Impacts of time-varying productivity on estimated stock-recruitment parameters and biological reference points

| Journal:               | Canadian Journal of Fisheries and Aquatic Sciences |
|------------------------|-----------------------------------------------------|
| Manuscript ID:         | crfas-2019-0104.R1                                  |
| Manuscript Type:       | Article                                             |
| Date Submitted by the | 08-Nov-2019                                         |
| Author:                |                                                     |
| Complete List of Authors: | Holt, Carrie; Department of Fisheries and Oceans, Michielsens, Catherine; Pacific Salmon Commission |
| Keyword:               | time-varying productivity, biological reference point, Kalman filter, stock-recruitment model, SALMON < Organisms |
| Is the invited manuscript for consideration in a Special Issue?: | Not applicable (regular submission) |
Impact of time-varying productivity on estimated stock-recruitment parameters and biological reference points

Carrie A. Holt\textsuperscript{1} and Catherine G.J. Michielsens\textsuperscript{2}

\textsuperscript{1} Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC, Canada; School of Resource and Environmental Management, Simon Fraser University, Burnaby, BC, Canada. Carrie.Holt@dfo-mpo.gc.ca

\textsuperscript{2} Pacific Salmon Commission, 600-1155 Robson St, Vancouver, BC, Canada. michielsens@psc.org
Abstract

Models with time-varying parameters are increasingly being considered in the assessment of fish stocks, but their reliability when used to derive biological reference points or benchmarks has not been thoroughly evaluated. Here, we evaluated stock-recruitment models with and without time-varying productivity in a simulation framework for sockeye salmon (*Oncorhynchus nerka*) under different scenarios of productivity and exploitation. Ignoring trends in productivity led to overestimates of productivity and underestimates of capacity when both exploitation rates and productivity declined over time, resulting in an underestimation on average of benchmarks of biological status. Despite being less biased, time-varying models had relatively poor fit based on AICc and BIC model selection criteria. Our simulation results were compared with empirical analyses of 12 Fraser River sockeye salmon stocks in British Columbia, Canada. Although benchmarks were less biased based on time-varying models, underlying true benchmarks based on spawner abundances at maximum sustainable yield, $S_{\text{MSY}}$, trend downwards when productivity declines, which may not be aligned with conservation objectives. We conclude with best practices when adapting biological benchmarks to time-varying productivity.

Keywords: time-varying productivity, stock-recruitment model, benchmarks, biological reference points, Ricker model, Kalman filter, sockeye salmon, Fraser River, simulation
Introduction

Accounting for environmental variability and its effects on stock productivity is increasingly being recognized as a priority for fisheries management (King et al. 2015). Success incorporating environmental drivers into stock assessments has however been limited due to correlations between environmental variables and biological processes that break down over time or are spurious (e.g., DeOliveria et al. 2005) leading to incorrect inference, prediction or management recommendations.

Several approaches for adjusting biological reference points to account for variability in productivity have been developed and implemented (Pinsky and Mantua 2014; Punt et al. 2013; Szuwalski and Hollowed 2016). For multi-decadal scale variability, biological reference points can be estimated using only data after a shift has been detected (e.g., as applied to Snow Crab in the eastern Bering Sea, Szuwalski and Punt 2012), or using a “moving window” approach that selects recent periods of data for inclusion in assessments (Punt et al. 2013). Alternatively, time-varying parameters reflecting annual to multiyear-scale variability can be incorporated into assessment and management using more complex analytical approaches where parameters vary according to a random walk, or more specifically, using a Kalman filter (Schnute 1994; Peterman et al. 2000; Johnston 2013). The Kalman filter is a set of recursive formulae for identifying time-varying changes within a state-space model for linear Gaussian systems. That technique derives annual parameters recursively from the previous year’s estimate, while accounting for uncertainty in that estimate and model variances.

Peterman et al. (2000) first introduced the use of the Kalman filter as an analytical tool to identify time-varying productivity in Pacific salmon stocks using a Ricker stock-recruit relationship. For these species, productivity tends to show more change over time compared with capacity (Adkison et al. 1996) and stock-recruit models with time-varying productivity perform better than models with time-varying capacity (Dorner et al. 2008). Capacity or $S_{max}$ is defined here as spawner abundances that will maximize
adult recruitment (Myers 2001). Several studies have applied the Kalman filter to track time-varying productivity for Pacific salmon (Peterman et al. 2003; Dorner et al. 2008; Peterman and Dorner, 2012), and to revise biological reference points, or benchmarks of status (Grant et al. 2011; Grant and Pestal 2013). In studies where the appropriateness of including time-varying productivity was considered, statistical model selection criteria were used without examining the contrast in the underlying data or possible correlations between parameter estimates (Jiao et al. 2009; Britten et al. 2016; Szuwalski et al. 2019). Peterman et al. (2000) found that including time-varying parameters in stock-recruitment models for Pacific salmon reduced parameter bias and increased precision when the true underlying productivity varies over time. Their analysis considered cyclic patterns in productivity but not long-term declines, which is the dominant pattern for salmon across its southern range, including Fraser River sockeye salmon. In addition, they did not consider the impact of variability in spawner and recruitment data as well as data contrast on the relative performance of time-varying models, nor did they assess the implications of including time-varying productivity on assessment and management decisions. These factors represents a gap in our understanding of time-varying productivity models for recruitment parameter estimation and management applications.

The goal of this paper is to evaluate the impact of including time-varying productivity on parameter uncertainty and bias when deriving benchmarks. We addressed this overarching goal with the following objectives.

1. Evaluate the bias and precision of parameter estimates from a standard stock-recruit model against one that includes time-varying productivity using simulated stock-recruitment data based on different scenarios of temporal changes in productivity and exploitation rates.

2. Evaluate the ability of statistical model-fit criteria to identify the “true” model when choosing between standard and time-varying models.
(3) Evaluate the bias and precision of biological benchmarks derived from the estimated parameters of standard and time-varying models.

(4) Compare deviations in parameter estimates between standard and time-varying model forms derived from the simulated data and from observed data for Pacific salmon to identify potential risks of biases in benchmarks when using empirical data.

(5) Provide suggestions on best practices when considering time-varying productivity in stock assessments and when deriving benchmarks.

