Fisheries Research 171 (2015) 170–187

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Evaluating the performance of data-moderate and catch-only assessment methods for U.S. west coast groundfish

Chantel R. Wetzel[a,b,*], André E. Punt[b]

[a] Northwest Fisheries Science Center, National Marine Fisheries Service, 2725 Montlake Boulevard East, Seattle, WA 98112, United States
[b] School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020, United States

ARTICLE INFO

Article history:
Received 27 March 2014
Received in revised form 4 June 2015
Accepted 9 June 2015
Available online 2 July 2015

Keywords:
Data-limited
Data-moderate
U.S. west coast
Groundfish
Depletion Corrected Average Catch
Depletion-Based Stock Reduction Analysis
Extended Depletion-Based Stock Reduction Analysis
Extended Simple Stock Synthesis
Management strategy evaluation

ABSTRACT

The estimation of sustainable harvest limits for stocks that have never been assessed and have limited data available can be challenging. Harvest limits for previously un-assessed U.S. west coast groundfish have been set using catch-only methods (Depletion Corrected Average Catch [DCAC] and Depletion-Based Stock Reduction Analysis [DB-SRA]) for data-limited stocks, as well as catch and index based methods (Extended Depletion-Based Stock Reduction Analysis [XDB-SRA] and Extended Simple Stock Synthesis [XSSS]) for data-moderate stocks. To account for uncertainty and to prevent overfishing the harvest levels for U.S. west coast groundfish are reduced based upon the estimation method and the amount of data used. A management strategy evaluation was applied to evaluate the performance of each estimation method to provide benchmark harvest levels for two life-history types (U.S. west coast flatfish and rockfish) under varying misspecifications of the parameter distributions used by these methods. Both, data-moderate and catch-only (data-limited) methods, resulted in overfishing > 0.50 (except XSSS in select scenarios) when the simulated stock was below the relative biomass target for both life-histories. Each of the data-moderate methods (XDB-SRA and XSSS) that applied biomass index data failed to estimate the correct stock status in the first assessment when an overly optimistic prior distribution about the stock status was assumed. However, during the projection period as the biomass index lengthened, the estimates of current stock size improved for both of these estimation methods, reducing the probability of overfishing to <0.50 (except XSSS for one scenario). The ability to incorporate index data by the data-moderate methods resulted in improved estimates, as the data became more informative, for stock status and the subsequent harvest limits, that when reduced to account for uncertainty resulted in population stock sizes that either remain stable or rebuild toward the biomass targets for both life-histories. A notable exception was the performance of XDB-SRA for the flatfish life-history when the stock was at the target level and the prior was assumed correctly. In this instance the index data were non-informative, resulting in overfishing due to overly optimistic estimates of relative population size.

Published by Elsevier B.V.

1. Introduction

The Pacific Fishery Management Council (PFMC) manages federal fisheries off the U.S. west coast. Currently, the groundfish fishery management plan includes 90 + species, of which only approximately a third have been formally assessed (PFMC, 2014a). The Council classifies groundfish into three categories based upon the type of assessment, and thus, harvest specification uncertainty: (1) category 1 stocks, where complex age-structured stock assessments that incorporate biomass indices and composition data (length and or age data) were applied and are assumed to have the lowest level of uncertainty regarding status; (2) category 2 stocks, where some biological indicators are present (e.g., index data) and estimation methods that incorporate some limited data were applied and have a moderate level of uncertainty about stock status; and (3) category 3 stocks, which have limited data (e.g., catch) and harvest estimates are calculated using catch-based statistics and have a high degree of uncertainty surrounding stock status. The level of uncertainty surrounding stock status is accounted for when setting harvest limits. Determining a harvest limit is typically done by two steps: (1) estimating an overfishing limit (OFL), a level of harvest that if exceeded would constitute overfishing and defines the maximum harvest level; and (2) setting an acceptable biological catch (ABC), a level of harvest less or equal

* Corresponding author at: Northwest Fisheries Science Center, National Marine Fisheries Service, 2725 Montlake Boulevard East, Seattle, WA 98112, United States. Tel.: +1 206 302 1753; fax: +1 206 860 6792.
E-mail address: Chantel.Wetzel@noaa.gov (C.R. Wetzel).

http://dx.doi.org/10.1016/j.fishres.2015.06.005
0165-7836/Published by Elsevier B.V.

to the OFL where reductions account for scientific uncertainty in the OFL. The PFMC has termed the reduction between the OFL and the ABC as the "buffer", where the size of that buffer should be directly related to the level of uncertainty about the status, which the council bases on the category of the stock (PFMC, 2010).

