Evaluating the consequences of common assumptions in run reconstructions on Pacific-salmon biological status assessments

Stephanie J. Peacock[1,*], Eric Hertz[2], Carrie A. Holt[3],

Brendan Connors[4], Cameron Freshwater[5], and Katrina Connors[6]

1. 8625 33 Ave NW, Calgary, AB T3B 1M2. Email: stephanie.j.peacock@gmail.com
2. Pacific Salmon Foundation, 300 – 1682 West 7th Avenue Vancouver, BC V6J 4S6. Email: ehertz@psf.ca
3. Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC V9T 6N7. Email: Carrie.Holt@dfo-mpo.gc.ca
4. Fisheries and Oceans Canada, Institute of Ocean Sciences, 9860 West Saanich Road, Sidney, BC V8L 4B2. Email: Brendan.Connors@dfo-mpo.gc.ca
5. Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC V9T 6N7. Email: Cameron.Freshwater@dfo-mpo.gc.ca
6. Pacific Salmon Foundation, 300 – 1682 West 7th Avenue Vancouver, BC V6J 4S6. Email: kconnors@psf.ca

* Corresponding author. Phone: (778) 266-0575

Abstract

Information on biological status is essential for designing, implementing, and evaluating management strategies and recovery plans for threatened or exploited species. However, the data required to quantify status are often limited, and it is important to understand how assessments of status may be biased by assumptions in data analysis. For Pacific salmon, biological status assessments based on spawner abundances and spawner-recruitment (SR) analyses often involve "run reconstructions" that impute missing spawner data, expand observed spawner abundance to account for unmonitored streams, assign catch to individual stocks, and quantify age-at-return. Using a stochastic simulation approach, we quantified how common assumptions in run reconstructions biased assessments of biological status based on spawner abundance. We found that status assessments were robust to most common assumptions in run reconstructions, even in the face of declining monitoring coverage, but that overestimating catch tended to increase rates of status misclassification. Our results lend confidence to biological status assessments based on spawner abundances and SR analyses, even in the face of incomplete data.

Introduction

Timely and effective management of fish and wildlife relies on accurate information about the current biological status of populations. Status assessments form the basis of many conservation decision-support tools, though they have limited influence without clear and binding linkages to policy and decision making. Methods to determine the allocation of resources that maximize biodiversity conservation, such as Priority Threat Management (Martin et al. 2018; Carwardine et al. 2019; Walsh et al. 2020), are most effective when baseline information on status can help identify which threatened populations are at highest risk of extirpation without management intervention, and which are most likely to benefit from recovery planning given limited resources. Metrics of biological status may also be useful for defining recovery goals and as performance measures when evaluating the effectiveness of different management strategies (e.g., Walsh et al. 2020). Knowledge of current biological status is therefore needed for both implementing and evaluating management strategies and recovery plans.

Despite the recognized need for information on biological status, complete data on abundance, trends, and demographic parameters are rarely available for populations or species that need to be assessed (IUCN 2017). Improving the analytical approaches used to quantify status is one way to reduce assumptions and limit the bias in status outcomes from low quality or quantity data (e.g., Fleischman et al. 2013; Staton et al. 2020), but the methods widely applied in assessing fish and wildlife can be slow to change (Peterman 2018) and analytical

resources are sometimes not available to implement new advanced approaches on all populations or species of interest. Thus, evaluating the impact of assumptions in the approaches that are widely applied is critical to understanding the potential biases, uncertainty, and limitations of status assessments (Chen et al. 2003; Wetzel and Punt 2011).

Closed-loop simulation models are useful tools for evaluating the potential biases in status assessments due to various ecological and management uncertainties (Walters 1986). The major advantages of simulation approaches (over, for example, retrospective analyses) are that the model can be tailored to different biological and management scenarios, allowing management strategy evaluation (e.g., Punt 1992; Holt and Peterman 2006, 2008), and the true status of the system is known perfectly, allowing quantification of bias in estimated parameters (e.g., Pyper and Peterman 1998; Dorner et al. 2009) and estimated status (e.g., Peacock and Holt 2012; Holt and Folkes 2015; Holt et al. 2018). Simulation approaches are extremely flexible and enable testing common assumptions that underpin status assessments at broad spatial scales in the face imperfect information, and have been applied extensively in fisheries research. In this study, we apply a simulation approach to understanding potential biases in biological status assessments of Pacific salmon introduced by assumptions in the analysis of spawner and recruitment data.

Assessing biological status of Pacific salmon

Pacific salmon (*Oncorhynchus* spp.) are a highly exploited group of species and many populations have experienced declines in recent decades due to overfishing, changing ocean conditions, and freshwater habitat degradation (e.g., COSEWIC 2016, 2017; Brown et al. 2019).

Pacific salmon are anadromous and semelparous, returning from ocean rearing grounds to spawn in freshwater before dying, and are typically vulnerable to fisheries upon their return to coastal waters. The data required to assess biological status of Pacific salmon can include annual estimates of the number of returning adult salmon to individual rivers, fisheries catch or harvest rates, and the age composition of returning salmon. Often, these data are incomplete and require imputation. "Run reconstructions" have been undertaken for many salmon stocks to expand spawner abundances to account for unmonitored streams and estimate the abundance of recruits to coastal fisheries prior to spawning (i.e., recruitment; Cave and Gazey 1994; English et al. 2007, 2016, 2018). The exact procedure undertaken depends on the life-history traits and available data for each population, and can include complexities such as spatial and temporal variability in returns among spawning populations. Status assessments that rely on run reconstructions have been adopted by, for example, local management organisations, the Marine Stewardship Council (www.msc.org), the Pacific Salmon Treaty (PSC 2019), COSEWIC (e.g., COSEWIC 2016, 2017), and the Pacific Salmon Foundation (Connors et al. 2013, 2018, 2019).

Fisheries and Oceans Canada (DFO) also relies on run reconstructions when assessing the biological status of Conservation Units (CUs) – groups of wild salmon that, if lost, are unlikely to recolonize within an acceptable timeframe – under Canada's Wild Salmon Policy (WSP; Fisheries and Oceans Canada 2005). Under the WSP's biological status assessment framework, quantifiable metrics are calculated from available data and compared against biological reference points, or "benchmarks", to arrive at a status outcome of red, amber, or

green (Fisheries and Oceans Canada 2005; Figure 1). A red status indicates that a CU has low spawner abundance and/or reduced spatial distribution and management intervention is required to avoid extirpation. A green status indicates that the CU is able to sustain maximum annual catch under existing environmental conditions. The specific benchmarks delineating these status zones must consider uncertainties in metrics and the unique biological characteristics (e.g., age composition) of the CU being assessed. Indeed, in their "integrated status assessments", DFO engages experts to consider the unique context of each CU and the quality and quantity of data that may be used to estimate status (e.g., DFO 2015, 2016, 2018a). However, the required resources and time (typically 1-3 years) have meant that integrated status assessments have been completed for only 9% of CUs since the WSP was released nearly 15 years ago, and reports are often 2-4 years out of date when they are released (DFO 2019).

Data-driven approaches focused on calculating quantitative metrics without expert elicitation can provide timely status updates that are standardized and comparable across CUs using transparent and repeatable methodology. For example, the Pacific Salmon Foundation (PSF) has undertaken a widespread effort to apply a data-driven approach to assessing spawner abundance under the WSP framework, with results for the north and central coast openly available through their Pacific Salmon Explorer (PSE) – www.salmonexplorer.ca (Connors et al. 2013, 2018, 2019), and are currently expanding their assessments to all remaining salmon CUs in British Columbia (BC). Despite the benefits of this approach, multiple assumptions in run reconstructions (Table 1) mean that potential biases can compound, and the impact on status outcomes needs to be quantified to lend confidence and credibility to data-driven assessments.

In this study, we quantified bias in biological status outcomes for Pacific salmon from data-driven assessments due to assumptions and uncertainty in run reconstructions (Table 1). We used a simulation approach that allowed us to explore how bias changed under different biological and management conditions. As a case study, we tailored our simulation model to represent a generic chum salmon CU from the central coast of BC because these CUs have recently been assessed using a data-driven approach (Connors et al. 2018) and have a relatively simple run-reconstruction model that does not include run timing. Furthermore, there are conservation concerns for both north and central coast chum salmon, which have not recovered despite significant reductions in harvest rates over the past two decades (DFO 2018b). Thus, central coast chum salmon offer a useful case study for an initial investigation of basic assumptions underlying biological status assessments. However, our simulation model is flexible enough to accommodate different species and life-history traits (e.g., age-at-return, density dependence) of Pacific salmon. We further explore a broad range of biological (e.g., trends in capacity, interannual variability in age-at-return) and management (e.g., declining monitoring coverage, bias in observed spawners and catch) scenarios to yield more general insight into the circumstances under which assumptions may bias status assessments.

Methods

Benchmark calculations and assumptions

Multiple metrics of biological status have been proposed under the WSP that cover four broad classes of indicators: current spawner abundance, trends in abundance over time, spatial

distribution of spawners, and fishing mortality (Holt et al. 2009). We focused our analysis only on the current spawner abundance metric and considered two types of benchmarks (Figure 1,2) that have been frequently applied to determine biological status of Pacific salmon CUs, including in the Pacific Salmon Explorer (PSE; Connors et al. 2013, 2018, 2019). The first type of benchmark is associated with maximum sustainable yield, derived from a SR relationship (Figure 2b). An upper SR benchmark of 80% of the spawner abundance that is projected to maintain long-term maximum yield, or 80% S_{MSY}, has been recommended by Holt et al. (2009, 2018) and will be applied to future assessments in the PSE (previous PSE assessments have applied S_{MSY}; Connors et al. 2018, 2019). S_{MSY} can be calculated explicitly from the productivity and density-dependence parameters of the Ricker SR relationship (Scheuerell 2016). Multiple lower benchmarks have been suggested (Holt et al. 2009, 2018), and we used a lower benchmark of the spawner abundance that leads to S_{MSY} in one generation in the absence of fishing mortality, called S_{GEN} (Korman and English 2013; DFO 2015), as applied in the PSE. These benchmarks delineate the green, amber and red status zones: If the geometric mean spawner abundance over the most recent generation (S_{AVG}) is above the upper benchmark of 80% S_{MSY}, then the CU is assigned a green status, if S_{AVG} is between 80% S_{MSY} and S_{GEN} , then the CU is assigned an amber status, and if S_{AVG} is less than S_{GEN}, then the CU is assigned a red status.

The second type of benchmarks we considered were percentile benchmarks (Clark et al. 2014; Holt and Folkes 2015; Holt et al. 2018), also called historic spawners benchmarks (e.g., Connors et al. 2018). The upper and lower benchmarks are the 50th and 25th percentiles of historical spawner abundance, respectively (Figure 2a). Calculating these benchmarks requires

fewer data, as the SR relationship need not be estimated (i.e. age-at-return and harvest data are not required). As such, the percentile benchmarks can be applied to data-limited CUs for which spawner data are patchy or age and harvest data are not available.

We focused on a basic run-reconstruction model and associated assumptions that are commonly made when assessing spawner abundance against the benchmarks above (Figure 2c; Table 1). At a minimum, in order to apply percentile benchmarks, a time series of total spawner abundance at the CU scale is required. Conservation Units are typically composed of multiple spawning populations that may or may not be monitored in any given year. Spawning populations in individual streams (henceforth "populations") may exhibit unique dynamics as their productivity is (in part) limited by density-dependent processes in freshwater. A simple sum of the observed spawner abundance within a CU may be misleading if the same populations are not monitored consistently. On BC's north and central coast, monitored populations have been designated as either "indicator streams" or "non-indicator streams", with indicator streams being prioritized for monitoring and thus having more continuous and reliable spawner estimates (English 2016). In addition, there may be populations that have never been monitored and for which spawner abundance is unknown. To reconstruct spawner abundance to the CU, three "Expansion Factors" have been applied to account for (1) spawners returning to indicator streams that are not monitored in a given year, (2) spawners returning to non-indicator streams, and (3) observation efficiency and populations that are never monitored (Table 1).

The application of SR benchmarks also requires time series of the total number of salmon returning to the CU to reconstruct recruitment, including those caught in fisheries and those that make it to spawn but are not counted. The number of returning salmon in a CU that are caught in fisheries is estimated based on the catch statistics for Pacific Fisheries Management Areas (PFMAs) adjacent to the geographic location of the CU (Figure 3). It is assumed that salmon caught in a PFMA were destined to spawn in streams that empty into that PFMA, although there is the potential for bias in that fish may be caught while migrating through the PFMA or fish destined for streams in the focal PFMA may be caught in other PFMAs. Furthermore, in most cases, there is not a perfect spatial correspondence between PFMAs and CUs (Figure 3). Streams in multiple CUs may flow into a single PFMA, which is common for small CUs, such as with sockeye salmon which spawn in individual lakes, each of which is associated with a unique CU (Holtby and Ciruna 2007). In the simplest case, the catch from that PFMA may be assigned to CUs based on the relative spawner abundance to each CU. However, differences in run-timing among CUs may complicate the assignment of catch and necessitate more complex run-reconstruction models. A single CU may also be composed of populations that are caught in multiple PFMAs, particularly for species with large CUs such as pink and chum salmon whose geographic boundary often spans hundreds of kilometers and includes dozens of spawning populations (Figure 3). In these cases, an average harvest rate across PFMAs may be applied which may not reflect the variable harvest rates experienced by the different populations within these large CUs. The impact of observation bias in the catch assigned to each CU on status assessments is unknown and is a focal aspect of this study (see Sensitivity analyses, below).

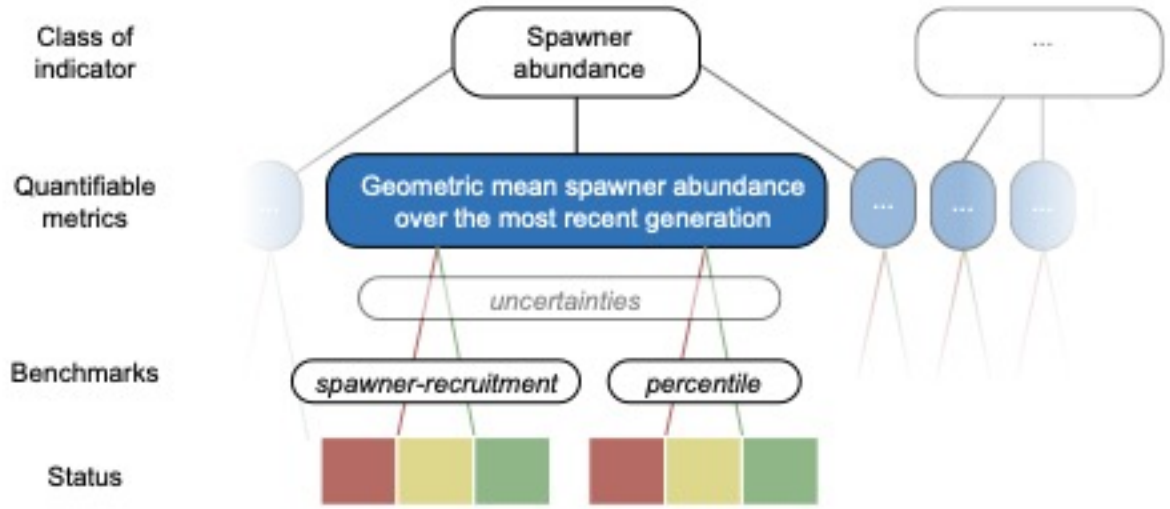

Figure 1. Illustration of the WSP status assessment framework (adapted from Holt et al. 2009). We focused on the geometric mean spawner abundance (metric, blue) under the spawner abundance indicator. This metric was assessed against two types of benchmarks: spawner-recruitment (SR) and percentile (see Figure 2). Faded boxes represent other types of metrics and indicators that may be included in integrated status assessments but were beyond the scope of what we considered.