To address these objectives, we used a Monte Carlo simulation framework. Simulation modelling allowed us to evaluate the relative biases of standard stock-recruitment models that assume constant productivity against time-varying models for a wide range of different stock-recruit data with different underlying population dynamics. We parameterized the simulation model (objectives 1-3) with information for sockeye salmon, *Oncorhynchus nerka*, a species of high commercial, ecological, and socio-economic value. We fit both standard and time-varying models to real times-series of stock-recruitment data (objective 4) from 12 sockeye salmon stocks from the Fraser River, British Columbia, Canada. We focus on these stocks because of strong evidence for time-varying productivity (Peterman and Dorner 2012), and the recent application of stock-recruitment models that include time-varying productivity to derive benchmarks (Grant and Pestal 2013). In addition, we indicate how deviations in parameter estimates between the two model forms compare between observed stock-recruitment data for Fraser River sockeye salmon, and the simulation results. The simulation results were generalized in sensitivity analyses to evaluate impacts on a broader range of population characteristics for Pacific salmon, and the methodology can be further expanded to evaluate impacts on other marine fish species.
Methods

Stock-recruitment models and associated benchmarks

We used the Ricker stock-recruitment model with log-normal recruitment deviations because of theoretical support for this model (Peterman 1981) and its extensive use representing population dynamics of Pacific salmon (e.g., Dorner et al. 2008; Rogers and Schindler 2011; Fleischman et al. 2013) and other marine fishes (e.g., Szuwalski et al. 2019). The standard Ricker model assumes productivity remains unchanged over time and is formulated:

$$\log_e \left( \frac{R}{S} \right) = \alpha - \beta S_t + \nu_t,$$

where $S_t$ is the total number of spawners in the brood year $t$, $R_t$ is the number of adult recruits produced by those spawners, $\alpha$ is the productivity at low spawner abundance in the absence of density dependence, $\beta$ is the rate at which recruitment is reduced by density-dependence and the inverse of the abundance of spawners at maximum recruitment, $S_{max}$, or capacity, and $\nu_t$ are random normal deviations with variance $\sigma^2_\nu$, $\nu_t \sim N(0, \sigma^2_\nu)$.

We incorporated time-varying productivity into the Ricker model, by allowing the $\alpha$ parameter to vary over time $t$ according to a random walk:

$$\log_e \left( \frac{R}{S} \right) = \alpha_t - \beta S_t + \nu_t,$$

$$\alpha_t = \alpha_{t-1} + \omega_t,$$

where $\omega_t$ are random normal process errors with variance $\sigma^2_\omega$, $\omega \sim N(0, \sigma^2_\omega)$. Parameters for the standard Ricker model (Eqn. 1) were estimated using least-square minimization. Parameters of the time-varying Ricker model (Eqns. 2 and 3) were estimated using a recursive Kalman Filter algorithm with maximum likelihood estimation (Harvey 1989; Peterman et al. 2000; Britten et al. 2016). A fixed-interval...
smoother was applied to the time-varying estimates of productivity to reduce random annual variations that obscure underlying long-term trends (Peterman et al. 2000; Britten et al. 2016).

We then calculated two biological benchmarks from the estimated stock-recruitment parameters, \( S_{MSY} \) and \( S_{gen} \). \( S_{MSY} \) is the spawner abundance required to achieve maximum sustainable yield,

\[
(4) \quad S_{MSY} = \left( 1 - W(e^{1-a}) \right) / \beta,
\]

where \( W \) is the Lambert W function (Scheuerell 2016). \( S_{gen} \) is the number of spawners that would result in recovery to \( S_{MSY} \) in one generation in the absence of fishing. This latter benchmark is used under Canada’s Wild Salmon Policy (WSP) for status assessments (DFO 2005; Holt et al. 2009). We estimated time-varying benchmarks, \( S_{MSY,t} \) by using time-varying \( \alpha_t \) values and \( S_{gen,t} \) by optimizing the following formulation of the Ricker equation with time-varying \( \alpha_t \) values,

\[
(5) \quad S_{MSY,t} = S_{gen,t} e^{(\alpha_t - \beta S_{MSY,t})}.
\]

**Simulation framework**

The Monte Carlo simulation framework used to evaluate the stock-recruitment models consisted of a simple operating model that represented the spawner and recruitment dynamics and a management procedure that included harvest decisions. We evaluated parameter biases and precision under three productivity scenarios (constant, declining and increasing) and three exploitation rate scenarios (constant low or high exploitation and a stepwise decline in exploitation). Our simulation model was a simplified version of models previously developed to evaluate benchmarks for southern British Columbia chum salmon (Holt and Folkes 2015; Holt et al. 2018), which captures the general population and management dynamics for Pacific salmon stocks, as described below.
The operating model simulated a hypothetical population of sockeye salmon with a Ricker spawner recruitment relationship assuming either constant or time-varying productivity (Eqns. 1 and 2). Because the majority (>90%) of adult sockeye return to spawn at age-4 in the Fraser River, we simulated only a single age-at-return using the spawner-recruitment model. Previous simulations for Pacific salmon have shown that variability in age-at-maturity has relatively little impact on benchmark performance compared with variability of other biological and management variables (Holt et al. 2018), though these models did not include time-varying parameters. To simplify the simulation model, the population dynamics within the operating model used generational instead of annual time-steps.

For the management procedure within the simulation framework, we incorporated three different exploitation rate scenarios: 1) a constant low target exploitation rate of 20% (our base case scenario), 2) a constant 60% target exploitation rate and 3) a step-wise decline in target exploitation rate from 80% to 20% half-way through the time series. This last scenario reflects the dramatic declines in exploitation that occurred for Fraser River sockeye salmon in the 1990s (Supp. Mat. 1, Fig. S1). We adapted the exploitation scenarios to broadly represent harvest control rules applied to Fraser River sockeye salmon. These rules include a relatively complex combination of escapement goals and target exploitation rates such that those exploitation rates, $h_t$, are not achieved annually. To account for this, we included annual deviations in realized exploitation rates, $h'_t$, equal to a proportion of the normally distributed random deviations in recruitment rate ($v_t$ in Eqn. 1), so that interannual variation in exploitation rates were positively related to recruitment (Supp. Mat. 1, Fig. S2):

$$ h'_t = h_t + \rho \cdot v_t, $$

where the scalar $\rho$ is set to 0.1 so that the standard deviation in $h'_t$ ($0.1 \cdot \sigma_v = 0.06$) is approximately equal to the standard deviation in annual exploitation rates observed historically for Fraser River sockeye salmon (PSC 2018, see also Supp. Mat. 1, Fig. S2). We set lower and upper limits on realized
exploitation rates to 5 and 85%, respectively. Following the implementation of this exploitation rate, spawner abundances in the subsequent generation, $t+1$, were calculated as:

$$S_{t+1} = R_t \cdot (1 - h_t').$$

**Model parameterization**

To parameterise the operating model of the simulation framework, we selected productivity ($\alpha$) and exploitation rate values under the base case scenario that reflect the population dynamics of sockeye salmon in general, with additional scenarios reflecting the population dynamics for Fraser River sockeye salmon. We assumed a productivity of $1.5 \log_e(R/S)$ for the constant (base case) productivity scenario, which equals the mean productivity estimated previously for 37 sockeye salmon stocks across North America (Dorner et al. 2008). For the second productivity scenario, we modelled a linear decline in productivity from 2 to $1 \log_e(R/S)$ over the time frame of the simulation (60 years), the magnitude of which is within the range of declines observed for Fraser River sockeye salmon stocks (Peterman and Dorner 2012). In contrast, certain stocks such as the Harrison River, have seen an increase in productivity (Peterman and Dorner 2012) and this is reflected in the third productivity scenario that simulates a linear increase in productivity from 1.5 to $2.5 \log_e(R/S)$ thereby approximating the dynamics of that stock.