In 2010, the PFMC took the first step to establish harvest limits for the previously un-assessed stocks included within their fishery management plan by applying catch-only harvest estimation methods, Depletion-Based Stock Reduction Analysis (DB-SRA [Dick and MacCall, 2011]) and Depletion Corrected Average Catch (DCAC [MacCall, 2009]) (Dick and MacCall, 2010). These methods are applicable in "data-free" situations because they require only limited information to estimate harvest levels. DB-SRA uses a full catch time-series, a pre-specified distribution of relative biomass, and biological parameter distributions to determine the unfished biomass level required to produce the pre-specified ending relative biomass level resulting in estimates of the OFL, while DCAC calculates a one-time estimate of the distribution for a yield that would likely be sustainable given the sum of the known catch history and assumed biological parameter distributions. Stocks where harvest levels have been estimated by either DB-SRA or DCAC are considered to have the highest uncertainty about their status (classified as category 3) and subject to the largest reduction between the OFL and the ABC. While this was an important first step in calculating ABCs for West Coast groundfish, ideally the estimation of OFLs would be based on methods that incorporate data where available. The benefit of applying data based estimation methods would be an improved estimate of stock status with a reduced uncertainty, resulting in less of a reduction between the OFL and ABC, minimizing "lost harvest". Currently, stocks with high uncertainty (category 3) result in ABC values that are reduced from the OFL by a buffer value of 0.69 (ABC = 0.69 × OFL), while stocks with a moderate level of uncertainty (category 2) apply a buffer value of 0.83 to reduce the OFLs.[1]

In 2013, two estimation methods that incorporate trend information, Extended Depletion-Based Stock Reduction Analysis (XDB-SRA) and Extended Simple Stock Synthesis (XSSS), were developed and applied to West Coast groundfish to estimate stock status and determine OFLs (Cope et al., 2015). Each method, similar to DCAC and DB-SRA, apply user-defined biological parameter and relative biomass prior distributions, but XDB-SRA and XSSS incorporate biomass indices to update these prior distributions. The incorporation of indices classifies these estimation methods as data-moderate (category 2) stock assessments, as defined by the PFMC for West Coast groundfish. The underlying structure of XDB-SRA mimics DB-SRA by applying a delay-difference model that assumes equal growth between the sexes with age-based knife-edge maturity and selectivity, but XDB-SRA allows for updating of the prior distributions based upon model fits to observed index data (PFMC, 2014b). Extended Simple Stock Synthesis (XSSS) involves a simplified implementation of Stock Synthesis, an integrated statistical catch-at-age model (Methot and Wetzel, 2013). XSSS is an age-structured model that assumes length-based maturity (although it can be parameterized to be age-based) with selectivity equal to maturity (PFMC, 2014b). Similar to XDB-SRA, XSSS calculates model fits to observed index data, which allows for updating of the prior distributions to estimate stock status and harvest levels.

One of the prior distributions that must be specified by the user for each of these data-moderate and catch-only (data-limited) approaches is the relative biomass level in a specific year, a quantity that is often an estimated result from an assessment. Several simulation studies have evaluated the performance of the catch-only methods (DCAC and DB-SRA) in regards to the assumed prior distributions at estimating sustainable levels of harvests (Wetzel and Punt, 2011; Carruthers et al., 2014). Wetzel and Punt (2011) determined that the OFL can be overestimated when the prior distribution for relative biomass is specified higher than the true value. However, catch-only assessments (DCAC and DB-SRA) are subject to the greatest reduction in the OFL by a buffer value of 0.69 to account for uncertainty. Therefore, if this reduction is large enough, prior misspecification does not necessarily lead to overfishing when the estimated OFL is too great.