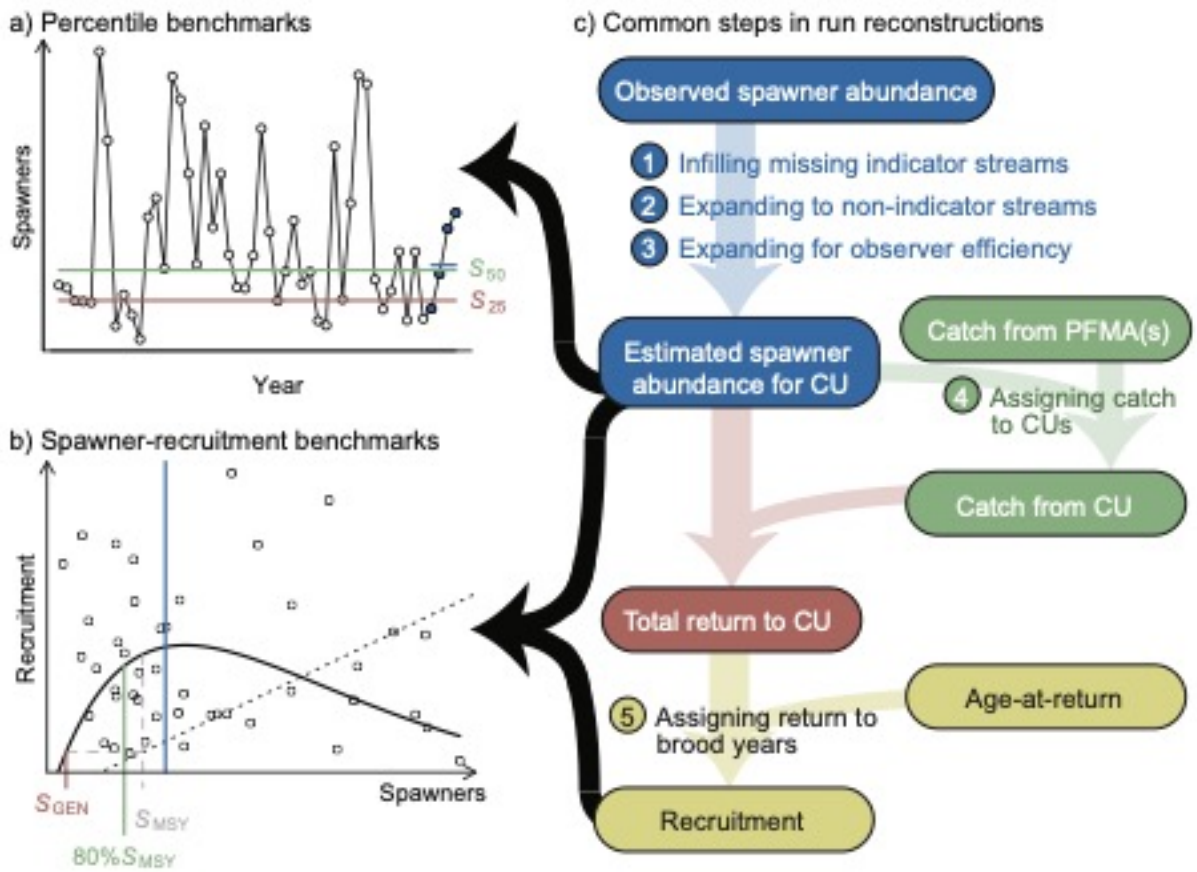

Figure 2. a) benchmarks are the 50th (horizontal green line) and 25th (horizontal red line) percentiles of historical spawner abundance (points). The current spawner abundance is calculated as the geometric mean spawner abundance over the most recent generation (4 years, blue points and line). b) -recruitment benchmarks are based on the shape of the Ricker relationship (solid line) fit to data on spawner abundance (x-axis) and corresponding recruitment (catch + escapement, y-axis). The upper and lower benchmarks are 80% S_{MSY} (green) and S_{GEN} (red), respectively. S_{GEN} is defined as the spawner abundance that leads to S_{MSY} (grey) in one generation in the absence of fishing mortality. Under both types of benchmarks, the current spawner abundance in the example shown is above the upper benchmark, and this CU would be assessed as 'green'. c) The calculation of benchmarks and benchmarks requires run reconstruction to expand observed spawners abundances, assign catch to CUs, and calculate recruitment (Table 1).

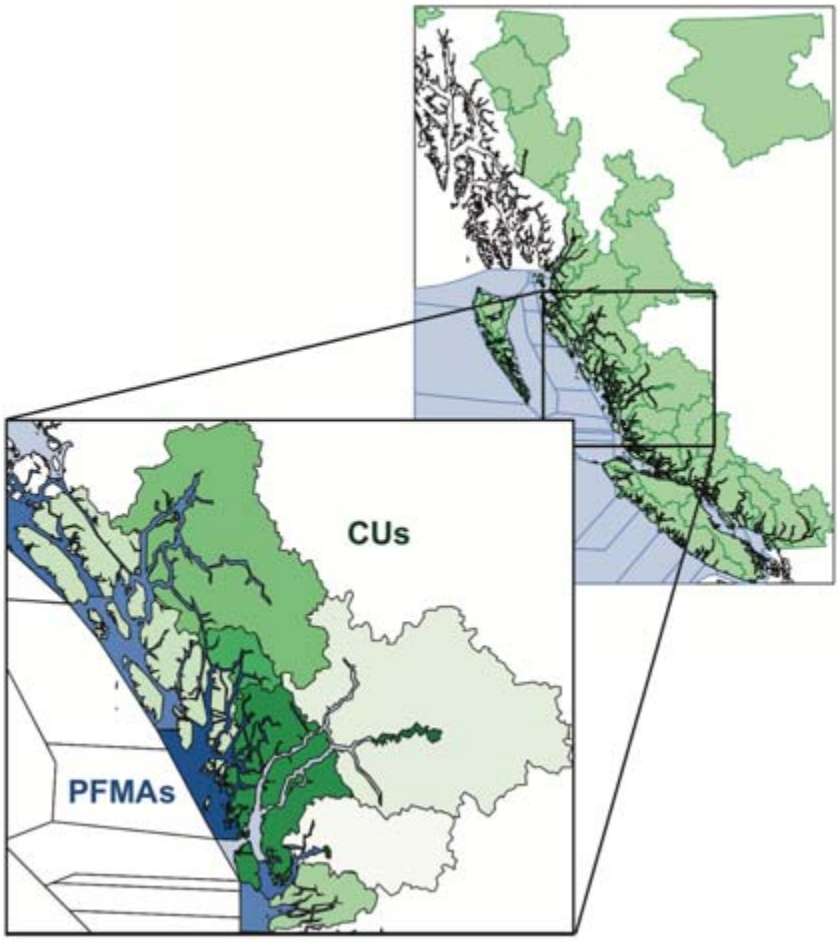

Figure 3. Our study focused on central coast chum salmon Conservation Units (CUs) (green; different CUs shaded differently in central coast inset),which are relatively large and do not correspond to the Pacific Fishery Management Areas (PFMAs; white or light blue shaded regions) for which catch is reported. One factor we investigated was how bias in assigning catch to CUs, resulting in under- or over-estimation of catch, affects estimates of biological status. Map produced using PBSmapping (Schnute et al. 2015) with data from Martin Huang (pers. comm. <Martin.Huang@dfo-mpo.gc.ca>).

Finally, to calculate recruitment for a given cohort of spawners, assumptions about the age-at-return of spawners in any given year are required (except in the case of pink salmon, which have a fixed 2-year generation time). The total return in a given year is assigned to brood

years based on the proportion of fish that return at a certain age, but these proportions are often not estimated every year. For chum salmon on the central coast, the distribution of age-at-return is assumed to be constant over time and is based on the average of available data (English et al. 2018). In this case, interannual variability in age-at-return may introduce uncertainty into the calculation of brood-year recruitment and bias resulting assessments of status (Zabel and Levin 2002).

Simulation model

We developed and applied a stochastic simulation model of salmon population dynamics that allows control over various biological and management factors that may influence the accuracy of status assessments. Our approach built on previous studies that evaluated uncertainties in fisheries management (e.g., Holt and Peterman 2008) and other factors influencing the performance of metrics and benchmarks under the Wild Salmon Policy (e.g., Peacock and Holt 2012; Holt and Folkes 2015; Holt et al. 2016, 2018). The simulation model is composed of submodels for salmon population dynamics, observation of spawners, assessment, harvest, and performance (Figure 4). In this subsection, we describe the general equations for the simulation model, with details of parameterization in the following subsection. The code for the model and results is available at https://github.com/sjpeacock/run-reconst-sim_PSF. Version 1.0, used in this paper, is archived at doi: 10.5281/zenodo.3971276 and was implemented in R version 4.0.2.

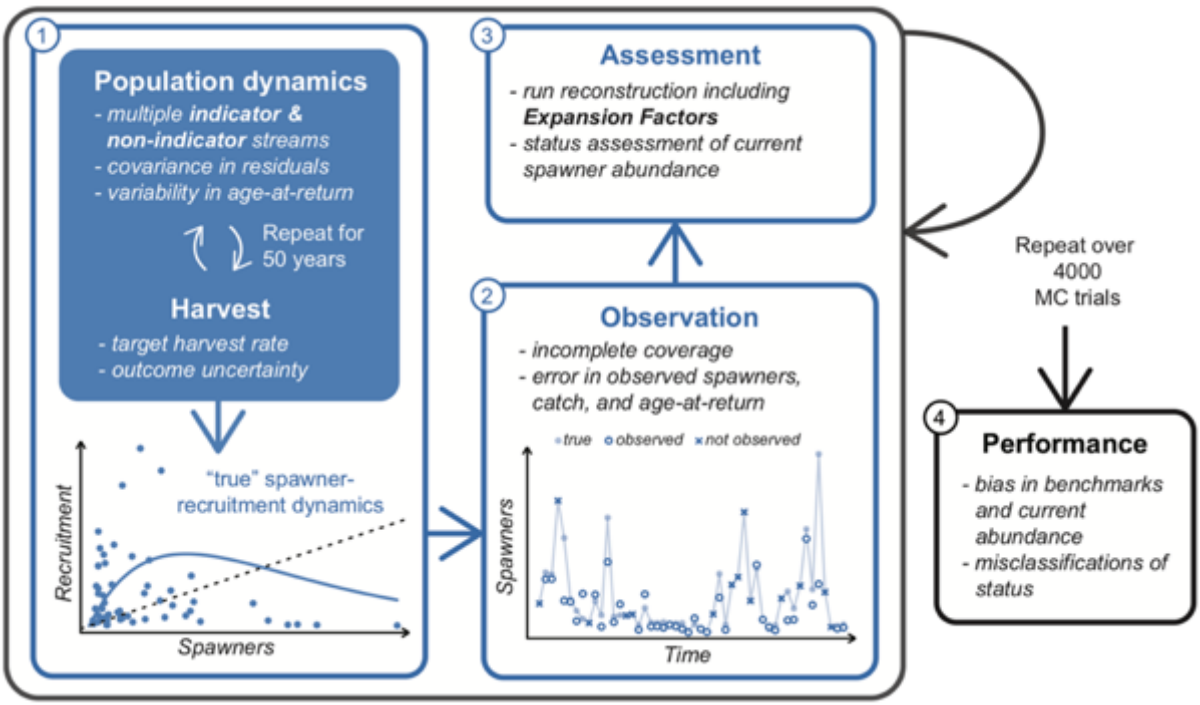

Figure 4. Schematic of the simulation model composed of submodels for population dynamics (including harvest), observation, assessment, and performance. Adapted from Holt et al. (2016).

Population dynamics

We simulated the dynamics of multiple spawning populations returning to indicator or non-indicator streams within a single hypothetical CU. Although some CUs consist of just a single spawning population (e.g., lake-type sockeye salmon), many CUs (especially pink and chum salmon) span thousands of square kilometers (Figure 3) and can include multiple spawning populations whose dynamics may differ due to local adaptation and finite rearing and spawning habitats.

We based our simulations on the life history of chum salmon, which generally return to spawn as 3-, 4-, or 5-year-olds. The number of salmon returning in return year t and population j, $R_{t,j}$, was calculated as:

$$R_{t,j} = R'_{t-3,j}\, p_{t-3,3} + R'_{t-4,j}\, p_{t-4,4} + R'_{t-5,j}\, p_{t-5,5}, \tag{1}$$

where $p_{y,g}$ is the proportion of recruits from brood year y returning at age g. Throughout this model description, we use R to denote returns, or catch plus escapement of fish returning in a year, and R' to denote recruitment, or the total number of offspring from a brood year that survive to maturity.

We assumed that the annual proportion of recruits returning at a given age was the same among populations, but incorporated interannual variability in age-at-maturity by allowing the proportion of recruits that return at age g to vary among brood years, y:

$$p_{y,g} = \frac{\overline{p}_g \exp(\overline{\omega}\, \varepsilon_{y,g})}{\sum_{G=3}^{5} \overline{p}_G \exp(\overline{\omega}\, \varepsilon_{y,G})}, \tag{2}$$

where $\overline{p}_g$ is the average proportion of individuals maturing as g year-olds, $\overline{\omega}$ is a parameter that controls interannual variability in proportions of fish returning at each age (Figure S4) and $\varepsilon_{y,g}$ are standard normal deviates (Schnute and Richards 1995).

The number of salmon that escape the fishery and return to spawn was calculated as:

$$S_{t,j} = \left(1 - h_{t,j}\right) R_{t,j}, \tag{3}$$

where $h_{t,j}$ is the realized harvest rate for population *j* in year t. We incorporated outcome uncertainty (i.e., deviations from the target harvest rates) by drawing the realized harvest rates for each year and population from a Beta distribution with mean equal to a target harvest rate, h'_t, and variance σ_h^2 (Holt and Peterman 2008).