We assumed a fixed spawner capacity of 10 000 salmon for $S_{\text{MAX}}$, a standard deviation in Ricker residuals of 0.6, which is within the range of values estimated for stocks in British Columbia and Alaska (0.17-1.63, Korman et al. 1995; Peterman et al. 2003). We used 100 fish as an extirpation threshold below which populations were assumed lost. We did not include lag-1 year autocorrelation among recruitment residuals in our base case given the limited evidence among Fraser River sockeye stocks, and hence its absence in their assessment models (DFO 2017; Grant et al. 2011).
We initialized the simulation model with spawner abundances at 20% of the spawner abundances at unfished equilibrium, \( S_{\text{eq}} \), and ran the simulation model over 60 years, the length of time series available for most Fraser River sockeye salmon stocks. The simulation model was run over 1000 Monte Carlo trials to stabilize resulting model parameters and performance metrics.

We evaluated the sensitivity of our simulation results to variations in the spawner capacity, the magnitude and form of residual variance in the stock-recruitment model including the presence of lag-1 year autocorrelation and scalar linking residual variance in recruitment and realized exploitation rates, productivity, initial spawner abundances, and the number of years in the simulation within plausible bounds for Pacific salmon (See Supp. Mat. 2, Table S2 for details).

**Performance metrics**

We first calculated summary statistics to characterize the simulated stock-recruitment data generated for each scenario. In particular, we calculated the proportion of the time series at high spawner abundances (above \( S_{\text{MAX}} \)) within each Monte Carlo trial, and the proportion of those high spawner abundances that resulted in recruitment events > \( R_{\text{MAX}} \), the maximum equilibrium recruitment. These metrics reflect two characteristics of the distribution of simulated spawner and recruitment data that help explain divergent results in parameter biases among scenarios.

Then, two different types of statistical performance metrics were used to evaluate and compare the performance of the stock-recruitment models with and without time-varying productivity: those indicating bias and precision of parameters and benchmarks and those informing model selection, indicating the fit of the models given the data. Biases can only be calculated through simulations when the true parameter values are known while the second type, model selection criteria, is commonly used
to compare model fit against observed data without knowledge of true parameters (Peterman and Dorner 2012; Britten et al. 2016).

We evaluated bias and precision in productivity by comparing estimated productivity, $\hat{\alpha}_t$, against the true productivity $\alpha_t$ for both standard and time-varying models. When productivity varies over time, the recent performance metrics may be most useful for informing future application of these models. We therefore evaluated the bias (percent error, $PE$) in productivity of the last generation, or year ($t=60$) against the true productivity of that year for both models.

$$PE_{\alpha_{60}} = \frac{(\hat{\alpha}_{60} - \alpha_{60})}{\alpha_{60}} \times 100.$$  

We also calculated the mean percent error, $MPE$, over the 60-year simulation,

$$MPE_{\alpha} = \frac{\sum_{t=1}^{60} (\hat{\alpha}_t - \alpha_t)}{\alpha_t} \times 100.$$  

We further evaluated parameter biases by comparing the estimate of $S_{MAX}$ against the true value for $S_{MAX}$ for both the standard and time varying models and calculating the percent error for each Monte Carlo trial. The $MPE$ and $PE$ estimates across the 1000 Monte Carlo simulation trials were presented in the form of boxplots where the 95% confidence intervals indicated the precision of the $MPE$ and $PE$ estimates across trials.

From a fisheries management perspective, biases in derived benchmarks may be of greater relevance than biases in the stock-recruitment parameters. We first showed how true benchmarks, $S_{MSY}$ and $S_{gen}$, calculated from true productivity and $S_{MAX}$ vary with changes in those underlying parameters. We then estimated biases in the benchmarks $S_{MSY}$ and $S_{gen}$ both in the most recent year (as $PE$, Eqn. 8) and across the time series (as $MPE$, Eqn. 9) under various productivity and exploitation rate scenarios.
In addition to evaluating the bias and precision of the parameters and benchmark estimates for both models, we also compared model performance using standard statistical model selection criteria. In particular, we compared model fits of the standard and time-varying model using the Akaike Information Criterion for small sample sizes, AICc (Burnham and Anderson 1998). In the calculation of AICc, we assumed one additional parameter for the time-varying model relative to the standard model to account for the additional variance parameter, $\sigma_\omega^2$ (Britten et al. 2016) and the AICc was calculated using an analytical solution for the log-likelihood from the filtering algorithm (Meinhold and Singpurwalla 1983; Britten et al. 2016). In sensitivity analyses, an alternative model selection criterion, BIC, was also applied (as in Britten et al. 2016). Ideally, model preferences based on statistical model selection criteria should correspond with parameter biases and precision from simulation results, with models that best fit the data containing parameters that are least biased and most precise.

**Data**

To illustrate the relevance of the simulation results for real-world applications, we applied the two stock-recruitment models to 12 stocks of sockeye salmon from the Fraser River watershed in British Columbia, Canada (DFO 2017) and compared the parameter estimates between models. We further evaluated historical trends in productivity and exploitation for each stock. Spawner numbers for Fraser River sockeye are expressed in terms of effective total spawners and are the product of total spawner numbers and the proportion of successfully spawned eggs based on spawning-ground carcass surveys. Effective total spawners are hereafter referred to as spawners. Recruitment estimates are based on the sum of the escapement, catch, and losses due to in-river mortality prior to spawning (Grant et al. 2011). We excluded Fraser River stocks with strongly cyclic patterns (Cass and Wood 1994), which are typically analyzed with more complex stock-recruitment models.
The 12 stocks included in this paper differed substantially in average spawning abundance (ranging from thousands to hundreds of thousands of spawners), current status (ranging from relatively depleted to healthy), number of years of stock-recruitment data (38-61 years), and quality of stock-recruitment data (high precision fence counts to lower precision visual surveys).

Results

Simulation Results (Objectives 1-3)

When we fit stock-recruitment models to simulated spawner and recruitment data for each Monte Carlo trial, the resulting parameter estimates varied due to the distribution of true underlying data. Figure 1 illustrates the impact of data at high spawner abundance on the models’ abilities to account for density dependence and produce reliable estimate of \( S_{\text{MAX}} \), for two Monte Carlo trials under two different scenarios. Under the first scenario (Fig. 1a), the simulated stock-recruit data contain a limited number of years with high spawner abundance above \( S_{\text{MAX}} \), but a relatively high portion of these years resulted in recruitment above \( R_{\text{MAX}} \) (75%). When fitting a standard Ricker model to these data (Fig 1a, red curve), the model underestimated the strength of density dependence (overestimated \( S_{\text{MAX}} \)) and underestimated productivity. The observed negative relationship between productivity and \( S_{\text{MAX}} \) is consistent with the well-documented correlation between these parameters (Walters and Martell 2006). Applying the time-varying Ricker model to the same data, the model identified the declining trend in productivity and correctly accounted for density dependence when estimating \( S_{\text{MAX}} \). Under the second scenario, the proportion of years with spawner abundances above \( S_{\text{MAX}} \) was similarly low, but only a small proportion of large spawning events resulted in recruitment above \( R_{\text{MAX}} \) (33%, Fig. 1b). The high spawner abundances occurred primarily in the later years when both exploitation and productivity were low. When fitting the standard Ricker model, this pattern of spawner and recruitment data was associated with the appearance of relatively strong density dependence in the form of...
overcompensation at high spawner abundances (Fig. 1b, red curve). As a result, the standard model tended to underestimate $S_{\text{MAX}}$ and overestimate productivity. These biases were reduced, but not eliminated for the time-varying model which could, to some extent track the temporal trends.

