegian spring spawning herring *Clupea harengus*. *Mar Ecol Progress Ser* 2005;**298**:277–86. https://doi.org/10.3354/meps298277

Fournier DA, Skaug HJ, Ancheta J *et al.* AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim Methods Software* 2012;**27**:233–49. https://doi.org/10.1080/10556788.2011.597854

Free CM, Mangin T, Wiedenmann J *et al.* Harvest control rules used in US federal fisheries management and implications for climate resilience. *Fish Fisheries* 2023;**24**:248–62. https://doi.org/10.1111/faf.12724

Fulton EA, Smith ADM, Smith DC *et al.* An integrated approach is needed for ecosystem based fisheries management: insights from ecosystem-level management strategy evaluation. *PLoS One* 2014;**9**:e84242. https://doi.org/10.1371/journal.pone.0084242

Gelman A, Carlin JB, Stern HS *et al. Bayeisan Data Analysis*. Boca Raton, FL: Chapman and Hall, 2014.

Goethel DR, Lucey SM, Berger AM *et al.* Recent advances in management strategy evaluation: introduction to the special issue Under pressure: addressing fisheries challenges with "Management Strategy Evaluation". *Can J Fisheries Aquat Sci* 2019;**01**:1689–96.

Haltuch MA, Brooks EN, Brodziak J *et al.* Unraveling the recruitment problem: a review of environmentally-informed forecasting and management strategy evaluation. *Fish Res* 2019;**217**:198–216. https://doi.org/10.1016/j.fishres.2018.12.016

Hicks A, Taylor N, Taylor I *et al.* Status of the Pacific hake (whiting) stock in U.S. and Canadian waters in 2013 International Joint Technical Committee for Pacific hake. Newport, OR: Pacific Fisheries Management Council, 2013. https://www.pcouncil.org/documents/2013/03/status-of-the-pacific-hake-whiting-stock-in-u-s-and-canadian-waters-in-2013-final-document-march-4-2013.pdf

Hilborn R, Walters CJ. *Quantitative FIsheries Stock Assessment*. Boca Raton, FL: Chapman and Hall, 1992.

Holland DS. Management strategy evaluation and management procedures. *OECD Food Agricult Fisheries Pap* 2010;**25**:67.

Hurtado-Ferro F, Hiramatsu K, Shirakihara K. Allowing for environmental effects in a management strategy evaluation for Japanese sardine. *ICES J Mar Sci* 2010;**67**:2012–7. https://doi.org/10.1093/icesjms/fsq126

ICES. Advice on fishing opportunities (2022). *General ICES Advice guidelines* 2022a. https://doi.org/10.17895/ices.advice.19928060.v1

ICES. ICES technical guidance for harvest control rules and stock assessments for stocks in categories 2 and 3. *ICES Technical Guidelines* 2022b. https://doi.org/10.17895/ices.advice.19801564.v1

Jacobsen NS, Marshall KN, Berger AM *et al.* Climate-mediated stock redistribution causes increased risk and challenges for fisheries management. *ICES Journal of Marine Science* 2022;**79**:1120–32. https://doi.org/10.1093/icesjms/fsac029

Keeney RL. A utility function for examining policy affecting salmon on the Skeena River. *J Fish Res Board Can* 1977;**34**:49–63. https://doi.org/10.1139/f77-006

Kell LT, Pastoors MA, Scott RD *et al.* Evaluation of multiple management objectives for Northeast Atlantic flatfish stocks: sustainability vs. stability of yield. *ICES J Mar Sci* 2005;**62**:1104–17. https://doi.org/10.1016/j.icesjms.2005.05.005

Kurota H, Hiramatsu K, Takahashi N *et al.* Developing a management procedure robust to uncertainty for southern bluefin tuna: a somewhat frustrating struggle to bridge the gap between ideals and reality. *Population Ecol* 2010;**52**:359–72. https://doi.org/10.1007/s10144-010-0201-1

Lane DE, Stephenson RL. A framework for risk analysis in fisheries decision-making. *ICES J Mar Sci* 1998;**55**:1–13. https://doi.org/10.1006/jmsc.1997.0237

McGowan DW, Branch TA, Haught S *et al.* Multi-decadal shifts in the distribution and timing of Pacific herring (*Clupea pallasii*) spawning in Prince William Sound. *Can J Fish Aquat Sci* 2021;**78**:1611–27. https://doi.org/10.1139/cjfas-2021-0047

Mildenberger TK, Berg CW, Kokkalis A *et al.* Implementing the precautionary approach into fisheries management: making the case for probability-based harvest control rules. *bioRxiv* 2020.