We considered two different scenarios for determining the target harvest rate (Figure 5). First, we considered a simple, abundance-based harvest control rule (HCR) where h'_t increased with the total return to the CU from a minimum of 0.05 (to account for bycatch and unavoidable mortality and also avoid problems associated with low target HRs when incorporating Beta-distributed outcome uncertainty) up to an asymptote, h_{MAX} (Holt and Peterman 2008):

$$h'_t = \max\left(h_{MAX} \left(1 - \exp\left[m \sum_j R_{t,j}\right]\right), 0.05\right), \tag{4}$$

where m is the shape parameter of the HCR. The low harvest rates at low returns under this HCR prevented the CU from declining to red status in simulations, and so as to broaden our results to include CUs with true red status, we also considered a constant high target harvest rate of $h'_t = 0.60$ regardless of the total return (Figure 5). In the Supplementary Material, we present an intermediate scenario with a constant moderate target harvest of $h'_t = 0.42$.

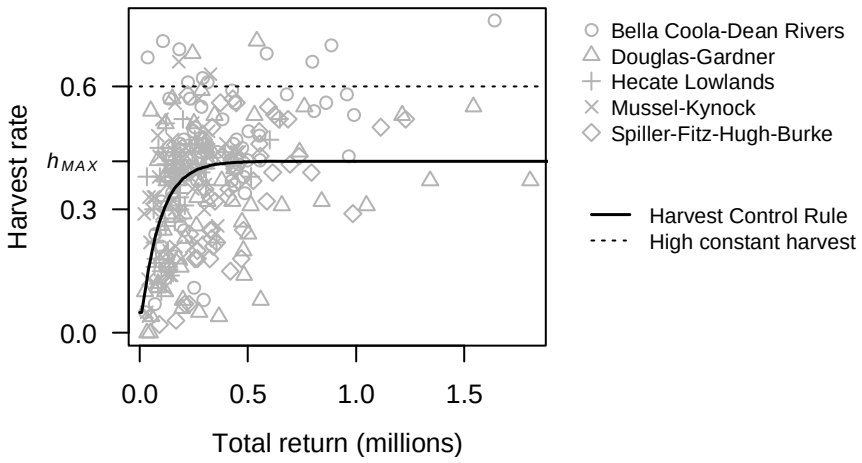

Figure 5. The two target harvest-rate cases we simulated were (1) a simple Harvest Control Rule (eq. (4); solid line) with parameters estimated from historical harvest rates and total return from five central coast chum CUs (grey points), (2) a constant high target harvest rate of $h_t' = 0.6$ *(dotted line).*

Each population in our model was harvested in proportion to its abundance, such that the true total catch of fish that would have returned to streams within the CU was calculated as:

$$C_t = \sum_j h_{t,j} R_{t,j}. \tag{5}$$

Although realized harvest rates differed among populations, we did not incorporate persistent biases in realized harvest rates among populations and thus assumed that all populations were equally vulnerable to the fishery. The extent to which this assumption is violated will depend on the size of the CU, the number of populations within it, the magnitude of variation in run-timing and body size among populations, and where fisheries are prosecuted. Any such biases among populations within a CU would likely be small because run-timing was a consideration

297 when delineating CUs (Holtby and Ciruna 2007). However, biases among CUs may be significant
298 and we investigate this by varying the observation bias in the total catch to the CU (see
299 Sensitivity analyses, below).

300 Finally, we assumed the spawner-recruitment dynamics followed a Ricker relationship
301 (Ricker 1954) yielding the number of recruits from brood year y and population j:

$$R'_{y,j} = S_{y,j} \exp\left(a_j - b_{j,y} S_{y,j} + \phi_{y,j}\right), \quad (6)$$

302 where a_j is the $\log_e$ recruits per spawner at low spawner abundance (i.e., productivity), $b_{j,y}$ is
303 the density-dependence parameter (which we allowed to change linearly through time; see
304 below), and $\phi_{y,j}$ are the recruitment deviates applied for year y and population j (eq. (7)).

305 We allowed productivity to differ among populations, where a_j was drawn from a
306 truncated normal distribution with mean $\overline{a}$ and variance σ_a^2, with a lower limit of $a_j = 0.4$
307 because SR benchmarks are not calculable for very low productivity (Holt and Ogden 2013; Holt
308 et al. 2018). For central coast chum salmon, we found only 1% of spawning populations (2/181)
309 had $a_j < 0.4$ (Online Supplement). A linear change in the density-dependence parameter was
310 simulated for some populations as a decline in the capacity of the freshwater habitat (i.e.,
311 $S_{MAX,j,y} = 1/b_{j,y}$, or the spawner abundance that leads to maximum recruitment). This decline
312 in capacities captured the potential consequences of cumulative stressors to freshwater habitat
313 among watersheds on the central coast (Connors et al. 2018). For all populations, the initial
314 capacity $S_{MAX,j,1}$ was drawn from a lognormal distribution whose mean and standard deviation
315 differed for populations in indicator versus non-indicator streams, as indicator streams tend to

be larger systems (English 2016; see Parameterization). The productivity and density-dependence parameters were drawn independently for each Monte Carlo (MC) iteration of the model.

We incorporated temporal autocorrelation in recruitment deviates:

$$\phi_{y,j} = \tau\, \phi_{y-1,j} + \sigma_{\upsilon}\sqrt{1-\tau^2}\, \upsilon_{y,j}, \tag{7}$$

where τ is the temporal autocorrelation coefficient, σ_{υ} is the standard deviation in recruitment deviates without temporal autocorrelation (Ricker 1975, Holt and Bradford 2011), and $\upsilon_{y,j}$ is a multivariate normal random variable with mean zero and variance-covariance matrix:

$$\Sigma_{j\times j} = \begin{bmatrix} \sigma_{\upsilon}^2 & \rho\,\sigma_{\upsilon}^2 & \dots & \rho\,\sigma_{\upsilon}^2 \\ \rho\,\sigma_{\upsilon}^2 & \sigma_{\upsilon}^2 & \dots & \rho\,\sigma_{\upsilon}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho\,\sigma_{\upsilon}^2 & \rho\,\sigma_{\upsilon}^2 & \dots & \sigma_{\upsilon}^2 \end{bmatrix}_{j\times j}. \tag{8}$$

Here, ρ is the correlation in recruitment deviates among populations.

We simulated the 'true' population dynamics over 50 years, after an initialization period of seven years to seed eq. (6) given the variable age-at-return of chum salmon. For each year in this initialization, we assumed that the number of spawners was equal to 20% of $S_{MAX,j,1}$. For the first year of the initialization, we set $\phi_{y-1,j}$ from eq. (7) to zero.

Observation submodel

In the observation submodel, we incorporated both incomplete monitoring coverage of streams and imperfect observation of spawners in streams that were monitored. In any given year, population *j* was observed with probability $\psi_{t,j}$. We included a linear decline in monitoring coverage (i.e., the probability of a population being observed) over time based on observations of declining monitoring coverage on the north and central coast of BC (Price et al. 2008, 2017; English 2016). We calculated the annual probability of being monitored separately for indicator and non-indicator streams based on observations that monitoring coverage of non-indicator streams is generally lower and has declined more severely than coverage of indicator streams (English 2016). See Parameterization for further details.

Spawner abundances were 'observed' with log-normal error:

$$\hat{S}_{t,j} = \begin{cases} \text{NA} & \text{if } z_{t,j} = 0 \\ S_{t,j} \exp\left(\delta_{t,j}\right) & \text{if } z_{t,j} = 1 \end{cases}, \tag{9}$$

where $z_{t,j} \sim \text{Bernoulli}(\psi_{t,j})$, $\delta_{t,j} \sim N(\overline{\delta} - \sigma_\delta^2/2,\ \sigma_\delta)$, $\overline{\delta}$ is the mean observation error, and σ_δ is the standard deviation in observation error of spawner abundances. Thus, this combines both the probability of a population being monitored and the distribution of observation errors around true spawner abundances if monitored. The mean observation error is corrected by $-\sigma_\delta^2/2$ so that the arithmetic mean observation bias is $\exp\left(\overline{\delta}\right)$. We included a negative bias in the observation of spawners ($\overline{\delta} \leq 0$) such that the observed spawner abundance is on average lower than the true spawner abundance. In general, it is challenging to enumerate spawners in

all reaches of a stream and in all streams within a CU. The reported spawner abundance is considered an underestimate of the total spawners in a CU, which motivates the application of Expansion Factor III for observer (in)efficiency when performing run reconstructions (Table 1). The calculation of Expansion Factor I required that we impose the constraint that at least one indicator stream was monitored each year, so if $z_{t,j}$ = 0 for all indicator streams in a year, we randomly selected one indicator stream to be monitored.

The catch to the entire CU in return year t was observed with log-normal error:

$$\hat{C}_t = C_t \exp(\chi_t), \tag{10}$$

where $\chi_t \sim N(\bar{\chi} - \sigma_\chi^2/2, \sigma_\chi)$, σ_χ is the standard deviation in catch error, and $\bar{\chi}$ is a bias in catch (corrected by $-\sigma_\chi^2/2$ as for observed spawners above). We assumed a default value of $\bar{\chi} = 0$, but in sensitivity analyses we varied this parameter to simulate scenarios in which fish are caught from other CUs or fish from the focal CU were caught in other PFMAs.

Previous models (e.g., Holt et al. 2018) have included error in the "estimated age-at-return" separately for each return year. For central coast chum salmon, annual age-at-return data are rarely sampled comprehensively so the same average is generally applied across all years (Peacock et al. 2014; English et al. 2018). Therefore, for each year in a MC trial we applied the same age-at-return, which was drawn independently for each MC trial using eq. (2) with observation error, $\overline{\omega_r}$ (Table A1).

Assessment submodel

As described above, the minimum requirement to calculate benchmarks and assess population status using percentile benchmarks is a time-series of spawner abundance to the CU. For SR benchmarks, harvest rates and age-at-return must also be estimated in order to calculate recruitment. The basic procedure of these run reconstructions is outlined in Table 1, and begins by expanding observations of spawners to indicator streams to the total spawner abundance to the CU by applying three Expansion Factors. The equations and criteria governing these Expansion Factors are detailed in the Online Supplement and in English et al. (2012, 2016, 2018). Briefly, Expansion Factor I ($F'_{t,d}$) imputes missing spawner abundances for populations in unmonitored indicator streams and is calculated for each year, *t*, within a decade, *d*, of the spawner time series. It relies on the decadal contribution of each indicator-stream population to the total escapement to all indicator streams (English et al. 2016). Expansion Factor II (F''_d) expands observations of spawners from indicator streams to include populations in non-indicator streams that are less frequently monitored, and is the same for each year within a decade, *d*. Expansion Factor II is calculated as the average proportion of total monitored spawners (in indicator and non-indicator streams) that are in non-indicator streams for that decade. For decades with insufficient information to calculate either of these Expansion Factors, for example due to declining monitoring coverage, a reference decade may be used. Expansion Factor III (F''') is determined by the regional DFO staff familiar with the escapement monitoring techniques used in each statistical area and is assumed to be constant through time (English et al. 2018). In our model, we assumed that all populations were at least

partially monitored, and that Expansion Factor III accounted for observation (in)efficiency, but in reality, Expansion Factor III may also account for populations in unmonitored streams.

The observed number of salmon returning in year t is the sum of observed catch and expanded escapement to the CU:

$$\hat{R}_t = \hat{C}_t + F'_{t,d}\, F''_d\, F''' \sum_{j \in I} \hat{S}_{t,j}, \tag{11}$$

where the summation includes observed spawner abundance to the I indicator streams only, with the non-indicator streams being accounted for through Expansion Factor II, F''_d.

We do not explicitly account for en route or pre-spawning mortality of fish that escaped the fishery, and assume that pre-spawning mortality is relatively small and accounted for in the productivity of the population through the Ricker SR dynamics. Observed recruitment for brood year y is calculated as the sum of age 3, 4, and 5 fish returning in years $y+3$, $y+4$, and $y+5$, respectively:

$$\hat{R}'_y = \hat{R}_{y+3}\, \hat{pr}_3 + \hat{R}_{y+4}\, \hat{pr}_4 + \hat{R}_{y+5}\, \hat{pr}_5, \tag{12}$$

yielding the 'reconstructed' spawner-recruit pairs for brood year y.

To calculate estimated SR benchmarks, we fit a linearized Ricker model to the observed data at the aggregate CU-level. The estimated productivity and density-dependence parameters were then used to calculate upper and lower SR benchmarks (80% of S_{MSY} and S_{GEN}) for the CU.

Performance submodel

For each MC simulation, we estimated status using both SR and percentile benchmarks calculated from the observed spawner-recruit pairs for the CU, including observation biases and incomplete monitoring coverage. Estimated status under both types of benchmarks was compared to the true status, which was calculated by comparing the current true spawner abundance (without observation error) against the upper and lower SR benchmarks (80% of S_{MSY} and S_{GEN}) from the underlying SR parameters. This meant that we evaluated estimates of percentile (and SR) benchmarks against the true underlying benchmarks derived from "true" SR parameters. Because we simulated the true dynamics at the scale of spawning populations and there was no "true" CU-level value for the Ricker parameters, calculating the true S_{MSY} (and thus the SR benchmarks) at the CU-level was not straightforward. We chose to calculate the true SR benchmarks for each population from the underlying Ricker parameters for that simulation and then summed the benchmarks across all populations to yield the "true" CU-level benchmarks. Although this CU-level benchmark will underestimate the level required to maintain all component populations above their individual benchmarks in any given year, the objective of the WSP is to maintain the overall CU and populations within CUs are assumed be recolonizable within reasonable time frames (Fisheries and Oceans Canada 2005). When declines in capacity were included in the simulation, we calculated true status from the initial capacity parameters before the decline. Although this may have increased the chance that estimated status was lower than true status (increasing the rate of pessimistic misclassifications), we took this approach to avoid a shifting baseline in benchmarks.

Performance was evaluated in two ways that capture the difference between estimated and true status: (1) the proportion of MC simulations for which status was correctly assessed as green, amber, or red, and the proportion of simulations for which status was either underestimated (pessimistic) or overestimated (optimistic), and (2) the percent relative bias of observed average spawner abundance (S_{AVG}) and of the four benchmarks (S_{GEN}, 80% S_{MSY}, S_{25}, S_{50}) compared to their true values for each MC simulation.

For each parameterization investigated (see below) we ran 4000 MC trials, which was sufficient to ensure the mean percent error in performance measures was < 3% (Figure S3).