Summarized over 1000 Monte Carlo trials, the time-varying model tended to result in less biased parameter estimates compared with the standard model, especially under declining productivity trends (Fig. 2, middle column). As expected, productivity in the most recent year was overestimated by the standard model when underlying productivity declined (Fig. 2a, top middle column). Although less biased, these productivity estimates were relatively imprecise for the time-varying model compared to the standard model (Fig. 2a, top middle column). The most severe overestimate of $S_{\text{MAX}}$ (median PE = 73%) occurred for the standard model under the scenario of constant high exploitation (60%) with declining productivity (Fig. 2c, middle boxplots). Under constant productivity, biases in parameters tended to be relatively small and similar between models in the last year and across the time series (Fig. 2, left column). Under the scenario of increasing productivity (Fig. 2, right column), both the standard and the time-varying models substantially underestimated productivity.

The direction of biases in productivity and $S_{\text{MAX}}$ estimates were similar among the sensitivity analyses we considered (Supp. Mat. 2, Figs. S6-S18). In particular, biases for the standard model were greatest when variances in recruitment residuals were low, the scalar between recruitment deviations and exploitation rate deviations was high, and lag 1-year autocorrelation in recruitment was included. Differences between the standard and time-varying models were minimal for populations with very low spawner capacity, as these populations were more likely to become extirpated. Time-series length, initial abundances, and level of true constant underlying productivity had little impact on results for the ranges considered here.
While the simulation allowed the assessment of model performance in terms of bias of stock-recruitment parameters, this information is not available when analyzing empirical data. In that case statistical model criteria such as AICc are commonly used to evaluate model performance. Model fits based on AICc were relatively poor for the time-varying model compared with the standard model across scenarios (Fig. 3), despite providing less biased parameters estimates. The alternative model selection criterion, BIC, showed even less support for the time-varying model because of increased penalty given to the additional variance term in that model (Supp. Mat. 1 Fig. S3). Overall, AICc and BIC provided an indicator of how well models describe data, but did not reflect the relative accuracy of parameter estimates.

**Derived benchmark estimates**

Prior to evaluating the impact of biases in parameter estimates on the benchmarks, we show the impact of changes in productivity on both $S_{MSY}$ and $S_{gen}$ (Fig. 4). When underlying productivity declines but density-dependence ($S_{MAX}$) remains constant, the value for $S_{MSY}$ declines and $S_{gen}$ increases (Fig. 4 moving left on the X-axis). Since productivity is defined by the slope of the stock-recruitment curve near the origin, as productivity declines, the stock-recruitment curve shifts closer to the 1:1 equilibrium line. As a result, fewer spawners are needed to maximize yield ($S_{MSY}$ declines) and the maximum sustainable yield declines. However, to recover to $S_{MSY}$ within one generation (the definition of $S_{gen}$), many more spawners are needed despite the modest decline in $S_{MSY}$, and so $S_{gen}$ increases towards $S_{MSY}$ (see Holt and Folkes 2015 for further details). In contrast, as $S_{MAX}$ increases, both benchmarks increase in value (Fig. 4, moving up the Y-axis). These patterns are expected following the equations for $S_{MSY}$ (Eqn. 4) and $S_{gen}$ (Eqn. 5).

As true productivity declined (or increased) in our simulation, benchmark estimates varied both because of those underlying changes (Fig. 4, shifts to the left or right, respectively) and due to associated biases.
in productivity and $S_{\text{MAX}}$, with the direction of those biases depending on the scenario. We illustrated the impact of biases in parameter estimates on benchmark estimates using the same two scenarios presented in Figure 1. Under the scenario of declining productivity and constant high exploitation, the overestimation of $S_{\text{MAX}}$ (Fig. 1a, Fig. 2c) by the standard model caused both benchmarks to be overestimated in the most recent year relative to the true benchmark (Fig. 4, orange dots labelled “standard” relative to black dots; Fig. 5, orange vertical lines relative to the black lines). Under the scenario of simultaneous declines in productivity and exploitation, the standard model underestimated $S_{\text{MAX}}$ (Fig. 1b, Fig. 2c) and overestimated productivity on average over the time series (Fig. 1b, Fig. 2b), resulting in underestimates of $S_{\text{gen}}$ but not $S_{\text{MSY}}$ in the most recent year (Fig. 4, green dots labelled “standard” relative to black dots), and underestimates of both benchmarks on average over the time series (Fig. 5, green vertical lines relative to the black lines). In contrast, for the time-varying model, slight underestimation of $S_{\text{MAX}}$ resulted in relatively small biases in benchmark values under both scenarios (Fig. 4, green dots labelled “Time-varying” relative to black dots).

Overall, benchmarks derived from the time-varying model tended to be less biased than those from the standard model, both in the recent year and on average over the time series (Fig. 6). For the standard model, biases in benchmarks were especially large under declines in productivity. As shown above, under declining productivity and constant high exploitation rates, the standard model overestimated $S_{\text{MAX}}$, resulting in large overestimates of benchmarks (Fig. 6, middle column). In contrast, under constant productivity, benchmarks were negatively biased for both models, especially under high exploitation rates (Fig. 6, left column) consistent with the underestimates of $S_{\text{MAX}}$ by both models (Fig. 2). Under increasing productivity, the direction of the bias depended on the benchmark. On average, stocks that experienced limited exploitation and constant productivity had stock recruitment data that contained more information to reliably estimate benchmarks using standard methods.
Empirical analyses (Objective 4)

Using data for Fraser River sockeye salmon, we evaluated historical trends in productivity and exploitation rate, and then compared estimates of productivity and benchmarks between the two models. Fitting time-varying stock-recruitment models we identified declines in productivity for 7 stocks while 4 stocks showed limited or no change in productivity data (Supp. Mat. 1, Table S1, Fig. S4). Only one stock, Harrison, showed an increase in productivity over the time series. All stocks experienced similar steep declines in exploitation in the 1990s, being harvested in mixed-stock fisheries (Supp. 1, Fig. S1).