Similar to the catch-only methods, the data-moderate approaches, XDB-SRA and XSSS, also apply an assumed prior distribution for relative biomass. However, XDB-SRA and XSSS differ from the catch-only methods by fitting to biomass indices to update the prior distribution based on the data. If the indices are informative, this process should reduce the influence of the user-specified priors, even when misspecified, resulting in posterior distributions based on the data and improving the OFL estimates. Data-moderate estimated OFLs are reduced by 0.83 to set ABCs, resulting in potentially higher harvest limits relative to the catch-only methods due to the reduced uncertainty about stock status and therefore potentially result in increased harvest limits relative to catch-only methods.

This paper applies a management strategy evaluation approach to evaluate the performance of data-moderate (XDB-SRA and XSSS) and catch-only estimation methods (DCAC and DB-SRA) and the subsequent ABC calculations for the management of West Coast groundfish stocks. The estimation methods are used to set ABCs for two simulated stocks, a flatfish and a rockfish, for a 25-year projection period. The status of the operating model population and error in parameter estimates at the end of the projection period are evaluated, and compared among methods to address three questions: (1) are the reductions applied to the OFLs sufficient to prevent overfishing for the data-moderate and catch-only methods; (2) what are the relative risks when parameters are misspecified for both data-moderate and catch-only methods; and, finally, (3) what are the potential trade-offs for performing either a data-moderate or catch-only assessment?

2. Materials and methods

2.1. General approach

The population was modeled using an age-structured operating model. An annual biomass index was observed with error for selected years, and was used by each index-based estimation method (XDB-SRA and XSSS) to estimate population size and OFLs. The estimated OFLs were then adjusted according to a pre-determined buffer (data-moderate 0.83 and catch-only [data-limited] 0.69) to determine ABCs. U.S. federally managed fisheries are require to set an annual catch limit (ACL) value, which can be set equal to the ABC or reduced further to account for additional uncertainty (e.g., management). In this simulation the ACL was set equal to the ABC and then applied to the simulated stock. For clarity, the level of harvest removed from the stock will be referred to as the ABC rather than an ACL. The data generation, OFL estimation and stock updating were conducted in an iterative fashion for 25 years (Fig. 1).

Two life-history types that are common to U.S. west coast groundfish were simulated; a fast growing, short-lived flatfish and a slow growing, long-lived rockfish (Table 1).

[1] The buffer values here are the default values applied by the PFMC based upon stock category. However, the PFMC has applied alternative buffer values that resulted in larger reductions between the ABC and OFL for stocks within each category when a higher level of precaution was deemed warranted.

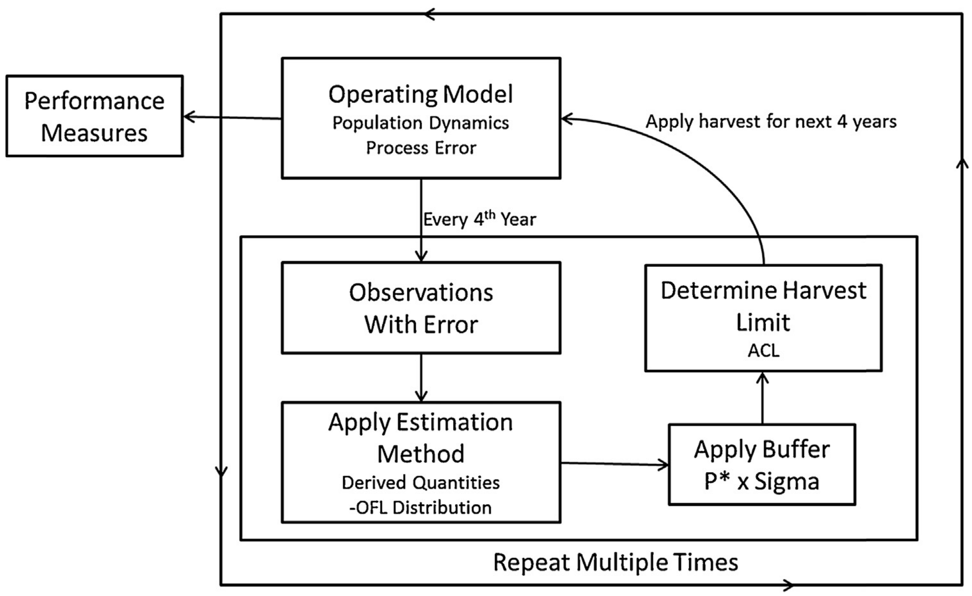

Fig. 1. The process and order of operations for the management strategy evaluation.