Parameterization

Some of the parameters in our simulation model were unknown or unknowable, in which case we followed assumptions made for southern BC chum salmon by Holt et al. (2018). Other parameters were available specifically for central coast chum salmon or could be estimated from available data; details of parameter estimation are given in the Online Supplement. As mentioned above, in order to understand the assessment biases under different true statuses we considered two cases: (1) high productivity and a conservative harvest control rule (HCR), which we refer to as the “base case” because it is most representative of central coast chum salmon, and (2) low productivity and a constant high target harvest rate, which represented a CU at high risk of extirpation. In the Online Supplement, we also present results from a third case intermediate between these two with low productivity and a constant moderate harvest rate. Unless otherwise indicated in the

sensitivity analysis (below), parameters defaulted to the values described here and listed in the Appendix, Table A1.

The mean proportion of adults maturing at ages 3, 4, and 5 in eq. (2) was held constant for all simulations at 0.23, 0.64, and 0.13, respectively, based on the average age-at-return applied in run reconstructions of central coast chum (Challenger et al. 2018; English et al. 2018). For the base case, we estimated the parameters in the HCR (Figure 5) from harvest rates and total return sizes for five central coast chum CUs (English et al. 2018; Salmon Watersheds Program - Pacific Salmon Foundation 2019). The high harvest case is consistent with the upper bound in harvest rates chosen in other studies (Holt et al. 2018) and represents the upper threshold in harvest rates for central coast chum (Figure 5).

To estimate the SR parameters for spawning populations (eq. (6)), we fit a linearized Ricker model to population-level spawner-recruit pairs from nine central coast chum CUs with individual productivity and density-dependence parameters for each population (Figure S5). Note that this approach may lead to biased SR parameters as we did not account for temporal autocorrelation and errors in variables when analysing the data, but they nonetheless provide a useful starting point for our simulation study. From the model fits, we calculated (1) mean and variance in productivity among populations, (2) mean and variance in the initial capacity (i.e., $1/S_{MAX}$) for indicator and non-indicator streams, (3) the residual variance within populations, (4) the average correlation in residuals between populations, and (5) the temporal autocorrelation in residuals within populations. The mean population-level productivity was $\bar{\alpha}$ = 1.40 (Figure S5), which we applied in our base case. For the low-productivity case, we chose $\bar{\alpha}$ = 0.56, which

462 was the 2.5th percentile of population-level productivity estimates (Figure S5). The residual
463 variance within populations was used as an estimate of the variance in recruitment deviations
464 (σ_{υ}^2) in simulations, but was likely an overestimate as we did not account for observation error
465 separately when fitting the Ricker model. Thus, we explore the results under a range of values
466 for σ_{υ} in the Supplemental Results (Online Supplement). The correlation in residuals among
467 populations was estimated at $\rho = 0.46$, but we also investigated lower ($\rho = 0$) and higher
468 ($\rho = 0.9$) levels of synchrony in the Online Supplement.

469 As the default case, we considered a decline in the capacity of streams (i.e., one over
470 the density-dependence parameter of the Ricker model) that reflected observed declines in
471 freshwater habitat (Office of the Auditor General of Canada 2004). Within the central coast
472 chum CUs, 29% of watersheds are considered to be at moderate risk, and 21% at high risk, of
473 cumulative habitat pressures over the last 60 years from stressors such as logging, water
474 licenses for withdrawal of water from streams, and stream crossings (Connors et al. 2018). We
475 hypothesized that declines in capacity that differ among spawning populations may affect the
476 accuracy of Expansion Factors I and II, particularly in combination with declining monitoring
477 coverage (Price et al. 2017). We incorporated a linear decline in capacity (i.e., $S_{MAX} = 1/b$) over
478 the 50-year time series of between 25% and 50%, representing a moderate decline, for 29% of
479 populations (the percentage of central coast chum watersheds deemed to be at moderate risk
480 of cumulative habitat pressures by Connors et al. 2018) and a linear decline between 50% and
481 75%, representing severe decline, for 21% of populations (the percentage deemed to be at high
482 risk). The exact percent decline for each population was randomly drawn from a uniform

distribution within the above range for each MC simulation. The remaining 50% of populations had stable capacity over the 50-year simulation. In a sensitivity analysis, we investigated four additional scenarios for declining capacity (Table 2).

In the observation submodel, we chose the bias in the observation of spawners to match Expansion Factor III, which corrects for observer efficiency, with a range of values explored in a sensitivity analysis (below). Because Expansion Factor III is simply a constant multiplier applied to spawner abundance, it does not affect estimated status by the percentile benchmarks. However, Expansion Factor III affects estimated harvest rates and recruitment, and therefore bias in this Expansion Factor was expected to lead to bias in estimated status under the SR benchmarks. The value of Expansion Factor III applied in past status assessments has been constant at F''' = 1.5 for all central coast chum CUs (English et al. 2016), and so we applied a default value of $\bar{\delta}$ = $\log_e(1/1.5)$ = -0.4. The majority of central coast chum streams are surveyed visually by ground (fish counts or other sampling) with some aerial (fish counts) or boat (fish counts or other sampling) surveys (English 2016), similar to southern BC chum streams (Holt et al. 2018). We assumed σ_δ = 0.5 following Holt et al. (2018), which is the maximum estimated uncertainty for visually surveyed spawners (Cousens et al. 1982; Szerlong and Rundio 2008).

We incorporated a linear decline over the last 27 years of simulations in the proportion of indicator and non-indicator streams monitored each year from 0.76 and 0.72, respectively, to 0.72 and 0.05 based on English (2016). These declines are representative of overall declines in

monitoring across species, but we also consider the trends specific to chum salmon in the Online Supplement (Figure S6).

We assumed no bias in the observation of catch ($\bar{\chi}$ = 0) by default, but consider a range of bias in sensitivity analyses (below). The standard deviation in the observation error of catch accounts for differences between observed (i.e., reported) and actual catch due to uncertainties with mixed-stock fisheries and in reporting and estimation of recreational fisheries and subsistence use. We set this to σ_{χ} = 0.2 (Holt et al. 2018), which is less than the observation error in spawners.

Sensitivity analyses

We examined the sensitivity of status assessments over a range of values for several different model parameters that we considered most likely to affect status assessments due to their influence on the assumptions in run reconstructions (Table 1; Table 2). We investigated each of the questions below under both the base case and low-productivity high-harvest case described above, with other parameters at their default values (Table A1) unless otherwise noted. The specific questions that we addressed were:

1. How does the number of spawning populations and the proportion designated as indicator streams affect status assessments? The lower the proportion of streams that are indicators, the greater the magnitude of Expansion Factor II.
2. How does a decline in monitoring coverage affect status assessments? The fewer indicator streams that are monitored, the greater the magnitude of Expansion Factor I

and the potential uncertainty in expanded spawner abundance. Here, we consider two scenarios (Table 2; Figure S6): constant monitoring coverage at historical proportions among all streams and an observed decline in coverage starting in the mid-1980s as has been observed on the north and central coast (English 2016; Figure S6). In the Online Supplement we consider two additional scenarios: observed declines in monitoring specific to chum salmon streams and a sharp, recent decline in monitoring of indicator streams.

3. How do declines in capacity affect status assessments? The application of Expansion Factors I and II assumes that the relative contributions of populations to aggregate abundance in the CU does not change over time, but declines in capacity that differ among populations may violate this assumption.
4. How does spawner observation bias affect status assessments, given that the value of Expansion Factor III is fixed over time and often the same among CUs (English et al. 2018)?
5. How does catch observation bias (e.g., over- or under-estimating catch of salmon) affect status assessments? This represents scenarios where there are errors in estimates of CU proportions in the aggregate catch in a mixed-stock fishery, or violation in the assumption of homogenous spatial and temporal distribution of CUs when CU proportions are not monitored in such fisheries.
6. How does interannual variability in age-at-return affect status assessments?

We investigated the impact of declines in monitoring coverage (question #2 above) in combination with declines in capacity of spawning populations (question #3) in a bivariate sensitivity analysis.

Results

The different productivity and harvest rate combinations we considered led to different true CU statuses. Under high productivity and an abundance-based harvest control rule (HCR) – the base case corresponding to central coast chum salmon – 86.4% of simulations resulted in true green status (Figure 6a,b). Conversely, under low productivity and high harvest rates, 67.0% of simulations resulted in true red status (Figure 6c,d).

Under the base case when true status was mostly green, misclassifications resulted in estimated status lower than the true status meaning assessments were biologically pessimistic (henceforth referred to as "pessimistic misclassifications"). This was particularly true of the percentile benchmarks, for which 55.2% of simulations resulted in a pessimistic misclassification with 12.7% of simulations having misclassified green status as red. Pessimistic misclassifications were due to positive bias in benchmarks and not bias in the current spawner abundance (Figure 7), resulting in status being underestimated. For productive populations (as in the base case), most observed spawner abundances tended to be far above lower benchmarks and closer to equilibrium values. As a result, lower and upper benchmarks of 25th and 50th percentile of historical spawner abundance tended to overestimate the "true" SR-based benchmarks.

When true status was mainly red, under low productivity and high harvest rates, biologically optimistic misclassifications (henceforth "optimistic misclassifications") were more common, which may be riskier from a conservation and management standpoint. For example, 44.8% and 43.0% of simulations had an estimated status higher than true status under the spawner-recruitment (SR) and percentile benchmarks, respectively (Figure 6c,d). These more frequent optimistic misclassifications were due to a negative bias in benchmarks, in particular the lower SR benchmark of S_{GEN} (Figure 7a), likely due to a poor ability to estimate S_{MSY} under low productivity when spawner abundances tend to cluster near the origin.

Under both types of benchmarks, bias did not decrease when monitoring coverage was held constant at 100% (Figure S9), suggesting that the application of Expansion Factors I and II were not contributing factors.

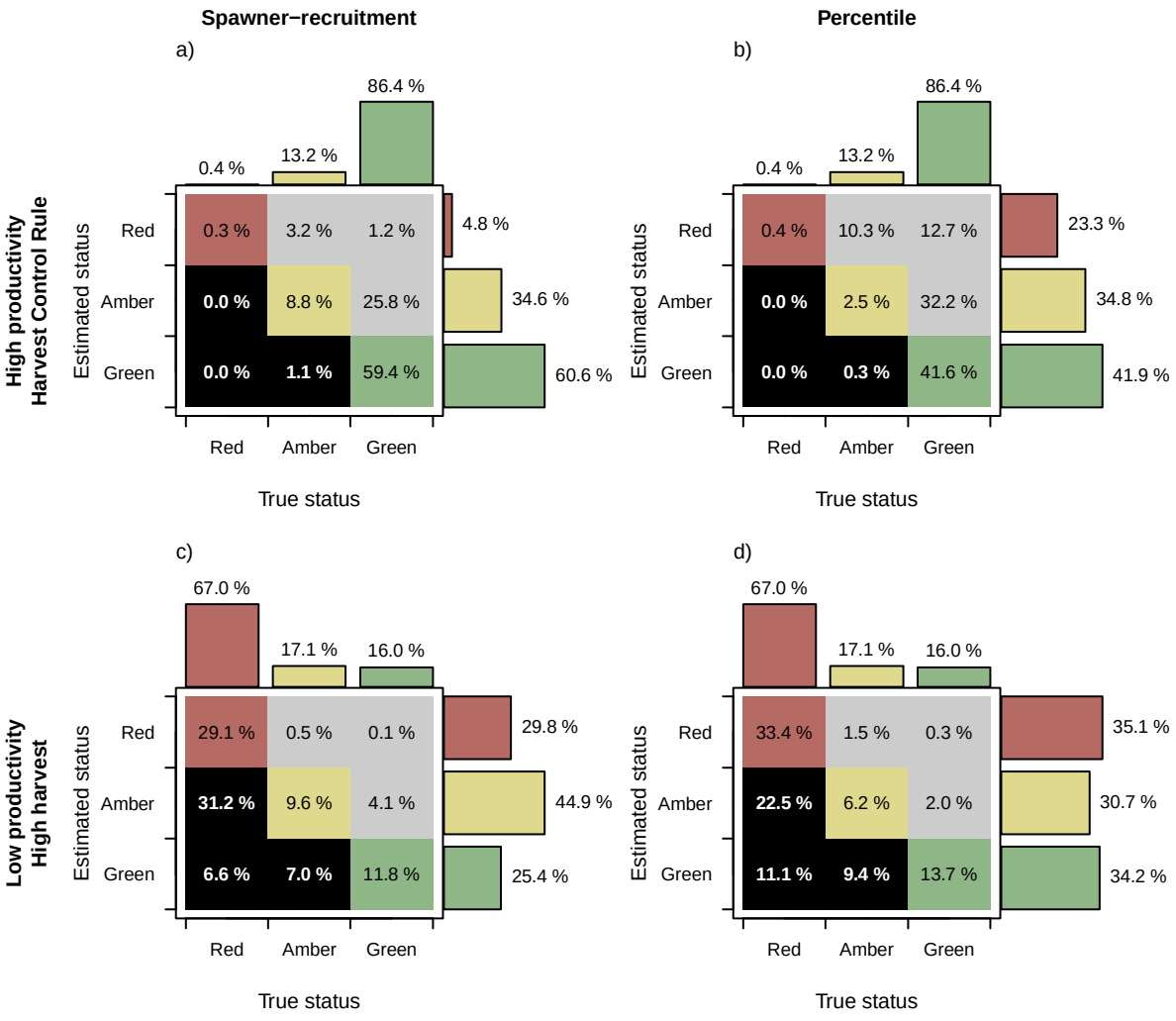

Figure 6. Estimated status according to the spawner-recruitment (SR) benchmarks (left) and the percentile benchmarks (right), over the true status for each of two cases (high productivity and an HCR (a-b) and low productivity and high harvest rates (c-d)). Grey cells indicate pessimistic misclassifications, which may lead to overly conservative management actions, and black cells indicate optimistic misclassifications, which may lead to overly risky management actions. Coloured cells indicate correct classifications for red, amber, and green zones, respectively.

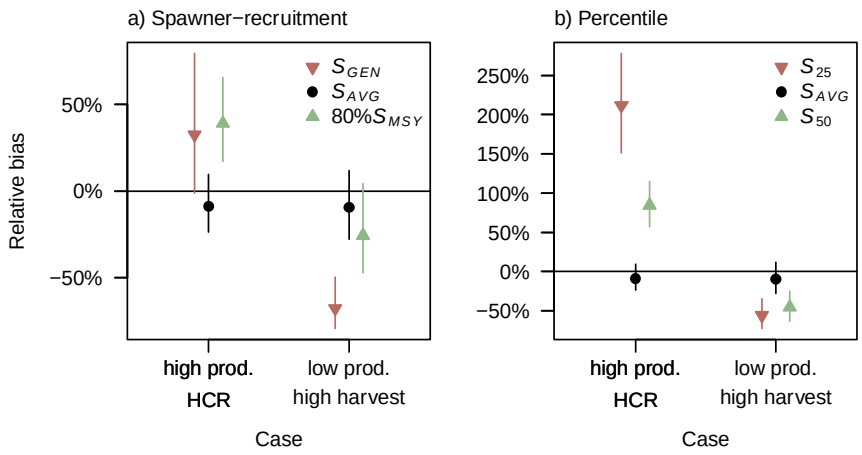

Figure 7. Relative bias in spawner-recruitment (SR) benchmarks (a), percentile benchmarks (b), and current spawner abundance (S_{AVG}; black) (median ± interquartile range among 4000 MC simulations) for high productivity and a harvest control rule (HCR) and the low-productivity high-harvest case.