Although it is not possible to identify true underlying parameters for Fraser River sockeye stocks, we can compare the relative difference in parameter estimates between standard and time-varying models for the simulation and empirical results to identify potential risks of estimation biases in benchmarks given known biases in the simulation results (Fig. 7). We found that for the 7 stocks where estimated productivity and exploitation rates had declined, the relative difference between parameters of the time-varying versus the standard model for the empirical data were the same as for simulated data when evaluated over the entire time series (Fig. 7) and for the most recent year (Supp. Mat 1, Fig. S5). If the underlying dynamics causing differences in parameter estimates are similar between the simulated and the empirical data, then this convergence highlights possible risk of estimation biases in empirical benchmarks derived from the standard model. In particular, for these stocks, estimates of productivity from the standard model were higher than those for the time-varying model averaged over the time series, and benchmarks were lower (Fig. 7, middle column), possibly due to negative estimation biases as found in the simulation results (Fig. 6c and d). Despite the compelling evidence from our simulations in favor of the time-varying model based on parameter biases, the AICc model selection criteria favored the standard model for 71% of the Fraser River sockeye stocks included in this paper (Supp. Mat. 1,
Table S1). For stocks where productivity was relatively constant or increasing, the empirical deviations in parameters between model forms also tended to converge with those from the simulation (Fig. 7).

**Discussion**

We found an association between the distribution of recruitment data at high spawner abundances and parameter biases in stock-recruitment relationships under various productivity and exploitation scenarios. We demonstrated how stock-recruitment models can underestimate productivity and underestimate the strength of density dependence (overestimate $S_{\text{MAX}}$) and vice versa, depending on the information content in the underlying data. Information content was described here as the proportion of years with spawner abundances above $S_{\text{MAX}}$ and proportions of those high spawner events that result in recruitment above $R_{\text{MAX}}$. However, without knowledge of the true underlying parameter values, it is not possible to identify parameter biases solely based on the distribution of spawner and recruitment abundances across their observed ranges. For example, a low proportion of recruits above $R_{\text{MAX}}$ among high spawner events may indicate declines in productivity under declining exploitation or overcompensation, i.e. strong density dependence at high abundances, as documented previously for some sockeye salmon stocks (Peterman and Dorner 2012).

Overall we found that the combination of historical trends in productivity and exploitation history impacts the distribution of spawner and recruitment data, which in turn impacts biases in estimated stock-recruitment parameters and benchmark values. Changes in exploitation rates over time can interact with temporal trends in productivity to cause unintuitive effects on parameter biases and precision. Our results indicated that stock-recruitment models with time-varying productivity tended to be less biased and more precise than standard models regardless of historical exploitation patterns.
improvements in parameter estimates were, however, greater for scenarios of declining productivity than increasing productivity because of the associated distribution of spawner and recruitment data. In this study, we simulated a limited set of trends in productivity and exploitation to drive population dynamics, which determined the resulting spawner and recruitment data and subsequent biases in parameter estimates. In reality, parameter estimates are driven by complex interactions between the distribution of the spawner and recruitment data, the length of time series, observation errors in spawner abundances, and time-series biases. In particular, time-series biases occur when spawner abundances in one generation are correlated with recruitment abundances in previous generations, biasing parameter estimates (Walters and Martell 2004), which may be accentuated with persistent trends in underlying parameters. In addition, our model assumed constant age-at-maturity and further extensions of this work could incorporate variability in age-at-maturity and time-varying trends in that parameter.

Model selection criteria have in general been the main method promoted in the literature to evaluate if time-varying productivity should be considered for particular stocks. In a global meta-analyses of fish stocks, Britten et al. (2016) used BIC to identify that 79% of the 262 stocks were best fit by a Ricker model with time-varying productivity or capacity, and using AICc, Szuwalski et al. (2019) found about half of 52 forage fish stocks were best fit with time-varying Ricker models (time-varying productivity or capacity). However, our results suggest caution when using information criteria to inform model choice for time-varying models. AICc and BIC tended to favour standard over time-varying models despite relatively large biases in parameter estimates, including under scenarios of declining productivity and reduced exploitation, which are common trends among many salmon populations (Dorner et al. 2008).

While our general results favoring time-varying models were in line with previous studies that evaluated time-varying productivity, our study greatly extended that work. Peterman et al. (2000) found that the
time-varying model parameters were less biased and more precise than those of the standard model and that performance depended on the temporal patterns in underlying true productivity. However, our work differed from theirs in four significant ways. First, we generalized the simulation scenarios to include both trends in exploitation as well as productivity, which was warranted given results differed substantially across exploitation scenarios. Second, we compared long-term declines and increases in productivity over the 60-year simulation period instead of cyclic patterns at a multi-decadal scale. Third, we compared model selections using standard information criteria against an approach using parameter biases derived from simulation, demonstrating challenges when relying on information criteria. And finally, we also evaluated the bias and precision of benchmarks derived from standard and time-varying models.

Implications for benchmarks

Observed declines in productivity have important implications for benchmark or biological reference point derivation due to changes in true underlying benchmarks and associated biases in benchmark estimates. For example, declines in true $S_{\text{MSY}}$ with productivity may increase conservation risks if this benchmark is used to inform harvest decisions since it will be exceeded at lower abundances compared to the $S_{\text{MSY}}$ value averaged over the time series that ignores time-varying productivity. In contrast, the increase in $S_{\text{gen}}$ under true declines in productivity can be considered more precautionary from a conservation perspective, since higher abundances are required to exceed that benchmark.

In addition to the impacts of time-varying productivity on true benchmarks, there are also associated impacts on benchmark estimates. Our results showed that under temporal trends in productivity, the direction and magnitude of biases in benchmark estimates depend on the exploitation history and direction of productivity trends. In general, a time-varying model may be preferred as it tended to provide less-biased parameter estimates than the standard model, especially under declining...
productivity. However, temporal changes in underlying true benchmarks can interact with biases in benchmark estimates by reducing them when productivity declines, as shown for the standard model estimates of $S_{\text{MSY}}$, or increasing them as for $S_{\text{gen}}$.

The choice of whether to adopt time-varying models for benchmark estimates to account for recent declines in productivity (i.e., by adopting lower $S_{\text{MSY}}$ and higher $S_{\text{gen}}$ values) will depend on the choice of benchmark, the underlying population dynamics, and how they are used to inform management decisions. If recent declines in productivity are due to reductions in marine survival, for example, associated with decadal-scale variability in oceanographic processes (Beamish et al. 1997), then more spawners may be needed to produce the same number of recruits compared with historical high-productivity regimes, and a precautionary response may be to adopt a time-varying $S_{\text{gen}}$ ($S_{\text{gen}}$ estimated with time-varying productivity) as a benchmark to inform management decisions since its value increases under declines in productivity. In contrast, if declines in productivity are irreversible, for example due to construction of dams that reduce the availability of freshwater rearing grounds, then it may be reasonable to adopt time-varying benchmarks related to $S_{\text{MSY}}$ to inform harvest objectives since its value declines with productivity.