Table 1
Life-history parameters used in the operating model for the flatfish and rockfish life-history.

Parameter	Equation form	Rockfish life-history		Flatfish life-history	
		Values		Values	
		Males	Females	Males	Females
Natural mortality (M)	M = constant	$M = 0.05$ (yr^{-1})		$M = 0.20$ (yr^{-1})	
Steepness (h)		$h = 0.511$		$h = 0.875$	
Compensation (c)		$c = 1.5$		$c = 1.75$	
Fishing rate at B_{MSY} (F_{MSY})		$F_{MSY} = 0.045$yr^{-1}		$F_{MSY} = 0.27$ yr^{-1}	
Mean length at age in cm ($L_{\gamma,a}$)[a]	$L_{\gamma,a} = L_{\infty,\gamma} + \left(L_{1,\gamma} - L_{\infty,\gamma}\right) e^{-K(a-a_3)}$	$a_3 = 1$ (yr)		$a_3 = 2.833$ (yr)	
Mean Asymptotic size in cm ($L_{\infty,\gamma}$)	$L_{\infty,\gamma} = L_{1,\gamma} + \frac{L_{2,\gamma}-L_{1,\gamma}}{1-e^{-K_\gamma(a_4-a_3)}}$	$a_4 = 80$ (yr)		$a_4 = 17.833$ (yr)	
Reference ages (a_3, a_4)		$L_{1,\gamma} = 6.64$ (cm)		$L_{1,\gamma} = 24.6210$ (cm)	
Mean length at a_3 ($L_{1,\gamma}$)					
Mean length at a_4 ($L_{2,\gamma}$)		$L_{2,\gamma} = 59.7094$ (cm)	$L_{2,\gamma} = 59.844$ (cm)	$L_{2,\gamma} = 40.6664$ (cm)	$L_{2,\gamma} = 55.4099$ (cm)
Growth coefficient (K_γ)		$K_\gamma = 0.2579$ (yr^{-1})	$K_\gamma = 0.1314$ (yr^{-1})	$K_\gamma = 0.299488$ (yr^{-1})	$K_\gamma = 0.14375$ (yr^{-1})
Coefficient of variation of length-at-age ($\sigma_{0\ \gamma a}$)	$\sigma_{0\gamma a} = \tilde{L}_{\gamma a} * CV$	$CV = 0.08$		$CV = 0.08$	
Body weight ($w_{l,\gamma}$)	$w_{l,\gamma} = \Omega_1 L_l^{\Omega_2}$	$\Omega_1 = 1.55 \times 10^{-6}$		$\Omega_1 = 7.17 \times 10^{-6}$	$\Omega_1 = 3.42 \times 10^{-6}$
Length in cm (L_l)		$\Omega_2 = 3.03$		$\Omega_2 = 3.346$	$\Omega_2 = 3.134$
Fraction mature (ϕ'_1)	$\varphi'_l = (1 + e^{\Omega_3(L_l - \Omega_4)})^{-1}$		$\Omega_3 = -0.50$ (yr^{-1})		$\Omega_3 = -0.734$ (yr^{-1})
Maturity slope (Ω_3)	$f_a = \sum_{i=1}^{A_l} \emptyset_{al} \varphi'_l w'_l$		$\Omega_4 = 33.5$ (cm)		$\Omega_4 = 33.10$ (cm)
Length at 50% maturity (Ω_4)			$Amat = 7$ (yr)		$Amat = 5$ (yr)
Age at 50% maturity					
Fecundity at age (f_a)					
Fishery selectivity[a]	Double normal	$\beta_1 = 43, \beta_2 = 3, \beta_3 = 5.05, \beta_4 = 6, \beta_5 = -12, \beta_6 = 70$		$\beta_1 = 43, \beta_2 = 3, \beta_3 = 5.05, \beta_4 = 6, \beta_5 = -12, \beta_6 = 70$	
Survey selectivity[a]		$\beta_1 = 33, \beta_2 = 3, \beta_3 = 5.05, \beta_4 = 6, \beta_5 = -12, \beta_6 = 70$		$\beta_1 = 33, \beta_2 = 3, \beta_3 = 5.05, \beta_4 = 6, \beta_5 = -12, \beta_6 = 70$	
Recruitment variation (σ_R)				$\sigma_R = 0.60$	
Autocorrelation (ρ)				$\rho = 0.717$	
Catchability coefficient (Q)				Q = 1	
Survey standard error (σ_s)				$\sigma_s = 0.25$	

[a] See Methot and Wetzel (2013) for the parameterization of selectivity and growth.