Sensitivity analyses

The number of spawning populations within the CU and the proportion of those populations spawning in indicator streams had little impact on status assessments (Figure S10). Under the base case, the relative bias in estimates of S_{MSY} and S_{GEN} were lower in larger CUs, with half as many pessimistic misclassifications for larger CUs under the SR benchmarks (43% for 10 populations versus 22% for 140 populations with 30% indicator streams; Figure S10a). This trend was not, however, observed in the low-productivity high-harvest case when true status was predominantly red (Figure S12).

The monitoring-coverage scenarios that we considered, representative of observed declines in monitoring on the north and central coast, had no effect on status outcomes or the relative bias in benchmarks. This was true in the base case (Figure 8, Figure S13) and under low

productivity (Figures S14 – S15). Even under severe declines in capacity of 50% to 75% for all spawning populations, our results suggest that the observed declines in monitoring coverage on the north and central coast are unlikely to bias status assessments. This result held regardless of whether the recruitment deviates among component populations within the CU were not correlated (ρ = 0; Figure S16) or highly correlated (ρ = 0.9; Figure S17).

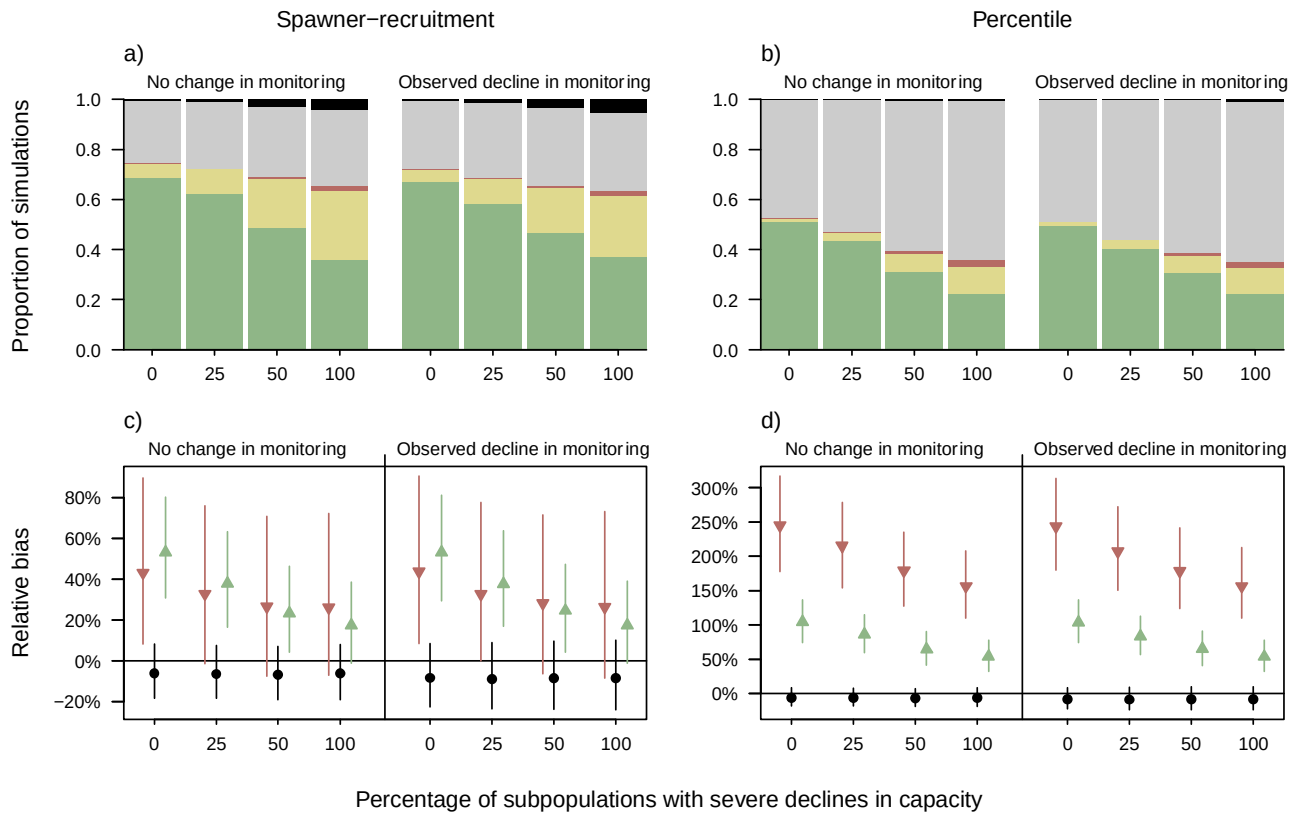

Figure 8. The effect of monitoring coverage (no change and decline; Table 2) and the percentage of spawning populations with severe declines in capacity (x-axis) on performance measures under the base case of high productivity and HCR. (a-b) The proportion of simulations with correct green, amber, or red status or pessimistic misclassifications (grey) and optimistic misclassifications (black) under the SR benchmarks (a) and percentile benchmarks (b). (c-d) The percent relative bias (median ± interquartile range among 4000 MC simulations) in the current spawner abundance (S_{AVG}; black) and lower and upper benchmarks (red and green, respectively) under the SR benchmarks (c) and the percentile benchmarks (d). See Online Supplement for results under the low-productivity high-harvest case.

Under the base case, declines in capacity of the CU were associated with poorer estimated status and an increase in misclassification rates (Figure 8a-b). Pessimistic misclassifications increased because CUs more frequently had a true status of amber but were

misclassified as red. Optimistic misclassifications increased, particularly under the SR benchmarks (Figure 8a), because the relative bias in the current spawner abundance (S_{AVG}) remained unchanged, but the bias in benchmarks decreased (Figure 8c-d). In the low-productivity high-harvest case, the results were similar but with biologically optimistic misclassifications dominating as status was predominantly amber or red (Figure S15).

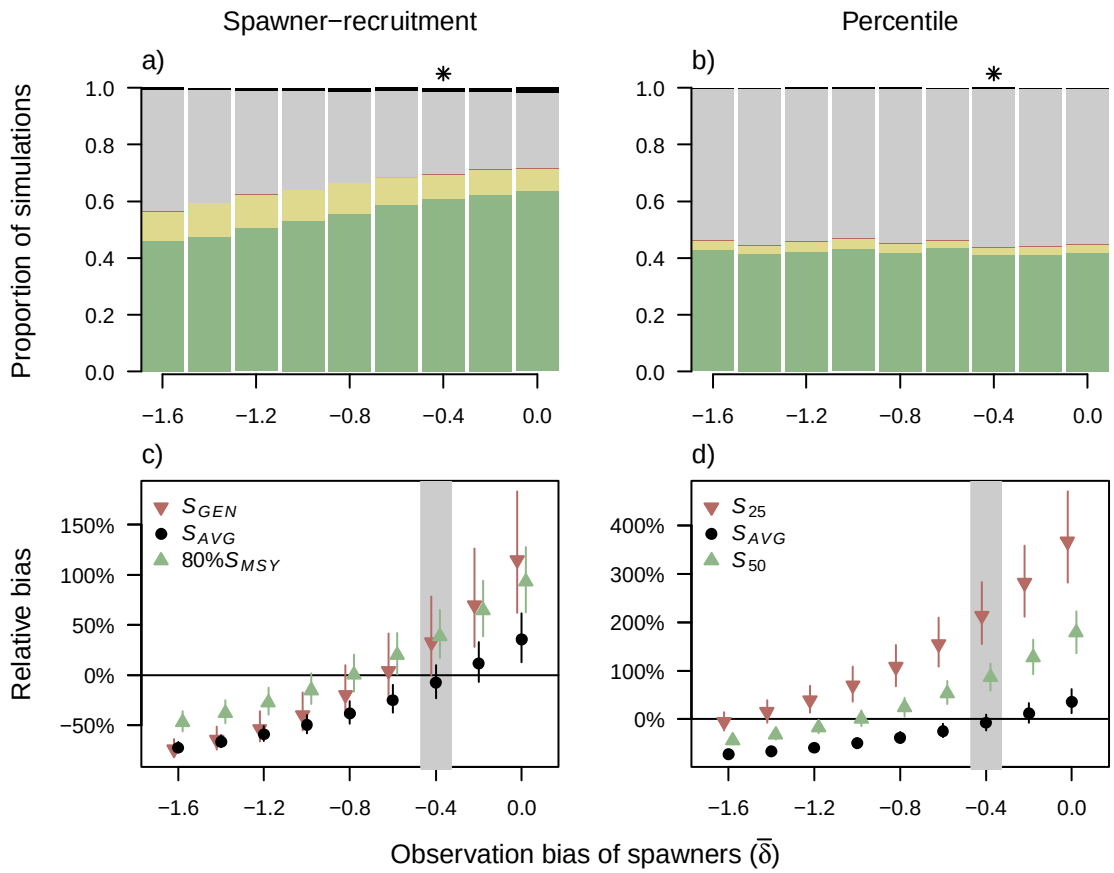

Figure 9. The effect of observation bias in the number of spawners (x-axis) on performance measures under the base case. (a-b) The proportion of simulations with correct green, amber, or red status or pessimistic misclassifications (grey) and optimistic misclassifications (black) under the SR benchmarks (a) and percentile benchmarks (b). (c-d) The percent relative bias (median ± interquartile range among 4000 MC simulations) in the current spawner abundance (S_{AVG}; black circle) and lower and upper benchmarks (red and green, respectively) under the SR benchmarks (c) and the percentile benchmarks (d). The asterisk in (a-b) and grey zone in (c-d) indicate the default parameter value of $\bar{\delta}$ = -0.4, and the bias that matches the Expansion Factor III of F''' = 1.5 applied in all simulations. See Online Supplement for results under the low-productivity high-harvest case (results were similar).

As the bias in the observation of spawners approached zero ($\bar{\delta} \rightarrow 0$), misclassifications under the SR benchmarks declined in all cases, even as the observation bias became less than the Expansion Factor III applied to correct for it (Figure 9a, Figures S18-S19). The relative bias in

the current spawner abundance was minimized when the observation bias matched the assumed value of Expansion Factor III ($\bar{\delta}$ = -0.4 corresponding to F''' = 1.5; Figure 9c-d). When observed spawner abundance was biased low ($\bar{\delta}$ = -1.6), the relative bias in the upper benchmark of 80% S_{MSY} was higher than the relative bias in current spawner abundance (S_{AVG}) or the lower benchmark (S_{GEN}), and thus CUs with a true green status were more likely to be misclassified as amber. Status outcomes under the percentile benchmarks were unaffected by changes in observation bias of spawners (Figure 9b); the bias in both benchmarks and current abundance showed similar changes as observation bias declined (Figure 9d) such that the resulting status outcome was unchanged. In the Online Supplement, we also investigated a change in observation bias halfway through the simulation (Figure S20), but a change from the base value of $\bar{\delta}$ = -0.4 to $\bar{\delta}$ = -1.6, -0.7, and 0 did not have any impact on status outcomes or biases in benchmarks (Figure S21).

Underestimation of catch (i.e., negative values of $\bar{\chi}$) resulted in fewer misclassifications than overestimating catch (Figure 10, Figures S22 – S23). As the catch bias increased from $\bar{\chi}$ = -1.0 (63% underestimation) to $\bar{\chi}$ = 1.0 (271% overestimation), the relative bias in the lower SR benchmark of S_{GEN} declined while the relative bias in the upper benchmark of 80% S_{MSY} increased (Figure 10b). This is due to the errors in variables that occur when catch is underestimated: productivity and recruitment tend to be underestimated, thus leading to lower estimates of S_{MSY} and higher estimates of S_{GEN} (Holt and Folkes 2015). Under the base case, the true status was green in the majority of simulations and so the increasing bias in the upper benchmark dominated the overall status assessments and led to the increase in

pessimistic misclassifications with increasing $\bar{\chi}$. In the low-productivity high-harvest case, true status was mostly red and so the increasingly negative bias in S_{GEN} resulted in more optimistic misclassifications as the observation bias in catch increased (Figure S23). In all cases, overestimating catch by ~50% (i.e., $\bar{\chi}$ = 0.4) led to a 5 – 8% increase in misclassification rate (Figure S24). Although these changes in misclassification rates may seem small, there is potential for large catch errors in run reconstructions, especially when multiple CUs overlap with a single PFMA. Catch does not factor into the calculation of percentile benchmarks, so status under the percentile benchmarks was unaffected by changing catch bias.

Finally, increasing interannual variability in age-at-maturity resulted in more frequent status misclassifications, but the effect was relatively small. Under the base case, increasing $\bar{\omega}$ from 0.2 to the default value of 0.8 resulted in an increase in misclassifications from 27.3% to 30.2%, but very little change in the bias in benchmarks (Figure S25). Further increasing the interannual variability to $\bar{\omega}$ = 1.6 led to 34.0% of simulations being misclassified, but this level of interannual variability is high (see Figure S4 for example) compared to data on age-at-return for central coast chum salmon (Challenger et al. 2018; English et al. 2018). The increase in misclassifications was smaller under the low-productivity high-harvest case (Figure S25).

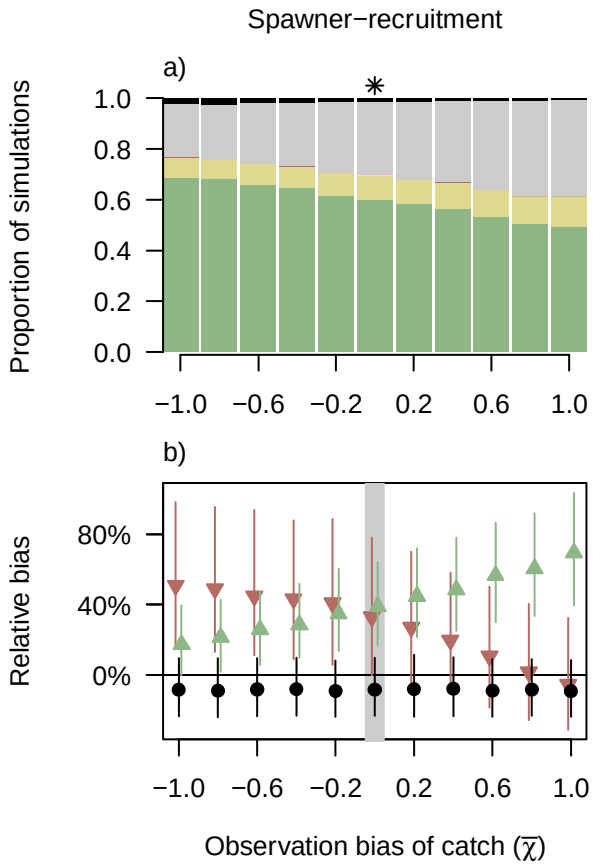

Figure 10. The effect of observation bias in catch (x-axis) on performance measures under the base case. (a) The proportion of simulations with correct green, amber, or red status or pessimistic misclassifications (grey) and optimistic misclassifications (black) under the SR benchmarks. (b) The percent relative bias (median ± interquartile range among 4000 MC simulations) in the current spawner abundance (S_{AVG}; black circle) and lower and upper benchmarks (red and green, respectively) under the SR benchmarks. The asterisk in (a) and grey zone in (b) indicate the default parameter value of $\bar{\chi}$ = 0. Observation bias in catch ranges from 63% underestimation ($\bar{\chi}$ = -1) to 271% overestimation ($\bar{\chi}$ = 1). See Online Supplement for results under the low-productivity high-harvest case.