Even in cases where accounting for time-varying productivity in benchmarks is deemed appropriate, large changes in benchmarks between successive years (or successive assessments), may not be palatable to harvesters or decision makers. Instead, it may be more effective to consider temporal variability in productivity by making management decisions robust to uncertainty in productivity (e.g., by implementing relatively low harvest targets if uncertainty in productivity is high), than by regularly revising benchmarks. In this case, time-varying models could be used annually to track changes in productivity (e.g., Dorner et al. 2008; Peterman and Dorner 2012), and only update benchmarks accordingly once sufficient evidence has accumulated to indicate a persistent shift.
When considering the impact of time-varying productivity on benchmarks, truncating data to the most recent productivity regime has been commonly suggested as an alternate approach (Grant et al. 2011; Punt et al. 2013) or to an acceptable reference period representing pristine on minimally exploited conditions. However, estimates of productivity are often negatively correlated with capacity, and spawner and recruitment data are typically not sufficiently informative to separately estimate these parameters especially when times-series are relatively short. Our results suggest that biases in $S_{\text{MAX}}$ in particular can result in significant biases in benchmark estimates, and therefore we recommend caution using this approach. Although regime-based stock-recruitment models that use the entire time series are a plausible alternative, rigorous simulation testing of these models including methods for detecting regime shifts is currently lacking (but see Szuwalski et al. 2019).

Overall, we recommend evaluating the impact of incorporating time-varying productivity on benchmark estimates within management strategy evaluations, MSE, before using them to inform management decisions. Like the simulation models used here, MSEs include simulations that incorporate impacts of harvest decisions on population dynamics with closed-loop feedback. MSEs, however, typically evaluate performance of various management procedures against quantified management objectives and incorporate more detailed specifications of the management decisions and assessment methods than considered here, often with the input of stakeholders, analysts, and decision makers (Punt et al. 2016).

**Best practices (Objective 5)**

To determine if time-varying productivity should be accounted for when estimating stock-recruitment parameters and deriving benchmark estimates, we suggest using simulation-evaluation as demonstrated here. For stocks whose population dynamics are adequately captured by the model presented in this paper (including sensitivity analyses), the parameter and benchmark biases estimated here can be applied according to observed trends in productivity and exploitation rates.
More broadly, we recommend several best practices when considering revisions to benchmarks or biological reference points due to persistent shifts in productivity to ensure decision making is sound and transparent (adapted from Duplisea and Cardigan (2012)):

- Document evidence of the degree of and change in exploitation of the stock using time series of exploitation rates, catch per unit effort and/or catches.
- Document evidence for the change in productivity by examining trends in residual variation of recruits per spawner and/or employing time-varying models such as the Kalman filter or regime detection algorithms (e.g., Sequential t-test Analysis of Regime Shifts, STARS, Rodionov 2004)
- Document underlying mechanism(s) that could be causing the change in productivity, using for example ecosystem status reports (Harvey et al. 2018; Grant et al. 2019) and include all critical habitats, which for anadromous species includes both the ocean and freshwater.
- Demonstrate that the change in the state of the productivity will persist long enough in comparison to the management plan, so that changing benchmarks will indeed respond to management needs. This step requires an understanding of the mechanism driving those changes and adequate forecasts of these variables into the future, which is lacking for many species.
- Calculate stock-recruitment estimates and benchmarks with and without time-varying productivity and compare values.
- Estimate the uncertainty in benchmarks by accounting for the underlying uncertainty in stock-recruitment parameters, and communicate the associated risks for assessment and management.
- Support decisions for benchmarks or reference points with simulation models that include management procedures with the assessment and application of those reference points (e.g., in
harvest control rules) and uncertainty in future trends in productivity (as shown for Pacific
salmon by Collie et al. (2012)).

By following these best practices, justification for changing benchmarks will be grounded in the strength
of evidence for temporal trends in productivity and/or changes in capacity with a transparent
understanding of the impacts of those changes on the ability to achieve management objectives. In the
absence of information on plausible drivers, identifying benchmarks or reference points that are robust
to plausible changes in productivity is favoured over the application of benchmarks that vary annually.

Our model focused on temporal trends in productivity and exploitation that have occurred for Fraser
River sockeye salmon, though these trends have also been found across other stocks and salmon species
and so results may be more broadly applicable (Peterman and Dorner 2012; Malick and Cox 2016;
Dorner et al. 2018). While our analyses focused on only two benchmarks commonly used to assess
status of Pacific salmon, the closed-loop simulation approach could be readily adapted to accommodate
other benchmarks, as well as other productivity trends and exploitation histories.

Observed changes in productivity among Pacific salmon stocks have prompted a need to account for
these changes in assessments and management. Given significant climate-driven changes in productivity
cost wide, standard models that ignore these changes will provide biased estimates of benchmarks,
and the direction and magnitude of the bias depends on exploitation history, direction of productivity
trends, and the choice of benchmark. Diagnostic metrics on the differences between parameter
estimates of the standard and time-varying models and on the distribution of spawner and recruitment
data can indicate the likely direction of biases in the standard model, and are more useful for
determining possible biases than traditional model selection criteria. However, analysts should carefully
weigh the impacts of using biased benchmark estimates that ignore time-varying productivity against
impacts of annually revising benchmarks, ideally within a quantitative simulation evaluation that
includes specific measurable objectives.
Acknowledgements

The authors are grateful to Catarina Wor, Ann-Marie Huang, Sue Grant, Brendan Connors, Kendra Holt, and Brooke Davis for their helpful comments on the work as well as the paper, and to Mike Lapointe for continuous feedback on the research topic. Four anonymous reviewers were especially helpful in refocusing an earlier version of the manuscript.

References

Adkison M.D., Peterman R.M., Lapointe M.F, Gillis D.M and Korman J. 1996. Alternative models of climatic effects on sockeye salmon productivity in Bristol Bay, Alaska and the Fraser River, British Columbia. Fish. Oceanogr. 5: 137-152.

Beamish R.J., Neville C.M. and Cass A.J. 1997. Production of Fraser River sockeye salmon (*Oncorhynchus nerka*) in relation to decadal-scale changes in the climate and ocean. Can. J. Fish. Aquat. Sci. 54:543-554.

Britten G.L., Dowd M. and Worm B. 2016. Changing recruitment capacity in global fish stocks. Proc. Natl. Acad. Sci. 113: 134-139.

Burnham, K.P., and Anderson, D.R. 1998. Model selection and inference: a practical information theoretic approach. Springer, New York.

Cass A. J. and Wood C.C. 1994. Evaluation of the depensatory fishing hypothesis as an explanation for population-cycles in Fraser River sockeye salmon (*Oncorhynchus nerka*). Can. J. Fish. Aquat. Sci. 51: 1839-1854.
Collie J.S., Peterman R.M. and Zuehlke B.M. 2012. A fisheries risk-assessment framework to evaluate trade-offs among management options in the presence of time-varying productivity. Can. J. Fish. Aquat. Sci. 69: 209-223.

De Oliveira J.A.A., Uriarte A. and Roel B.A. 2005. Potential improvements in the management of Bay of Biscay anchovy by incorporating environmental indices as recruitment predictors. Fish. Res. 75:2-14.

DFO (Department of Fisheries and Oceans). 2005. Canada’s Policy for the conservation of wild Pacific salmon. DFO, Vancouver. DFO (Department of Fisheries and Oceans). 2016. Proceedings of the Pacific regional peer review on the Assessment of Southern British Columbia Chinook Salmon Conservation Units, Benchmarks and Status; February 4-6, 2014. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2016/029.