Mohn RK, Chouinard GA. Harvest control rules for stocks displaying dynamic production regimes. *ICES J Mar Sci* 2007;**64**:693–7. https://doi.org/10.1093/icesjms/fsm042

Monnahan CC, Branch TA, Thorson JT *et al.* Overcoming long Bayesian run times in integrated fisheries stock assessments. *ICES J Mar Sci* 2019;**76**:1477–88. https://doi.org/10.1093/icesjms/fsz059

Monnahan CC, Kristensen K. No-U-turn sampling for fast Bayesian inference in ADMB and TMB: introducing the adnuts and tmbstan R packages. *PLoS One* 2018;**13**:e0197954. https://doi.org/10.1371/journal.pone.0197954

Monnahan CC, Thorson JT, Branch TA. Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods Ecol Evol* 2017;**8**:339–48. https://doi.org/10.1111/2041-210X.12681

Morella J. *Prince William Sound Science Center*. Cordova, Alaska: ADF&G Pacific herring fishery monitoring, 2022.

Morstad S, Sharp D, Wilcock J *et al. Prince William Sound Management Area 1995 Annual Finfish Management Report. Alaska Department of Fish and Game, Division of Commercial Fisheries. Regional Information Report 2A96-25*. Alaska: Anchorage, 1996.
Muradian ML, Branch TA, Moffitt SD *et al.* Bayesian stock assessment of Pacific herring in Prince William Sound. *PLoS One* 2017;**12**:e0172153. https://doi.org/10.1371/journal.pone.0172153
Muradian ML, Branch TA, Punt AE. A framework for assessing which sampling programmes provide the best trade-off between accuracy and cost of data in stock assessments. *ICES J Mar Sci* 2019;**76**:2102–13. https://doi.org/10.1093/icesjms/fsz163
Pearson WH, Deriso RB, Elston RA *et al.* Hypotheses concerning the decline and poor recovery of Pacific herring in Prince William Sound. *Rev Fish Biol Fisheries* 2012;**22**:95–135. https://doi.org/10.1007/s11160-011-9225-7
Pegau WS. 2022 Prince William Sound forage fish observations. Prince William Sound Regional Citizens Advisory Council, 2022, 29pp. https://www.pwsrcac.org/wp-content/uploads/900.431.221128.PegauForageRpt.pdf
Pikitch E, Boersma P, Boyd IL *et al.* Little fish, big impact: managing a crucial link in ocean food webs. *Lenfest Ocean Prog* 2012;**108**:16–30
Plaganyi E, Deng R, Campbell R *et al.* Evaluating an empirical harvest control rule for the Torres Strait *Panulirus ornatus* tropical rock lobster fishery. *Bull Mar Sci* 2018;**94**:1095–120. https://doi.org/10.5343/bms.2017.1101
Polacheck T, Klaer NL, Millar C *et al.* An initial evaluation of management strategies for the southern bluefin tuna fishery. *ICES J Mar Sci* 1999;**56**:811–26. https://doi.org/10.1006/jmsc.1999.0554
Prager MH, Porch CE, Shertzer KW *et al.* Targets and limits for management of fisheries: a simple probability-based approach. *North Am J Fisheries Manag* 2003;**23**:349–61. https://afspubs.onlinelibrary.wiley.com/doi/abs/10.1577/1548-8675%282003%29023%3C0349%3ATALFMO%3E2.0.CO%3B2
Privitera-Johnson KM, Punt AE. Leveraging scientific uncertainty in fisheries management for estimating among-assessment variation in overfishing limits. *ICES J Mar Sci* 2020;**77**:515–26. https://doi.org/10.1093/icesjms/fsz237
Punt AE, A'mar T, Bond NA *et al.* Fisheries management under climate and environmental uncertainty: control rules and performance simulation. *ICES J Mar Sci* 2014;**71**:2208–20. https://doi.org/10.1093/icesjms/fst057
Punt AE, Butterworth DS, de Moor CL *et al.* Management strategy evaluation: best practices. *Fish Fisheries* 2016;**17**:303–34. https://doi.org/10.1111/faf.12104
Punt AE, Cui G, Smith ADM. Defining robust harvest strategies, performance indicators and monitoring strategies for the SEF. Canberra, Australia: Australian Fisheries Management Authority, 2001. https://www.frdc.com.au/sites/default/files/products/1998-102-DLD.pdf
Punt AE, Donovan GP. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. *ICES J Mar Sci* 2007;**64**:603–12. https://doi.org/10.1093/icesjms/fsm035