Discussion

In this study, we quantified the impact of common assumptions in run reconstructions (Figure 2c) on biological status outcomes. In general, the assumptions that we hypothesized

678 might introduce biases (Table 1) had little effect on status outcomes, suggesting that attempts
679 to assess status in the face of limited or uncertain data are worthwhile. In particular, there was
680 almost no effect of declines in monitoring coverage on the accuracy of benchmarks or resulting
681 status outcomes, even in the face of different trends in capacity and reduced synchrony (i.e.,
682 zero autocorrelation in recruitment deviates) among spawning populations within the CU.
683 Extreme declines in monitoring may impact assessments – certainly, if no spawning populations
684 are monitored then our ability to assess status will inevitably be compromised – but the current
685 approach to imputing and expanding spawner abundances appeared robust to declines in
686 monitoring coverage in the range documented for the north and central coast (English 2016;
687 Price et al. 2017).

688 Perhaps unexpectedly, misclassifications were not minimized when the value of
689 Expansion Factor III used in run reconstructions, correcting for observer efficiency, perfectly
690 offset the observation bias in the underlying simulation. The bias in estimated current spawner
691 abundance was lowest when Expansion Factor III matched observation bias, but the positive
692 bias in benchmarks under the base case meant that misclassifications under the spawner-
693 recruitment (SR) benchmarks were minimized when there was no observation bias in spawners.
694 Status outcomes under the percentile benchmarks were unaffected by observation bias, as this
695 bias was assumed to be constant over time and affected the current and historical spawner
696 abundances equally.

697 Under our base case of high productivity and an abundance-based harvest control rule
698 (HCR) consistent with historical central coast chum salmon harvest rates, most simulations had

a true green status, but pessimistic misclassifications as amber were common under both SR and percentile benchmarks. The estimated status from our simulations roughly matched the status outcomes for seven central coast chum CUs from that period, with the majority of CUs having green status under the SR benchmarks and amber status under the percentile benchmarks (Connors et al. 2018). (Note that status of central coast chum salmon CUs has since declined; see the Pacific Salmon Explorer (www.salmonexplorer.ca) for the most up-to-date assessments.) Consistent with the real status assessments, status under the percentile benchmarks tended to be poorer than status under the SR benchmarks. Our simulations attributed this to a higher relative bias in percentile benchmarks, as found for south-coast chum salmon under high-productivity scenarios (Holt et al. 2018). Pessimistic misclassifications may lead to overly conservative management actions, consistent with the Precautionary Principle (Foster et al. 2000), but may also result in foregone harvest and adverse socio-economic consequences (Walters et al. 2019).

For simulations with low productivity and high harvest rates (i.e., true red status), the bias in benchmarks was negative, resulting in a high proportion of optimistic misclassifications. This represents a high-risk management scenario from a conservation standpoint, particularly as the true underlying status is red. Further, the absolute biases in benchmarks were more severe under declines in capacity when status was predominantly red (productivity low) compared to predominantly green (productivity high). The bias in SR benchmarks was particularly sensitive to the underlying true status, presumably because productivity and recruitment, which influence status, also affect the bias in SR parameters (Subbey et al. 2014)

720 that arises due to recruitment-spawner linkage inherent in the data (Walters 1985; Korman et
721 al. 1995) and/or due to error in spawner estimates (Walters and Ludwig 1981; Kehler et al.
722 2002; Kope 2006). More sophisticated modelling approaches, such as the state-space models
723 that can account for errors in variables and allow for information to be shared among
724 populations and/or CUs, may reduce bias in benchmarks and lead to more robust status
725 assignments (Staton et al. 2020).

726 Our results suggest that overestimating catch should be avoided. In particular, under
727 low productivity and high harvest rates, optimistic misclassifications associated with
728 overestimating catch and therefore underestimating the lower benchmark, S_{GEN}, may put
729 populations at further risk. Under the base case of high productivity and an HCR,
730 overestimating catch resulted in more frequent pessimistic misclassifications as the upper
731 benchmark (S_{MSY}) was overestimated, resulting in CUs with a true green status being estimated
732 as amber. In both cases, the impact of overestimating catch has the potential to significantly
733 bias assessments: overestimating catch by ~50% led to a 5 – 8% increase in misclassification
734 rates. This level of catch overestimation (and higher levels) may occur when fish caught in a
735 Pacific Fisheries Management Area (PFMA) and assigned to the CU that overlaps with that
736 PFMA were actually bound for other CUs. This could occur in mixed-stock fisheries if genetic
737 stock identification is not undertaken to validate assumptions regarding run-timing and
738 migration patterns. Increased efforts to quantify catch composition, run timing, and spatial
739 distribution of Pacific salmon CUs are therefore needed to more accurately estimate harvest
740 rates and minimize misclassifications of biological status.

Limitations, challenges, and future research

Closed-loop simulation models are powerful tools for evaluating management strategies and quantifying the biases in parameter estimation and status outcomes (Walters 1986; Peterman 2018), but are not without their weaknesses. As with any model, our simulation model was an approximation of reality and thus we had to make a number of assumptions. Both the true population dynamics and assessment submodel assumed a Ricker SR relationship (Ricker 1954), but we recognize that there is considerable model uncertainty. Distinguishing among different models (e.g., Ricker versus Beverton Holt) in assessing status would be challenging, particularly under the declines in capacity that we simulated to capture observed changes in freshwater habitat. The Ricker model is commonly used to set management targets and for simulating population dynamics of Pacific salmon (e.g., Peterman et al. 2000; Peacock and Holt 2012; Fleischman et al. 2013; Holt and Folkes 2015), and a full exploration of other true models was beyond our scope.

We did not implement a bias correction when simulating log-normal recruitment deviates, consistent with other Pacific-salmon simulation studies (e.g., Dorner et al. 2013; Cunningham et al. 2019). This introduced a positive bias in mean recruitment and subsequent estimates of productivity. Alternatively, one could include a bias correction when simulating recruitment by assuming the mean of $\upsilon_{y,j}$ in eq. (7) is $-\sigma_{\upsilon}^{2}/2$, and this may be best practice moving forward (e.g., Hicks et al. 2020). However, there is no clear consensus about whether this correction is necessary when parameters are estimated from simulated data in an assessment model without a $+\sigma_{\upsilon}^{2}/2$ adjustment (Hilborn and Walters 1992).

We considered true SR dynamics to operate at the spatial scale of spawning populations (i.e., individual streams), but there is evidence that the processes influencing productivity and density-dependence may operate at broader, regional spatial scales (e.g., Malick and Cox 2016). Other simulation models have incorporated straying among spawning populations within a CU (e.g., Peacock and Holt 2012; Holt and Folkes 2015). While that approach incorporates density-dependence that may occur at broader spatial scales, it also requires additional assumptions to be made about the probabilities of straying among streams, which are largely unknown.

Simulating true dynamics at the scale of individual spawning populations also complicates the calculation of true status at the CU-level from SR benchmarks. We chose to calculate SR benchmarks at the spawning-population-level and then sum across spawning populations to calculate S_{MSY} and S_{GEN} at the CU-level. There are other approaches to calculate aggregate benchmarks, but each has its own potential biases. For example, SR relationships could be fit to the "true" data aggregated at the CU-level and SR benchmarks calculated from the resulting CU-level estimates of productivity and density dependence. The way in which spawning-population-level benchmarks are aggregated to CU-level benchmarks may affect performance in our simulations, and a full exploration of how different methods of aggregation affect our results warrants future consideration.

Despite these limitations, the simulation model that we adapted and applied is flexible enough that it can accommodate different Pacific salmon species and life-history traits, opening the door to future work investigating the impact of different assumptions and the impact of the assumptions that we focused on under additional scenarios. Here, we considered a relatively

simple run-reconstruction model, but further work is needed to quantify how observation errors and uncertainty in the spatial and temporal distribution of returns affects status outcomes when more complex run-reconstruction models are used. Temporal shifts in biological parameters, including age-at-return (e.g., associated with environmental change and selective fisheries) and productivity (non-stationarity has been widely observed in Pacific salmon; e.g., Peterman and Dorner 2012; Malick and Cox 2016; Dorner et al. 2017), are also areas that warrant further exploration. Additional simulations could also help inform methods in the assessment process, such as the optimal time-series length for detecting changes, whether benchmarks should be updated with each assessment, and the best analytical approach to calculate SR benchmarks (e.g., Bayesian hierarchical models vs. single-stock ordinary least squares).

Conclusions

Pacific salmon are one of the most data-rich groups of fish due to their high economic, social, and cultural value, but nonetheless our knowledge of their dynamics is uncertain. Assessing the biological status of Pacific salmon CUs is a conservation and management priority given the continued declines of many stocks (e.g., COSEWIC 2016, 2017) and escalating threats to salmon conservation. Current government-led integrated status assessments under the WSP include expert opinion on the data uncertainties and risks unique to each CU (e.g., DFO 2015, 2016, 2018a), making the process time and resource intensive. This has limited their timely application to all 460+ Pacific salmon CUs in Canada (DFO 2019), and also means that the process is not entirely transparent or reproducible. The Pacific Salmon Foundation has

implemented data-driven biological status assessments based on a subset of the indicators and benchmarks recommended under the WSP, with the results accessible through the Pacific Salmon Explorer (www.salmonexplorer.ca). These data-driven assessments, similar to those undertaken by other management and conservation organizations (e.g., Marine Stewardship Council, COSEWIC, Pacific Salmon Commission), require assumptions to estimate spawner and recruit time series for CUs. Given the lack of expert scrutiny of the data for each individual CU under this data-driven approach, it was important to understand how these assumptions may bias status outcomes under a range of biological and management scenarios.

We found that the data-driven biological status assessments were relatively insensitive to common assumptions in expanding spawner abundances within the parameter ranges we explored, but the rate and type of status misclassifications depended on the underlying status of the CU and may be of greater concern for CUs with poor status. To ensure the accuracy of data-driven status assessments, increased efforts to collect data on catch composition, age-at-return, and spawner abundances are needed. Such information will help, for example, to define plausible ranges of error in catch estimation to lend confidence to estimates of recruitment and thus assessments under SR benchmarks. Nonetheless, our research suggests that current efforts to assess status in the face of imperfect and incomplete data are worthwhile for central coast chum salmon and other similar Pacific salmon populations and can provide a timely approach to assessing status for CUs that complements more thorough integrated status assessments.

Acknowledgements

The authors thank the Pacific Salmon Foundation's Population Science Advisory Committee, in particular David Peacock and Randall Peterman, and Steven Cox-Rogers and Charmaine Carr-Harris for their valuable feedback throughout this project. Thanks to Doug Stewart and Kate McGivney for sharing their insights on central coast chum salmon and Brooke Davis and two anonymous reviewers for feedback on earlier drafts. This research was funded by the Pacific Salmon Foundation's Salmon Watersheds Program.

Literature cited

Brown, G.S., Baillie, S.J., Thiess, M.E., Bailey, R.E., Candy, J.R., Parken, C.K., and Willis, D.M. 2019. Pre-COSEWIC review of southern British Columbia Chinook Salmon (*Oncorhynchus tshawytscha*) conservation units, Part I: background. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/11: vii + 67 pp. Available from http://publications.gc.ca/collections/collection_2019/mpo-dfo/fs70-5/Fs70-5-2019-011-eng.pdf [accessed 1 December 2019].

Carwardine, J., Martin, T.G., Firn, J., Reyes, R.P., Nicol, S., Reeson, A., Grantham, H.S., Stratford, D., Kehoe, L., and Chadès, I. 2019. Priority Threat Management for biodiversity conservation: A handbook. J. Appl. Ecol. 56(2): 481–490. doi:10.1111/1365-2664.13268.

Cave, J.D., and Gazey, W.J. 1994. A preseason simulation model for fisheries on Fraser River sockeye salmon (*Oncorhynchus nerka*). Can. J. Fish. Aquat. Sci. **51**(7): 1535–1549. doi:10.1139/f94-153.

Challenger, W., Mochizuki, T., English, K., and Bychkov, Y. 2018. North and Central Coast Salmon Database and Analysis System User Manual. Sidney, BC. Available from https://salmonwatersheds.ca/libraryfiles/lib_449.pdf [accessed 1 December 2019].

Chen, Y., Chen, L., and Stergiou, K.I. 2003. Impacts of data quantity on fisheries stock assessment. Aquat. Sci. 65(1): 92–98. doi:10.1007/s000270300008.

Clark, R., Eggers, D., Munro, A., Fleischman, S., Bue, B., and Hasbrouck, J. 2014. An evaluation of

the percentile approach for establishing sustainable escapement goals in lieu of stock productivity information. Alaska Dep. Fish Game, Fish. Manuscr. No. 14–06. Anchorage, AK. doi:10.4027/amdlfs.2016.08.

Connors, B.M., Pickard, D.C., Porter, M., Farrell, G., Bryan, K., Casely, S., and Huang, S. 2013. Skeena Salmon Conservation Unit Snapshots. Report prepared by ESSA Technologies Ltd. for the Pacific Salmon Foundation. v + 70p. Available from https://salmonwatersheds.ca/libraryfiles/lib_420.pdf [accessed 1 December 2019].

Connors, K., Hertz, E., Jones, E., Honka, L., Kellock, K., and Alexander, R. 2019. The Nass Region: Snapshots of salmon population status. The Pacific Salmon Foundation, Vancouver, BC. Available from https://salmonwatersheds.ca/libraryfiles/lib_453.pdf [accessed 1 December 2019].

Connors, K., Jones, E., Kellock, K., Hertz, E., Honka, L., and Belzile, J. 2018. BC Central Coast: A Snapshot of Salmon Populations and Their Habitats. The Pacific Salmon Foundation, Vancouver, BC. Available from https://salmonwatersheds.ca/library/lib_442/ [accessed 1 December 2019].