DFO (Department of Fisheries and Oceans). 2017. Pre-season run size forecasts for Fraser River sockeye (Oncorhynchus nerka) and Pink (O. gorbuscha) salmon in 2017. DFO Can. Sci. Advis. Sec. Sci. Resp. 2017/016.

Dorner, B., M.J. Catalano, and Peterman, R.M. 2018. Spatial and temporal patterns of covariation in productivity of Chinook salmon populations of the northeastern Pacific Ocean. Can. J. Fish. Aquat. Sci. 75:1082-1095.

Dorner B., Peterman R.M. and Haeseker S.L. 2008. Historical trends in productivity of 120 Pacific pink, chum and sockeye salmon stocks reconstructed by using a Kalman filter. Can. J. Fish. Aquat. Sci. 65: 1842-1866.

Duplisea D. and Cadigan N. 2012. Proceedings of the national workshop for technical expertise in stock assessment (TESA): maximum sustainable yield (MSY) reference points and the precautionary approach when productivity varies. Can. Sci. Adv. Sec. Proc. Ser. 2012/055.
Fleischman, S. J., M. J. Catalano, R. A. Clark, and D. R. Bernard. 2013. An age-structured state-space stock–recruit model for Pacific salmon (Oncorhynchus spp.). Can. J. Fish. Aquat. Sci. 70:401-414

Grant S.C.H., MacDonald B.L., Cone T.E., Holt C.A., Cass A., Porszt E.J., Hume J.M.B. and Pon L.B. 2011. Evaluation of uncertainty in Fraser sockeye WSP Status using abundance and trends in abundance metrics. Can. Sci. Advis. Sec. Res. Doc 2011/087. VIII + 183 p.

Grant, S.C.H. and Pestal G. 2013. Integrated biological status assessment under the Wild Salmon Policy using standardized metrics and expert judgement: Fraser River sockeye salmon (Oncorhynchus nerka) case studies. Can. Sci. Advis. Sec. Res. Doc. 2012/106. v + 132 p.

Grant, S.C.H., MacDonald, B.L., and Winston, M.L. 2019. State of the Canadian Pacific Salmon: Responses to Changing Climate and Habitats. Can. Tech. Rep. Fish. Aquat. Sci. 3332: ix + 50 pp.

Harvey, C., N. Garfield, G. Williams, N. Tolimieri, I. Schroeder, E. Hazen, K. Andrews, K. Barnas, S. Bograd, R. Brodeur, B. Burke, J. Cope, L. deWitt, J. Field, J. Fisher, T. Good, C. Greene, D. Holland, M. Hunsicker, M. Jacox, S. Kasperski, S. Kim, A. Leising, S. Melin, C. Morgan, B. Muhling, S. Munsch, K. Norman, W. Peterson, M. Poe, J. Samhouri, W. Sydeman, J. Thayer, A. Thompson, D. Tommasi, A. Varney, B. Wells, T. Williams, J. Zamon, D. Lawson, S. Anderson, J. Gao, M. Litzow, S. McClatchie, E. Ward, and S. Zador. 2018. Ecosystem Status Report of the California Current for 2018: A Summary of Ecosystem Indicators Compiled by the California Current Integrated Ecosystem Assessment Team (CCEIA). U.S. Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-145. https://doi.org/10.25923/mvhf-yk36

Harvey, A. C. 1989. Forecasting, structural time series models, and the Kalman filter. Cambridge University Press, Cambridge, U.K.

Ho Y.C. and Lee R.C.K. 1964. A Bayesian approach to problems in stochastic estimation and control. IEEE Trans. Autom. Control, 9: 333-339.
Holt C., Cass A., Holtby B. and Riddell B. 2009. Indicators of status and benchmarks for conservation units in Canada’s Wild Salmon Policy. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/058. VIII + 74p.

Holt, C.A., Davis, B., Dobson, D., Godbout, L., Luedke, W., Tadey, J., and Van Will, P. 2018. Evaluating Benchmarks of Biological Status for Data-limited Conservation Units of Pacific Salmon, Focusing on Chum Salmon in Southern BC. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/011. IX + 77 p.

Holt C.A. and Folkes M.J.P. 2015. Cautions on using percentile-based benchmarks of status for data-limited populations of Pacific salmon under persistent trends in productivity and uncertain outcomes from harvest management. Fish. Res. 171:188-200.

Jiao Y., Reid K., and Smith E. 2009. Model selection uncertainty and Bayesian model averaging in fisheries recruitment modeling. In: Beamish R.J. and Rothschild B.J. (Eds.) The future of fisheries science in North America. Fish and Fisheries Book Series No. 31, Springer, New York. pp. 505–524.

Johnston, N.T. 2013. Management reference points for the Thompson and Chilcotin late summer-run steelhead (Oncorhynchus mykiss) stock aggregates. BC Fish and Wildlife Branch, Fisheries Project Report RD139, Victoria, BC.

King J.R., McFarlane G.A. and Punt A.E. 2015. Shifts in fisheries management: adapting to regime shifts. Phil. Trans. Roy. Soc. B-Biol. Sci. 370 (art. 20130277).

Malick M.J. and Cox S.P. 2016. Regional-scale declines in Productivity of Pink and Chum Salmon stocks in western North America. PLoS ONE 11(1): e0146009. doi:10.1371/journal.pone.0146009

Meinhold, R.J., and Singpurwalla, N.D. 1983. Understanding the Kalman filter. Am. Stat. 37: 124–127.

Myers R.A. 2001. Stock and recruitment: generalizations about maximum productivity rate, density dependence, and variability using meta-analytic approaches. ICES, J. Mar. Sci. 58: 937-951.
PSC, Pacific Salmon Commission, 2018. Report of the Fraser River Panel to the Pacific Salmon Commission on the 2017 Fraser River Sockeye Salmon Fishing Season. Vancouver, BC.

Peterman, R. M. 1981. Form of random variation in salmon smolt-to-adult relations and its influence on production estimates. Can. J. Fish. Aquat. Sci. 38:1113-1119.

Peterman R.M. and Dorner B. 2012. A widespread decrease in productivity of sockeye salmon (Oncorhynchus nerka) populations in western North America. Can. J. Fish. Aquat. Sci. 69: 1255-1260.

Peterman R.M., Pyper B.J. and Grout J.A. 2000. Comparison of parameter estimation methods for detecting climate-induced changes in productivity of Pacific salmon (Oncorhynchus spp.). Can. J. Fish. Aquat. Sci. 57: 181-191.

Peterman R.M., Pyper B.J. and MacGregor B.W. 2003. Use of the Kalman filter to reconstruct historical trends in productivity of Bristol Bay sockeye salmon (Oncorhynchus nerka). Can. J. Fish. Aquat. Sci. 60: 809-824.

Pinsky M.L. and Mantua N.J. 2014. Emerging Adaptation Approaches for Climate-Ready Fisheries Management. Oceanography 27: 146-159.