Punt AE, Smith ADM. Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solandri*). *ICES J Mar Sci* 1999;**56**:860–75. https://doi.org/10.1006/jmsc.1999.0538
Quinn T, Marty G, Wilcock J *et al.* Disease and population assessment of Pacific herring in Prince William Sound, Alaska. In: Funk F, Blackburn J, Hay D, Paul AJ, Stephenson R, Toreson R, Witherell D (eds.), *Herring Expectations for a new Millennium*. Fairbanks, Anchorage, AK: University of Alaska Sea Grant, 2001, 363–79.
Quinn TJ, Fagen R, Zheng J. Threshold management policies for exploited populations. *Can J Fish AquatSci* 1990;**47**:2016–29. https://doi.org/10.1139/f90-226
Rademeyer RA, Plagányi ÉE, Butterworth DS. Tips and tricks in designing management procedures. *ICES J Mar Sci* 2007;**64**:618–25. https://doi.org/10.1093/icesjms/fsm050
Rodionov SN. A sequential algorithm for testing climate regime shifts. *Geophys Res Lett* 2004;**31**:19448. https://doi.org/10.1029/2004GL019448
Roel BA, De Oliveira JAA. Harvest control rules for the Western horse mackerel (*Trachurus trachurus*) stock given paucity of fishery-independent data. *ICES J Mar Sci* 2007;**64**:661–70. https://doi.org/10.1093/icesjms/fsm016
Schindler DE, Hilborn R, Chasco B *et al.* Population diversity and the portfolio effect in an exploited species. *Nature* 2010;**465**:609–12. https://doi.org/10.1038/nature09060
Shannon CE, Weaver W. *A Mathematical Theory of Communication*. Urbana: University of Illinois Press, 1949.
Shertzer K, Prager M, Williams E. Probabilistic approaches to setting acceptable biological catch and annual catch targets for multiple years: reconciling methodology with national standards guidelines. *Mar Coastal Fisheries* 2010;**2**:451–8. https://doi.org/10.1577/C10-014.1
Siple MC, Koehn LE, Johnson KF *et al.* Considerations for management strategy evaluation for small pelagic fishes. *Fish Fisheries* 2021;**22**:1167–86. https://doi.org/10.1111/faf.12579
Subbey S, Devine JA, Schaarschmidt U *et al.* Modelling and forecasting stock–recruitment: current and future perspectives. *ICES J Mar Sci* 2014;**71**:2307–22. https://doi.org/10.1093/icesjms/fsu148
Surma S, Pitcher TJ, Pakhomov EA. Trade-offs and uncertainties in Northeast Pacific herring fisheries: ecosystem modelling and management strategy evaluation. *ICES J Mar Sci* 2021;**78**:2280–97. https://doi.org/10.1093/icesjms/fsab125
Szuwalski CS, Britten GL, Licandeo R *et al.* Global forage fish recruitment dynamics: a comparison of methods, time-variation, and reverse causality. *Fish Res* 2019;**214**:56–64. https://doi.org/10.1016/j.fishres.2019.01.007
Szuwalski CS, Punt AE. Fisheries management for regime-based ecosystems: a management strategy evaluation for the snow crab fishery in the eastern Bering Sea. *ICES J Mar Sci* 2013;**70**:955–67. https://doi.org/10.1093/icesjms/fss182
Tommasi D, Stock CA, Pegion K *et al.* Improved management of small pelagic fisheries through seasonal climate prediction. *Ecol Appl* 2017;**27**:378–88. https://doi.org/10.1002/eap.1458
Trochta JT, Branch TA. Applying Bayesian model selection to determine ecological covariates for recruitment and natural mortality in stock assessment. *ICES J Mar Sci* 2021;**78**:2875–94. https://doi.org/10.1093/icesjms/fsab165
Trochta JT, Groner ML, Hershberger PK *et al.* A novel approach for directly incorporating disease into fish stock assessment: a case study with seroprevalence data. *Can J Fish AquatSci* 2022;**79**:1–20.
Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science* 1981;**211**:453–8.
Xia M, Carruthers T, Kindong R *et al.* How can information contribute to management? Value of information (VOI) analysis on Indian Ocean striped marlin (*Kajikia audax*). *Front Mar Sci* 2021;**8**:646174. https://doi.org/10.3389/fmars.2021.646174
Zahner JA, Branch TA. *Prince William Sound BASA: 2023 Model*, v2022 edn. Genève: Zenodo, 2023.