COSEWIC. 2016. COSEWIC Assessment and Status Report on the Coho Salmon *Oncorhynchus kisutch*, Interior Fraser population, in Canada. Committee on the Status of Endangered Wildlife in Canada Ottawa, ON. Available from https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments-status-reports/coho-salmon-interior-fraser-2016.html [accessed 1 December 2019].

COSEWIC. 2017. COSEWIC assessment and status report on the Sockeye Salmon *Oncorhynchus nerka*, 24 Designatable Units in the Fraser River Drainage Basin, in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa, On. Available from https://www.sararegistry.gc.ca/virtual_sara/files/cosewic/srSockeyeSalmon2017e.pdf [accessed 1 December 2019].

Cousens, N.B.F., Thomas, G.A., Swann, C.G., and Healey, M.C. 1982. A Review of Salmon Escapement Estimation Techniques. Can. Tech. Rep. Fish. Aquat. Sci. **1108**: 1–128.

Cunningham, C.J., Anderson, C.M., Wang, J.Y.L., Link, M., and Hilborn, R. 2019. A management strategy evaluation of the commercial sockeye salmon fishery in Bristol Bay, Alaska. Can. J. Fish. Aquat. Sci. **76**(9): 1669–1683. doi:10.1139/cjfas-2018-0133.

DFO. 2015. Wild Salmon Policy Biological Status Assessment for Conservation Units of Interior Fraser River Coho Salmon (*Oncorhynchus kisutch*). DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/022. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/364851.pdf [accessed 1 December 2019].

DFO. 2016. Integrated Biological Status of Southern British Columbia Chinook Salmon (*Oncorhynchus tshawytscha*) Under the Wild Salmon Policy. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2016/042. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/40595419.pdf [accessed 1 December 2019].

DFO. 2018a. The 2017 Fraser sockeye salmon (*Oncorhynchus nerka*) integrated biological status re-assessment under the Wild Salmon Policy. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep.

2018/017. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/40712163.pdf [accessed 1 December 2019].

DFO. 2018b. Northern BC Salmon Integrated Fisheries Management Plan. Fisheries and Oceans Canada, Vancouver, BC. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/40797168.pdf [accessed 1 December 2019].

DFO. 2019. Wild Salmon Policy 2018-2022 Implementation Plan: Annual Report 2018-2019. Available from https://www.pac.dfo-mpo.gc.ca/fm-gp/species-especes/salmon-saumon/wsp-pss/annual-annuel/2018-2019-eng.html [accessed 1 December 2019].

Dorner, B., Catalano, M.J., and Peterman, R.M. 2017. Spatial and temporal patterns of covariation in productivity of Chinook salmon populations of the northeastern Pacific Ocean. Can. J. Fish. Aquat. Sci. **75**(7): 1082–1095. doi:10.1139/cjfas-2017-0197.

Dorner, B., Holt, K.R., Peterman, R.M., Jordan, C., Larsen, D.P., Olsen, A.R., and Abdul-Aziz, O.I. 2013. Evaluating alternative methods for monitoring and estimating responses of salmon productivity in the North Pacific to future climatic change and other processes: A simulation study. Fish. Res. **147**: 10-23. doi: 10.1016/j.fishres.2013.03.017

Dorner, B., Peterman, R.M., and Su, Z. 2009. Evaluation of performance of alternative management models of Pacific salmon (Oncorhynchus spp.) in the presence of climatic change and outcome uncertainty using Monte Carlo simulations. Can. J. Fish. Aquat. Sci. **66**(12): 2199–2221. doi:10.1139/F09-144

English, K.K. 2016. Review of escapement indicator streams for the North and Central Coast

910 salmon monitoring program. Report prepared by LGL Limited for the Pacific Salmon
911 Foundation and Fisheries and Oceans Canada. Available from
912 https://salmonwatersheds.ca/libraryfiles/lib_440.pdf [accessed 1 December 2019].

913 English, K.K., Bailey, R.E., and Robichaud, D. 2007. Assessment of Chinook Salmon returns to the
914 Fraser River watershed using run reconstruction techniques, 1982-04. Can. Sci. Advis. Sec.
915 Res. Doc. 2007/020: vi + 76. Available from
916 https://www.psf.ca/sites/default/files/RES2007_020_e.pdf [accessed 1 December 2019].

917 English, K.K., Peacock, D., Challenger, W., and Mochizuki, T. 2016. North and Central Coast
918 Salmon Escapement, Catch, Run Size and Exploitation Rate Estimates for each Salmon
919 Conservation Unit for 1954-2014. Report prepared by LGL Limited for the Pacific Salmon
920 Foundation and Fisheries and Oceans Canada. Available from
921 https://salmonwatersheds.ca/libraryfiles/lib_435.pdf [accessed 1 December 2019].

922 English, K.K., Peacock, D., Challenger, W., Noble, C., Beveridge, I., Robichaud, D., Beach, K.,
923 Hertz, E., and Connors, K. 2018. North and Central Coast Salmon Escapement, Catch, Run
924 Size and Exploitation Rate Estimates for each Salmon Conservation Unit for 1954-2017.
925 Report prepared by LGL Limited for the Pacific Salmon Foundation. Available from
926 https://salmonwatersheds.ca/library/lib_451/ [accessed 1 December 2019].

927 Fisheries and Oceans Canada. 2005. Canada's Policy for Conservation of Wild Pacific Salmon.
928 Vancouver, BC. Available from https://www.pac.dfo-mpo.gc.ca/fm-gp/species-
929 especes/salmon-saumon/wsp-pss/policy-politique/index-eng.html [accessed 1 December

2019].

Fleischman, S.J., Catalano, M.J., Clark, R.A., and Bernard, D.R. 2013. An age-structured state-space stock recruit model for Pacific salmon Oncorhynchus spp. Can. J. Fish. Aquat. Sci. **70**: 401–414. doi:10.1139/cjfas-2012-0112.

Foster, K.R., Vecchia, P., and Repacholi, M.H. 2000. Science and the Precautionary Principle. Science **288**(5468): 979–981. Available from https://www.jstor.org/stable/3075092 [accessed 1 December 2019].

Hicks, A., Carpi, P., and Berukoff, S. 2020. Technical Details of the International Pacific Halibut Commission Management Strategy Evaluation Framework. Report number IPHC-2020-SRB016-INF01. Available from https://iphc.int/uploads/pdf/srb/srb016/iphc-2020-srb016-inf01.pdf [Accessed 28 July 2020].

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment and management: choice, dynamics, and uncertainty. Routledge, Chapman & Hall, Inc., New York.

Holt, C., Davis, B., Dobson, D., Godbout, L., Luedke, W., Tadey, J., and Van Will, P. 2018. Evaluating Benchmarks of Biological Status for Data-limited Populations (Conservation Units) of Pacific Salmon, Focusing on Chum Salmon in Southern BC. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/011: ix + 87. Available from http://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2018/2018_011-eng.html [accessed 1 December 2019].

Holt, C.A., Brooke, D., and Cox-Rogers, S. 2016. Adapting benchmarks of biological status for

persistent changes in productivity and variability in exploitation history with a focus on data-limited populations (Conservation Units) of Pacific salmon in north and central BC. Draft final report for the PSC. Can. Sci. Advis. Secr. Res. Doc. 86. Available from https://www.psc.org/fund-project/adapting-benchmarks/ [accessed 1 December 2019].

Holt, C.A., Cass, A., Holtby, L.B., Riddell, B.E., and Holtby, B. 2009. Indicators of status and benchmarks for Conservation Units in Canada's Wild Salmon Policy. Can. Sci. Advis. Sec. Res. Doc. 2009/058: viii+74. Available from https://www.psf.ca/sites/default/files/339096.pdf [accessed 1 December 2019].

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. doi:10.1016/j.fishres.2015.01.002.

Holt, C.A., and Ogden, A. 2013. Software for Assessing Status of Conservation Units under Canada's Wild Salmon Policy: Instructional Manual. Can. Tech. Rep. Fish. Aquat. Sci. *In Prep.*

Holt, C.A., and Peterman, R.M. 2006. Missing the target: uncertainties in achieving management goals in fisheries on Fraser River, British Columbia, sockeye salmon (*Oncorhynchus nerka*). Can. J. Fish. Aquat. Sci. **63**(12): 2722–2733. doi:10.1139/F06-155

Holt, C.A., and Peterman, R.M. 2008. Uncertainties in population dynamics and outcomes of regulations in sockeye salmon (*Oncorhynchus nerka*) fisheries: implications for

970 management. Can. J. Fish. Aquat. Sci. **65**(7): 1459–1474. Available from
971 https://www.nrcresearchpress.com/doi/full/10.1139/F08-053 [accessed 1 December
972 2019].

973 Holtby, L.B., and Ciruna, K.A. 2007. Conservation Units for Pacific salmon under the Wild
974 Salmon Policy. Can. Sci. Advis. Sec. Res. Doc. 2007/070: viii+74. Available from
975 http://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2007/2007_070-
976 eng.htm [accessed 1 December 2019].

977 IUCN. 2017. Guidelines for Using the IUCN Red List - Categories and Criteria. Version 6.1.
978 Available from https://www.iucnredlist.org/resources/redlistguidelines [accessed 13 April
979 2020].

980 Kehler, D.G., Myers, R.A., and Field, C.A. 2002. Measurement error and bias in the maximum
981 reproductive rate for the Ricker model. Can. J. Fish. Aquat. Sci. **59**(5): 854–864.
982 doi:10.1139/f02-060.

983 Kope, R. 2006. Cumulative effects of multiple sources of bias in estimating spawner-recruit
984 parameters with application to harvested stocks of Chinook salmon (*Oncorhynchus*
985 *tshawytscha*). Fish. Res. **82**(1–3): 101–110. doi:10.1016/j.fishres.2006.08.016.

986 Korman, J., and English, K. 2013. Benchmark Analysis for Pacific Salmon Conservation Units in
987 the Skeena Watershed. A report prepared for the Pacific Salmon Foundation, Vancouver,
988 BC. Available from https://salmonwatersheds.ca/libraryfiles/lib_318.pdf [accessed 1
989 December 2019].

Korman, J., Peterman, R.M., and Walters, C.J. 1995. Empirical and theoretical analyses of correction of time-series bias in stock-recruitment relationships of sockeye salmon (*Oncorhynchus nerka*). Can. J. Fish. Aquat. Sci. **52**(10): 2174–2189. doi:10.1139/f95-810.

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.

Martin, T.G., Kehoe, L., Mantyka-Pringle, C., Chades, I., Wilson, S., Bloom, R.G., Davis, S.K., Fisher, R., Keith, J., Mehl, K., Diaz, B.P., Wayland, M.E., Wellicome, T.I., Zimmer, K.P., and Smith, P.A. 2018. Prioritizing recovery funding to maximize conservation of endangered species. Conserv. Lett. **11**(6): 1–9. doi:10.1111/conl.12604.

Office of the Auditor General of Canada. 2004. 2005 Report of the Commissioner of the Environment and Sustainable Development. In Fisheries and Oceans Canada — salmon stocks, habitat, and aquaculture. Chapter 5 [online]. Available from http://www.oag-bvg.gc.ca/internet/docs/c20041005ce.pdf [accessed 8 April 2019].

Peacock, S.J., Connors, B.M., Krkošek, M., Irvine, J.R., and Lewis, M.A. 2014. Can reduced predation offset negative effects of sea louse parasites on chum salmon? Proc. R. Soc. B Biol. Sci. **281**: 20132913. doi:10.1098/rspb.2013.2913.

Peacock, S.J., and Holt, C.A. 2012. Metrics and sampling designs for detecting trends in the distribution of spawning Pacific salmon (*Oncorhynchus* spp.). Can. J. Fish. Aquat. Sci. **69**(4): 681–694. doi:10.1139/f2012-004.

Peterman, R.M. 2018. Continuous learning, teamwork, and lessons for young scientists. ICES J. Mar. Sci. **76**(1): 28–40. doi:10.1093/icesjms/fsy141.

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**(8): 1255–1260. doi:10.1139/F2012-063.

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**(1): 181–191. doi:10.1139/cjfas-57-1-181.

Price, M.H.H., Darimont, C.T., Temple, N.F., and MacDuffee, S.M. 2008. Ghost runs: management and status assessment of Pacific salmon (*Oncorhynchus* spp.) returning to British Columbia's central and north coasts. Can. J. Fish. Aquat. Sci. **65**: 2712–2718. doi: 10.1139/F08-174.

Price, M.H.H., English, K.K., Rosenberger, A.G., MacDuffee, M., and Reynolds, J.D. 2017. Canada's Wild Salmon Policy: an assessment of conservation progress in British Columbia. Can. J. Fish. Aquat. Sci. **74**(10): 1507–1518. doi:10.1139/cjfas-2017-0127.

PSC. 2019. Treaty Between the Government of Canada and the Government of the United States of America Concerning Pacific Salmon. Pacific Salmon Commission, Vancouver, BC. Available from http://www.psc.org/ [accessed 1 December 2019].

Punt, A.E. 1992. Selecting management methodologies for marine resources, with an illustration for Southern African hake. South African J. Mar. Sci. **12**(1): 943–958.

doi:10.2989/02577619209504754.

Pyper, B.J., and Peterman, R.M. 1998. Comparison of methods to account for autocorrelation in correlation analyses of fish data. Can. J. Fish. Aquat. Sci. **55**(9): 2127–2140. doi:10.1139/cjfas-55-9-2127.

Ricker, W.E. 1954. Stock and recruitment. J. Fish. Res. Board Canada **11**(5): 559–623. Available from https://www.nrcresearchpress.com/doi/abs/10.1139/f54-039 [accessed 1 December 2019].

Salmon Watersheds Program - Pacific Salmon Foundation. 2019. Salmon Watersheds Data Library. Available from http://data.salmonwatersheds.ca/data-library/ [accessed 1 December 2019].

Scheuerell, M.D. 2016. An explicit solution for calculating optimum spawning stock size from Ricker's stock recruitment model. PeerJ **4**(1): e1623. doi:10.7717/peerj.1623.

Schnute, J.T., Boers, N., and Haigh, R. 2015. PBSmapping: Mapping Fisheries Data and Spatial Analysis Tools. Available from https://cran.r-project.org/package=PBSmapping [accessed 1 December 2019].

Schnute, J.T., and Richards, L.J. 1995. The influence of error on population estimates from catch-age models. Can. J. Fish. Aquat. Sci. **52**(10): 2063–2077. doi:10.1139/f95-800.

Staton, B.A., Catalano, M.J., Connors, B.M., Coggins, L.G., Jones, M.L., Walters, C.J., Fleischman, S.J., and Gwinn, D.C. 2020. Evaluation of methods for spawner-recruit analysis in mixed-stock Pacific salmon fisheries. Can. J. Fish. Aquat. Sci. (*in press*): 1–43. doi:10.1139/cjfas-

2019-0281.