Punt A.E., A'mar T., Bond N.A., Butterworth D.S., de Moor C.L., De Oliveira J.A.A., Haltuch M.A., Hollowed A.B. and Szuwalski C. 2013. Fisheries management under climate and environmental uncertainty: control rules and performance simulation. ICES J. Mar. Sci.71: 2208-2220.

Punt, A. E., D. S. Butterworth, C. L. Moor, J. A. A. De Oliveira, and M. Haddon. 2016. Management strategy evaluation: best practices. Fish and Fisheries 17:303-334.

Rodionov, S. N. 2004. A sequential algorithm for testing climate regime shifts. Geophys. Res. Let. 31:5. L09204.
Rogers, L. A., and D. E. Schindler. 2011. Scale and the detection of climatic influences on the productivity of salmon populations. Global Change Biology 17:2546-2558.

Schnute J.T. 1994. A general framework for developing sequential fisheries models. Can. J. Fish. Aquat. Sci. 51:1676-1688.

Scheuerell, M. D. 2016. An explicit solution for calculating optimum spawning stock size from Ricker’s stock recruitment model. PeerJ 4:e1623-e1623.

Szuwalski C. S. and Hollowed A. B. 2016. Climate change and non-stationary population processes in fisheries management. ICES J. Mar. Sci. 73: 1297-1305.

Szuwalski C.S. and Punt A.E. 2012. Fisheries management for regime-based ecosystems: a management strategy evaluation for the snow crab fishery in the eastern Bering Sea. ICES J. Mar. Sci. 70: 955-967.

Szuwalski, C. S., G. L. Britten, R. Licandeo, R. O. Amoroso, R. Hilborn, and C. Walters. 2019. Global forage fish recruitment dynamics: A comparison of methods, time-variation, and reverse causality. Fish. Res. 214:56-64.

Walters, C.J. and Martell, S.J.D. 2004. Fisheries ecology and management. Princeton University Press, Princeton.
Figures

Figure 1. Example stock-recruitment data for two Monte Carlo trials under scenarios of (a) declining productivity and constant high exploitation rate (60%) and (b) declining productivity and exploitation rates. Circles represent spawner-recruitment data and are shaded darker towards the end of the time series. For each of the top two panels, the solid red curve is the standard Ricker model, blue curves are those based on the time-varying Ricker model and are similarly shaded from light to dark as time progresses, and the grey curves are the true underlying time-varying relationships shaded also from light to black. The vertical dashed line represents the true S_{MAX}, the spawner abundances that would result in maximum recruitment and pS_{max} is the proportion of observations above S_{MAX}. The horizontal dashed line is the median true maximum recruitment, R_{MAX} and pR_{max} is the proportion of high spawning.
events with recruitment > $R_{\text{MAX}}$ (in the upper right quadrant). For the time-varying curves, only every 5th curve is shown. The bottom panels depict the scenarios for (a) of declining true productivity and constant exploitation, and (b) of declining true productivity and step decline in exploitation.
Figure 2. Percent error and mean percent error, MPE, in estimated productivity (a,b) and $S_{\text{MAX}}$ (c), from standard (white) and time-varying (blue) stock-recruitment models under various scenarios of...
underlying true productivity and exploitation rate ($h_t$) trends, labelled at the top. Errors in productivity are estimated for the most recent year (a) and averaged over the time series (b). Boxes represent the upper and lower bounds of the interquartile range with the midline at the median. Upper and lower whiskers approximate upper and lower 95% confidence intervals. $pS_{max}$ is the median proportion of spawner abundances $> S_{MAX}$ among Monte Carlo trials and $pR_{max}$ is the median proportion of recruitments $> R_{MAX}$ among those high spawning events. The horizontal lines under pairs of boxplots identify scenarios where the standard and time-varying models had median errors that differed by $\geq 5\%$ but $<10\%$ (thin lines) and $\geq 10\%$ (thick lines).
Figure 3. Difference in AICc between models under various scenarios of productivity and exploitation rates ($h_t$) where $\Delta$AICc is the AICc of the standard minus the AICc of the time-varying stock-recruitment model. Boxes represent the upper and lower bounds of the interquartile range of $\Delta$AICc, with the midline at the median. Upper and lower whiskers are approximate 95% confidence intervals.
Figure 4. Contours of (a) $S_{MSY}$ and (b) $S_{gen}$ along gradients in productivity and $S_{max}$. The black dots represent the true median benchmark for the most recent year under a scenario of declining productivity. Estimated benchmarks from the standard and time-varying models are denoted with “Standard” and “Time-varying” respectively, and were median values over all Monte Carlo trials for the most recent year, generated for scenarios of declining productivity while exploitation remained consistently high (orange) or declined stepwise (green).
Figure 5. Spawner-recruitment curves and associated benchmarks for the standard model estimated from spawner-recruitment data generated under a scenario of declines in productivity and constant exploitation (orange shaded area, 95% CI of Monte Carlo predictions), and a scenario of simultaneous declines in productivity and exploitation (green shaded area, 95% CI). The shaded curves were derived from distributions of stock-recruitment parameters over Monte Carlo trials (mean estimate of productivity over the time-series and $S_{\text{MAX}}$). The black curve is the underlying true stock-recruitment relationship based on the true mean productivity. Vertical lines represent the estimated values of $S_{\text{MSY}}$ (solid lines) and $S_{\text{gen}}$ (dotted lines) with colours aligned by scenario. The length of the horizontal lines above the benchmark estimates represent the 95% confidence intervals.
Figure 6. Percent error in benchmarks $S_{MSY}$ (a,c) and $S_{gen}$ (b,d), derived from standard and time-varying stock-recruitment models, relative to the true benchmark values under various scenarios of underlying true productivity and exploitation rates, $h_t$. Percent error, $PE$, in the most recent year (a,b) and mean percent error, $MPE$, across the time series (c,d) are shown. Boxes represent the upper and lower bounds of the interquartile range with the midline at the median. Upper and lower whiskers are the approximate 95% confidence intervals. The horizontal lines under pairs of boxplots identify scenarios where the standard and time-varying models had median errors that differed by $\geq 5\%$ but $<10\%$ (thin lines) and $\geq 10\%$ (thick lines).
Figure 7. Differences in estimated productivity, $\alpha$ (a), $S_{MSY}$ (b), and $S_{gen}$ (c) between the standard (S) and time-varying (T) stock-recruitment models averaged over the time series under various scenarios of true
underlying productivity trends and exploitation histories. Both simulation results (blue) and empirical results for Fraser River sockeye salmon (white) are shown. Given the declines in exploitation for Fraser River sockeye salmon, the empirical results can be compared against the simulation results based on data generated under the same exploitation scenario. Boxes represent the upper and lower bounds of the interquartile range with the midline at the median. Upper and lower whiskers are approximate 95% confidence intervals. When parameter estimates for both models converged, the boxes collapsed into a single line at zero. With only 1 Fraser River stock exhibiting increasing trends in productivity, results are similarly collapsed into a single line. Asterisks represent upper whiskers above the upper limit of the plot at *= 174%, and **= 357%.