Appendix 1: Reduced Bayesian sampling routine

Best practices in Bayesian modeling often call for running multiple concurrent model chains for many thousands of samples, with additional checks for autocorrelation between samples within each model chain. These practices, if followed, result in high computational overhead in the form of model runtime. This high runtime often makes running large number of Bayesian models infeasible. This study ultimately ran 45 000 Bayesian models and, thus, required models to run quickly, while still yielding accurate estimates of important derived quantities.

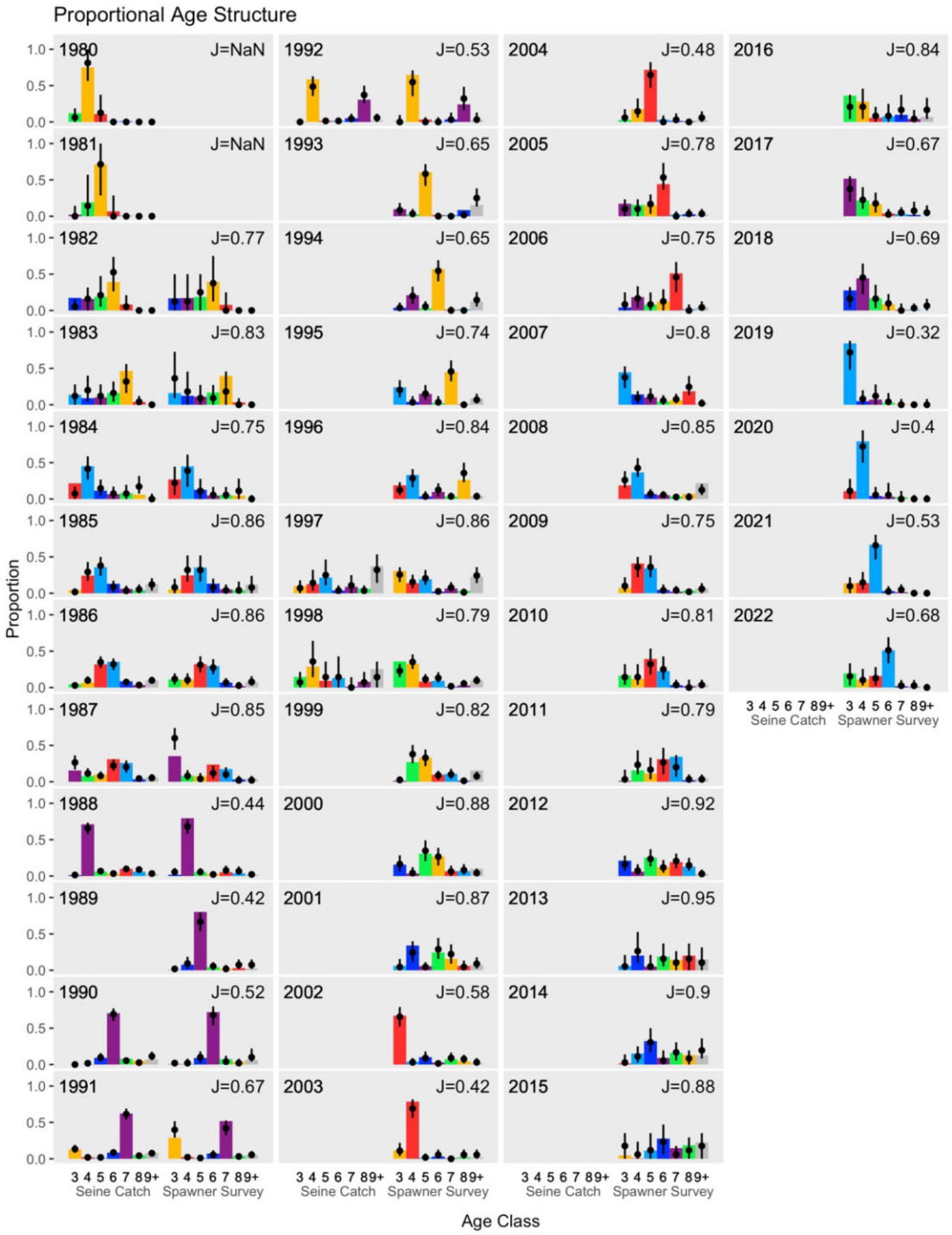

Figure A1. Historical age compositions and corresponding evenness (J) values for each year. Evenness values calculated from the age composition from the ASL survey for each year. Each color bar follows a single cohort through the fishery.