Subbey, S., Devine, J.A., Schaarschmidt, U., and Nash, R.D.M. 2014. Modelling and forecasting stock–recruitment: current and future perspectives. ICES J. Mar. Sci. **71**(8): 2307–2322.

Szerlong, R.G., and Rundio, D.E. 2008. A statistical modeling method for estimating mortality and abundance of spawning salmon from a time series of counts. Can. J. Fish. Aquat. Sci. **65**: 17–26. doi:10.1139/F07-141.

Walsh, J.C., Connors, K., Hertz, E., Kehoe, L., Martin, T.G., Connors, B., Bradford, M.J., Freshwater, C., Frid, A., Halverson, J., Moore, J.W., Price, M.H.H., and Reynolds, J.D. (2020). Prioritizing conservation actions for Pacific salmon in Canada. J. Appl. Ecol. (*in press*): 1–37. doi:10.1101/2020.02.03.931691.

Walters, C., English, K., Korman, J., and Hilborn, R. 2019. The managed decline of British Columbia's commercial salmon fishery. Mar. Policy **101**: 25–32. doi:10.1016/j.marpol.2018.12.014.

Walters, C.J. 1985. Bias in the Estimation of Functional Relationships from Time Series Data. Can. J. Fish. Aquat. Sci. **42**(1): 147–149. doi:10.1139/f85-018.

Walters, C.J. 1986. Adaptive Management of Renewable Resources. MacMillan Publishing Company, New York.

Walters, C.J., and Ludwig, D. 1981. Effects of Measurement Errors on the Assessment of Stock-Recruitment Relationships. Can. J. Fish. Aquat. Sci. **38**(6): 704–710.

Wetzel, C.R., and Punt, A.E. 2011. Performance of a fisheries catch-at-age model (Stock Synthesis) in data-limited situations. Mar. Freshw. Res. **62**(8): 927. doi:10.1071/mf11006.

Zabel, R.W., and Levin, P.S. 2002. Simple assumptions on age composition lead to erroneous conclusions on the nature of density dependence in age-structured populations. Oecologia **133**(3): 349–355. doi:10.1007/s00442-002-1051-0.

Tables

1075 *Table 1. Summary of common steps in run reconstructions (Figure 2) and associated assumptions and potential biases that we investigated.*

Step in run recon-struction	**Description**	**Assumption**	**Potential biases**	**Associated factor(s) investigated in this study**
(1) Expansion Factor I	Infills observed spawner abundances in indicator streams to account for indicator streams that were not monitored in a given year.	The contribution of each indicator-stream population to total escapement is constant within a decade.	May be biased if contributions are changing due to, for example, changes in capacity or productivity that differ among populations. Under low monitoring coverage, the magnitude of expansion is greater and thus we expect that any potential bias would be larger.	→ Diverging capacities of populations through time → Declining monitoring coverage
(2) Expansion Factor II	Expands spawner abundances from (1) to include spawners in non-indicator streams.	The contribution of non-indicator-stream populations to total escapement is constant within a decade.	May be biased if contributions are changing due to, for example, changes in capacity or productivity that differ between populations in indicator and non-indicator streams.	→ Diverging capacities of populations through time → Declining monitoring coverage → Number of streams and proportion that are

				indicators
(3) Expansion Factor III	Expands spawner abundances from (2) to account for observer efficiency and for populations that are never monitored, yielding the estimated total spawner abundance for the CU.	The proportion of total spawner abundance estimated in (2) is known, constant over time, and independent of spawner abundance.	May be biased if observer efficiency is not known or poorly estimated, if survey methods change over time (e.g., a weir to overhead counts), or if observation bias varies with total spawner abundance.	→ Bias in the observation of spawners (under the same value of Expansion Factor III) → Declining monitoring coverage
(4) Catch assignment to CUs	Catch from PFMAs is assigned to CUs in proportion to the spawner abundance for that CU.	Fish caught in an PFMA were destined to spawn in streams that flow into that PFMA.	Over- or under-estimation of catch due to different run timing among CUs that flow into the same PFMA (if information on run timing is uncertain, unavailable, or not incorporated into run-reconstruction models).	→ Bias in the observation of catch

(5) Calculating recruitment using age-at-return	The total return to the CU is assigned to brood years based on the proportion of fish returning at different ages.	Often, annual age-at-return data are not available for each CU, and so age-at-return is assumed to be constant over time, using the average of available data.	Variability in brood year recruitment will be underestimated if there is high inter-annual variability, or temporal changes, in age-at-return that is not accounted for in assessments.	→ Inter-annual variability in age-at-return

1076 *Table 2. Summary of factors that we investigated in sensitivity analyses to determine their impact on bias in status*
1077 *assessments.*

Factor	Scenario	Details
Number of populations and proportion that spawn in indicator streams	Central coast chum*	35 populations, 15 (43%) indicator streams
	Small/Low	10 populations, 3 (30%) indicator streams
	Small/High	10 populations, 8 (80%) indicator streams
	Large/Low	140 populations, 42 (30%) indicator streams
	Large/High	140 populations, 119 (85%) indicator streams
Monitoring coverage	Constant	**Indicator**: historical 76% with no change
		Non-indicator: historical 72% with no change
	Observed decline*	**Indicator**: 76% with a change of -5% over the last 26 years
		Non-indicator: 72% with a change of -67% over the last 26 years
Declines in capacity†	0	All spawning populations have stable capacities
	25	25% of populations severe and 25% moderate declines, and 50% stable capacity
	50	50% of populations severe and 50% moderate declines in capacity
	100	100% of populations display severe declines in capacity
Bias in observation of spawners	Range in bias from $\bar{\delta}$ = -1.6, which would correspond to the maximum value of Expansion Factor III that has been applied (F''' = 5.0,; English et al. 2018) to $\bar{\delta}$ = 0 in increments of 0.2 (*default value $\bar{\delta}$ = -0.4, corresponding to F'''=1.5)	
Bias in observation of catch	Range in bias from $\bar{\chi}$ = -1.0 (63% underestimation) to $\bar{\chi}$ = 1.0 (271% overestimation) in increments of 0.2 (*default value $\bar{\chi}$ = 0).	
Interannual variability in age-	Range in variability from $\bar{\omega}$ = 0.2 to $\bar{\omega}$ = 1.6 in increments of 0.2 (*default value $\bar{\omega}$ = 0.8; Figure S4)	

at-return

*Default values. †The default values for decline in capacity did not correspond exactly to the scenarios considered in the sensitivity analyses, but were based on habitat assessments for central coast (21% of populations having severe declines and 29% of populations having moderate declines (Connors et al. 2018)).

Figure captions

Figure 1. Illustration of the WSP status assessment framework (adapted from Holt et al. 2009). We focused on the geometric mean spawner abundance (metric, blue) under the spawner abundance indicator. This metric was assessed against two types of benchmarks: spawner-recruitment (SR) and percentile (see Figure 2). Faded boxes represent other types of metrics and indicators that may be included in integrated status assessments but were beyond the scope of what we considered.

Figure 2. a) benchmarks are the 50th (horizontal green line) and 25th (horizontal red line) percentiles of historical spawner abundance (points). The current spawner abundance is calculated as the geometric mean spawner abundance over the most recent generation (4 years, blue points and line). b) -recruitment benchmarks are based on the shape of the Ricker relationship (solid line) fit to data on spawner abundance (x-axis) and corresponding recruitment (catch + escapement, y-axis). The upper and lower benchmarks are 80% S_{MSY} (green) and S_{GEN} (red), respectively. S_{GEN} is defined as the spawner abundance that leads to S_{MSY} (grey) in one generation in the absence of fishing mortality. Under both types of benchmarks, the current spawner abundance in the example shown is above the upper benchmark, and this CU would be assessed as 'green'. c) The calculation of benchmarks and benchmarks requires run reconstruction to expand observed spawners abundances, assign catch to CUs, and calculate recruitment (Table 1).

Figure 3. Our study focused on central coast chum salmon Conservation Units (CUs) (green; different CUs shaded differently in central coast inset),which are relatively large and do not correspond to the Pacific Fishery Management Areas (PFMAs; white or light blue shaded regions) for which catch is reported. One factor we investigated was how bias in assigning catch to CUs, resulting in under- or over-estimation of catch, affects estimates of biological status. Map produced using PBSmapping (Schnute et al. 2015) with data from Martin Huang (pers. comm. <Martin.Huang@dfo-mpo.gc.ca>).

Figure 4. Schematic of the simulation model composed of submodels for population dynamics (including harvest), observation, assessment, and performance. Adapted from Holt et al. (2016).

Figure 5. The two target harvest-rate cases we simulated were (1) a simple Harvest Control Rule (eq. (4); solid line) with parameters estimated from historical harvest rates and total return from five central coast chum CUs (grey points), (2) a constant high target harvest rate of $ht' = 0.6$ (dotted line).

Figure 6. Estimated status according to the spawner-recruitment (SR) benchmarks (left) and the percentile benchmarks (right), over the true status for each of two cases (high productivity and an HCR (a-b) and low productivity and high harvest rates (c-d)). Grey cells indicate pessimistic misclassifications, which may lead to overly conservative management actions, and black cells indicate optimistic misclassifications, which may lead to overly risky management actions. Coloured cells indicate correct classifications for red, amber, and green zones, respectively.

Figure 7. Relative bias in spawner-recruitment (SR) benchmarks (a), percentile benchmarks (b), and current spawner abundance (S_{AVG}; black) (median $\pm$ interquartile range among 4000 MC

simulations) for high productivity and a harvest control rule (HCR) and the low-productivity high-harvest case.

Figure 8. The effect of monitoring coverage (no change and decline; Table 2) and the percentage of spawning populations with severe declines in capacity (x-axis) on performance measures under the base case of high productivity and HCR. (a-b) The proportion of simulations with correct green, amber, or red status or pessimistic misclassifications (grey) and optimistic misclassifications (black) under the SR benchmarks (a) and percentile benchmarks (b). (c-d) The percent relative bias (median $\pm$ interquartile range among 4000 MC simulations) in the current spawner abundance (S_{AVG}; black) and lower and upper benchmarks (red and green, respectively) under the SR benchmarks (c) and the percentile benchmarks (d). See Online Supplement for results under the low-productivity high-harvest case.

Figure 9. The effect of observation bias in the number of spawners (x-axis) on performance measures under the base case. (a-b) The proportion of simulations with correct green, amber, or red status or pessimistic misclassifications (grey) and optimistic misclassifications (black) under the SR benchmarks (a) and percentile benchmarks (b). (c-d) The percent relative bias (median $\pm$ interquartile range among 4000 MC simulations) in the current spawner abundance (S_{AVG}; black circle) and lower and upper benchmarks (red and green, respectively) under the SR benchmarks (c) and the percentile benchmarks (d). The asterisk in (a-b) and grey zone in (c-d) indicate the default parameter value of δ = -0.4, and the bias that matches the Expansion Factor III of F''' = 1.5 applied in all simulations. See Online Supplement for results under the low-productivity high-harvest case (results were similar).

Figure 10. The effect of observation bias in catch (x-axis) on performance measures under the base case. (a) The proportion of simulations with correct green, amber, or red status or pessimistic misclassifications (grey) and optimistic misclassifications (black) under the SR benchmarks. (b) The percent relative bias (median $\pm$ interquartile range among 4000 MC simulations) in the current spawner abundance (S_{AVG}; black circle) and lower and upper benchmarks (red and green, respectively) under the SR benchmarks. The asterisk in (a) and grey zone in (b) indicate the default parameter value of χ = 0. Observation bias in catch ranges from 63% underestimation (χ = -1) to 271% overestimation (χ = 1). See Online Supplement for results under the low-productivity high-harvest case.

Appendices

Appendix A: Table of default parameter values

Table A1. Default values for parameters in the simulation model that were used unless otherwise specified (e.g., in sensitivity analyses). See text for further explanation of the values and the Online Supplement for details of estimation for those based on raw data. For parameters that were part of sensitivity analyses, the range in parameter values that was explored is highlighted.

Submodel	**Parameter**	**Symbol**	**Default value**	**Range**	**Ref.**
General	Years over which the simulation is run	t : return year y : brood year	50	-	
	Number of indicator streams	I	15	1 – 41	1
	Number of non-indicator streams	J	20	1 - 100	1
Population dynamics	Productivity $a_i \sim N(\bar{a}, \sigma_a^2)$	$\bar{a}$	1.40	-	2
		σ_a^2	0.22	-	2
	Density dependence (initial) $b_{i,t=0} = 1/S_{MAX_i}$ $\log S_{MAX_i} \sim N(\mu_{b,I}, \sigma_{b,I}^2)$	$\mu_{b,I}$	7.95	-	2
		$\sigma_{b,I}^2$	2.07	-	2
		$\mu_{b,J}$	6.95	-	2
		$\sigma_{b,J}^2$	1.39	-	2

	Temporal autocorrelation	τ	0.422	-	2
	Variance in recruitment deviates within spawning populations	σ_{υ}^{2}	1.28	-	2
	Correlation among spawning populations in recruitment deviates	ρ	0.46	-	2
	Average proportions for age-at-maturity	$\overline{p}_{g}$	$\overline{p}_{3}$= 0.23 $\overline{p}_{4}$= 0.64 $\overline{p}_{5}$= 0.13	-	3
	Interannual variability in age-at-maturity	$\overline{\omega}$	0.8	-	4
	Maximum target harvest rate	h'_{MAX}	0.42	-	5
	Shape parameter for that harvest rule	m	1.13×10^{-5}		5
	Standard deviation in outcome uncertainty around harvest rate	σ_{h}	0.13	-	5
Observation	Lognormal bias in observed spawners	$\bar{\delta}$	- 0.4	-1.6 to 0.0	6
	Lognormal variance in observed spawners	σ_{δ}^{2}	0.25	-	4

	Lognormal bias in observed catch	$\bar{\chi}$	0	-1.0 to 1.0	
	Lognormal variance in observed catch	σ_{χ}^{2}	0.04	-	4
	Variability in observed age-at-return	$\overline{\omega}_{r}$	0.1	-	4

References:

1. Based on the range of indicator and non-indicator streams reported in the Pacific Salmon Explorer (www.salmonexplorer.ca) for the eight central coast chum CUs.

2. Calculated from river-level stock-recruitment data for central coast chum CUs. See Online Supplement for details.

3. From the NCCDBV2 (Challenger et al. 2018).

4. Same as assumed in Holt et al. (2018) for south-coast chum salmon.

5. Calculated from CU-level harvest rates and total return size, from the Salmon Watersheds Data Library (Salmon Watersheds Program - Pacific Salmon Foundation 2019). See Online Supplement for details.

6. Based on expert opinion. Expansion Factor of F’’’ = 1.5 in Table A3-A4 of English et al. (2016).