2.2. The operating model

An age-structure population was simulated with stochastic recruitment subject to annual fishery removals where an observed biomass index with observation error was available for select years (See supplementary material for operating model details). The population in the operating model at the start of year 51 (immediately prior to the start of management based on OFLs and ABCs) was defined to be a specified proportion of the virgin biomass, with relative biomass level depending on the life-history and the simulation scenario (see Section 2.5 below).

The population was assessed using each estimation method at the start of year 51 based on an annual survey biomass index for years 31–50. The index length was a strategic selection because it covered a period long enough to offer contrast, if present, in the trajectory of the population. The catch-only estimation methods calculate either a single OFL (DCAC) or an OFL for each projection year (DB-SRA) and were not reapplied in the projection period because they do not use index data and do not update the prior distributions. In contrast, the data-moderate estimation methods (XDB-SRA and XSSS) estimated biomass and OFLs for the subsequent four years based on the available index data. A buffer factor, as applied by the PFMC (approximately 0.83 for category 2 data-moderate stock estimation methods [XDB-SRA and XSSS], and approximately 0.69 for category 3 catch-only stock estimation methods [DCAC and DB-SRA]), was applied to each OFL to calculate the ABCs. The ABCs were removed without error and the population projected forward for four years. An additional four years of index data were then generated and provided to each data-moderate estimation method, which then re-assessed the population and calculated OFLs. This iterative process continued over twenty-five years, at which point the performance of the estimation methods were evaluated using performance measures (see Section 2.6 below). DCAC, DB-SRA, and XDB-SRA were applied separately to each of the 100 simulated operating model populations. XSSS was applied to 60 simulated operating model populations. The number of XSSS simulations conducted was limited because it was more time intensive (due to it being age-structured) compared to the other estimation methods.

2.3. Estimation methods

2.3.1. DCAC

DCAC (MacCall, 2009) allows for the estimation of a likely sustainable yield for data-limited stocks based upon average observed catches, distributions for three biologically-based life-history parameters, and an assumed distribution for relative stock status. DCAC calculates a yield as:

$$\text{Sustainable yield} = \frac{\sum_{t=1}^{N} C_t}{n + (1-\delta_{50})\left(\frac{B_{MSY}}{B_0}\left(\frac{F_{MSY}}{M}\right)M\right)^{-1}} \quad (1)$$

where C_t is the catch during year t, n is the length of the catch history, δ_{50} is the relative biomass in year 50, B_{MSY}/B_0 is the biomass that corresponds to maximum sustainable yield relative to carrying capacity, M is the instantaneous rate of natural mortality, and F_{MSY}/M is the ratio of the fishing mortality rate that corresponds to B_{MSY}/B_0 and M. Eq. (1) has been re-parameterized from the version by MacCall (2009) where the original version applied a parameter Δ (defined as the difference between the relative biomass at a previous point in time and the current status) which is $1-\delta_{50}$, assumed to be the relative biomass from the unfished state. This re-parameterization has been applied to DCAC, DB-SRA, and XDB-SRA to create consistency among all methods and how the PFMC defines their target relative biomass levels.

A Monte Carlo approach was applied to account for the parameter uncertainty. For each life-history type and simulation scenario a total of 10,000 random draws were conducted based on distributions for each parameter, which generated a distribution of sustainable yield values.

DCAC was applied at the start of year 51 to calculate a yield based on specified distributions for each life-history type and scenario (Tables 2 and 3). The median value of the resulting distribution was then applied as the OFL. The OFL was adjusted to an ABC value by applying the buffer (0.69 * OFL) that was then removed from the stock for each of the 25 future years. DCAC was not reapplied during the projection period because this method defines a one-time only calculation of yield, and is not an updateable calculation.