An external analysis performed using the Bayesian Age Structured Assessment (BASA) model for PWS herring with data through 2022 found the median value of the critical derived quantity (pre-fishery spawning biomass) to be consistently estimated for models run using a variable number of model chains and chain lengths (Table A1). Model convergence was also not substantially different across different numbers of chains or chain lengths. As such, we chose to run the Bayesian estimation model with 1 chain and 1000 samples, allowing for all model runs to be completed in 7 days, and accepting that the median estimate of pre-fishery biomass may be slightly imprecise (Table A1). We believe that variability in the OM and across OM simulations ultimately swamps slight imprecision in the model parameter estimates caused by using a reduced number of chains and samples, and that, thus, the results presented in this study are robust to the exact Bayesian sampling scheme used.

Table A1. Final year biomass, MCMC diagnostics (Divergences, Min ESS, and Max Rhat), single model run time (in seconds), and total MSE runtime (in days) for different numbers of MCMC chains and different chain lengths.

N chains	N samples	Final year biomass	Divergences	Min ESS	Max Rhat	Time (s)	Total time (dy)
1	**1,000**	**20,386.25**	**0**	**196**	**1.0239**	**53.60**	**1.395**
1	2,000	20,100.95	0	549	1.0111	102.70	2.674
1	5,000	20,125.95	0	253	1.0039	242.89	6.325
1	10,000	20,170.90	0	1988	1.0012	461.24	12.011
2	1,000	20,178.85	0	703	1.0075	62.20	3.239
2	2,000	20,143.50	0	1215	1.0043	112.80	5.875
2	5,000	20,070.20	0	1470	1.0022	263.39	13.718
2	10,000	20,153.40	0	4142	1.0008	514.72	26.808
4	1,000	20,161.00	0	1071	1.0057	67.48	7.029
4	2,000	20,172.50	0	2284	1.0037	122.92	12.804
4	5,000	20,145.35	0.0001	5913	1.0010	291.56	30.371
4	10,000	20,145.35	0.0001	5562	1.0007	571.43	59.524

Total time is the expected total duration of the MSE for 10 control rules each simulated 150 times for 30 years (for a total of 45 000 assessments) parallelized across 20 compute cores. The total time is an underestimate due to sampling time increasing as additional years of data are added. The bolded row indicates the sampling routine used in this study.

Table A2. Evenness values (J) computed using the Shannon-Weiner evenness formula for the PWS Pacific herring population since 1980.

Year	J	Year	J	Year	J
1980		1994	0.65	2008	0.85
1981		1995	0.74	2009	0.75
1982	0.77	1996	0.84	2010	0.81
1983	0.83	1997	0.86	2011	0.79
1984	0.75	1998	0.79	2012	0.92
1985	0.86	1999	0.82	2013	0.95
1986	0.85	2000	0.88	2014	0.90
1987	0.85	2001	0.87	2015	0.88
1988	0.44	2002	0.58	2016	0.84
1989	0.41	2003	0.42	2017	0.67
1990	0.52	2004	0.48	2018	0.69
1991	0.67	2005	0.78	2019	0.32
1992	0.53	2006	0.75	2020	0.40
1993	0.65	2007	0.80	2021	0.53

The age composition from the ASL survey was used as the age structure used for the calculation of J in each year (Zahner and Branch 2023).

Appendix 2: Historical age-structure evenness for Prince William Sound Pacific herring

Like many forage fish populations, the population of Pacific herring in PWS is characterized by unpredictable large recruitment events. Due to the relatively short life span of the species, this leads to the population age structure frequently being dominated by individuals from a single age class. The Evenness HCR attempts to take this directly into account by rescaling the allowable harvest rate based on the Shannon-Weiner definition of population evenness (Shannon and Weaver 1949), applied to annual age structure. Historically, the age structure of PWS Pacific herring has ranged widely, with many years being largely dominated by fish of a singular age and some years with nearly equal proportions of fish in each age class. The Shannon-Weiner evenness metric (J) has consequently ranged from J = 0.35 (2019) to J = 0.95 (2013) (Table A2, Fig. A1). The distribution of evenness values shows that the population tends towards more even age structures (large values of J), with over 30% of all years exhibiting an evenness value in the range of 0.80–0.90, and just 2 years (<5%) exhibiting evenness values <0.40.

Handling Editor: Pamela Woods