2.3.2. DB-SRA

DB-SRA is based on stock reduction analysis (SRA) (Kimura et al., 1984; Walters et al., 2006). Dick and MacCall (2011) adapted the concepts with the addition of a relative biomass based parameter. Similar to DCAC, DB-SRA uses Monte Carlo draws from four parameter distributions (for M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) to create probability distributions for current relative biomass and OFLs. DB-SRA is based on the following delay-difference production model that includes a time lag for recruitment and mortality:

$$B_t = B_{t-1} + P_t(B_{t-amat}) + (1-e^{-M})(B_{t-amat} - B_{t-1}) - C_{t-1} \quad (2)$$

where B_t is the biomass at the start of year t, M is the instantaneous rate of natural mortality, and $P(B_{t\text{-}amat})$ is the latent annual production based on a function of biomass in year t-$amat$, where $amat$ is the age at maturity. The latent annual production is determined by a hybrid between the Pella–Tomlinson–Fletcher model and the Schaefer surplus production model (see Dick and MacCall, 2011 for additional detail).

Unfished biomass, B_0, is calculated separately for each parameter draw from the distributions by solving the equation $B_{t=50}/B_0 = \delta_{50}$ for B_0. The estimated OFL value for each year t is calculated as:

$$OFL_t = (1 - e^{-(M+F_{MSY})})\left(\frac{F_{MSY}}{M+F_{MSY}}\right)B_t \quad (3)$$

This estimation method produces estimates of annual biomass and the corresponding OFLs for the entire projection period (years 51–75). The vector of OFLs was reduced by the PFMC category 3 data-limited buffer factor, 0.69, to produce ABCs that were annually removed from the simulated population.

2.3.3. XDB-SRA

XDB-SRA (PFMC, 2014b) builds upon the basic structure of DB-SRA, using the same four parameters and underlying population dynamics model. However, unlike DB-SRA, it updates the prior distributions for the parameters by fitting the model to a biomass index using adaptive importance sampling (see Section 2.4 below). The application of adaptive importance sampling was based on 1000 initial population trajectories created from draws for each of the four prior distributions and 1000 draws at each step (except the last draw which was based on 2000 draws). The parameter related to the relative biomass (δ_{50}) was always assumed to pertain to year 50 for consistency with how DCAC and DB-SRA were applied, even though XDB-SRA was applied in future years, unlike the catch-only estimation methods. The OFLs for the next four years were set equal to the median of the distribution derived from the estimated population size from each trajectory.

The OFLs for XDB-SRA were determined using the PFMC harvest control rule policy. The PFMC has defined a life-history specific harvest control rule which applies a linear reduction in harvest when a stock is below a pre-specified target relative biomass level. Harvest was reduced when the flatfish stock was below 0.25 of virgin biomass (B_0) or the rockfish stock was below $0.40B_0$, with no fishing when the relative biomass level was below the life-history specific threshold value (flatfish: $0.05B_0$, rockfish: $0.10B_0$). The appropriate life-history harvest control rule for West Coast groundfish was applied based on the posterior median relative biomass value. The OFLs, as determined by the harvest control rule, were reduced based upon the buffer value (0.83) to determine the ABCs for category 2 data-moderate assessments. The ABCs were removed without error from the simulated population.

Table 2
Scenarios (assumed prior distributions) for each assessment method for the flatfish life-history. The mean values of the prior distributions by scenario are given along with the standard deviation, and assumed distribution with the misspecified parameters in italics by scenario. The scenarios are denoted by abbreviation to indicate the prior assumption (T = true values, P = the productivity parameter [F_{MSY}/M or h], and δ_{50} = relative stock status in year 50) and subscript that refers to whether the stock was at or below the management target biomass level in year 50. The true values are given in bold below the corresponding scenario.

Scenarios	Parameters														
	All models						XDB-SRA DB-SRA DCAC						XSSS		
	M	*(sd)*	*Dist.*	δ_{50}	*(sd)*	*Dist.*	F_{MSY}/M	*(sd)*	*Dist.*	B_{MSY}/B_0	*(sd)*	*Dist.*	*h*	*(sd)*	*Dist.*
Scenario T_{at}	0.20	0.25	Lognormal	0.25	0.20	Beta	1.35	0.30	Lognormal	0.26	0.10	Beta	0.88	0.10	Beta
True values	**0.20**			**0.25**			**1.35**			**0.26**			**0.88**		
	M	*(sd)*	*Dist.*	δ_{50}	*(sd)*	*Dist.*	F_{MSY}/M	*(sd)*	*Dist.*	B_{MSY}/B_0	*(sd)*	*Dist.*	*h*	*(sd)*	*Dist.*
Scenario T_{below}	0.20	0.25	Lognormal	0.10	0.20	Beta	1.35	0.30	Lognormal	0.26	0.10	Beta	0.88	0.10	Beta
Scenario D_{below}	0.20	0.25	Lognormal	*0.25*	0.20	Beta	1.35	0.30	Lognormal	0.26	0.10	Beta	0.88	0.10	Beta
Scenario DP_{below}	0.20	0.25	Lognormal	*0.25*	0.20	Beta	*0.80*	0.30	Lognormal	0.26	0.10	Beta	*0.75*	0.10	Beta
True values	**0.20**			**0.10**			**1.35**			**0.26**			**0.88**		

Table 3
Scenarios (assumed prior distributions) for each assessment method for the rockfish life-history. The mean values of the prior distributions by scenario are given along with the standard deviation, and assumed distribution with the misspecified parameters in italics by scenario. The scenarios are denoted by abbreviation to indicate the prior assumption (T = true values, P = the productivity parameter [F_{MSY}/M or h], and δ_{50} = relative stock status in year 50) and subscript that refers to whether the stock was at or below the management target biomass level in year 50. The true values are given in bold below the corresponding scenarios.

Scenarios	Parameters														
	All models						XDB-SRA DB-SRA DCAC						XSSS		
	M	*(sd)*	*Dist.*	δ_{50}	*(sd)*	*Dist.*	F_{MSY}/M	*(sd)*	*Dist.*	B_{MSY}/B_0	*(sd)*	*Dist.*	*h*	*(sd)*	*Dist.*
Scenario T_{at}	0.05	0.25	Lognormal	0.40	0.20	Beta	0.89	0.30	Lognormal	0.45	0.10	Beta	0.50	0.10	Beta
True values	**0.05**			**0.40**			**0.89**			**0.45**			**0.50**		
	M	*(sd)*	*Dist.*	δ_{50}	*(sd)*	*Dist.*	F_{MSY}/M	*(sd)*	*Dist.*	B_{MSY}/B_0	*(sd)*	*Dist.*	*h*	*(sd)*	*Dist.*
Scenario T_{below}	0.05	0.25	Lognormal	0.20	0.20	Beta	0.89	0.30	Lognormal	0.45	0.10	Beta	0.50	0.10	Beta
Scenario D_{below}	0.05	0.25	Lognormal	*0.40*	0.20	Beta	0.89	0.30	Lognormal	0.45	0.10	Beta	0.50	0.10	Beta
Scenario DP_{below}	0.05	0.25	Lognormal	*0.40*	0.20	Beta	*0.60*	0.30	Lognormal	0.45	0.10	Beta	*0.40*	0.10	Beta
True values	**0.05**			**0.20**			**0.89**			**0.45**			**0.50**		

2.3.4. XSSS

XSSS simplifies Stock Synthesis for application in data-moderate situations (PFMC, 2014b). The population model underlying XSSS is sex- and age-structured, with a Beverton–Holt stock recruitment relationship, length-based maturity, with fishery and survey selectivity assumed equal to maturity. The Bayesian analyses were based on equivalent biological parameters to those used in XDB-SRA: M, the instantaneous rate of natural mortality, h (also known as steepness), and the relative biomass in year 50 (δ_{50}). This parameterization has been assumed when applying XSSS for the West Coast data-moderate assessments (Cope et al., 2015) and in simpler extensions (Cope, 2013). The application of adaptive importance sampling for XSSS was based on 1200 initial population trajectories created from draws for each of the four prior distributions and 700 draws at each step (except the last draw which was 1200). The fewer draws for XSSS than XDB-SRA reduced model estimation time while preserving convergence in the posterior distributions. The value for *log*(*R0*) was calculated so that the generated and model-predicted relative biomass levels for year 50 were the same for each parameter draw. The estimated OFLs for the next four years were calculated from the medians of the posterior distributions for the biomass trajectory as for XDB-SRA. The harvest control rule and the calculation of ABCs for category 2 data-moderate stocks were applied the same as for XDB-SRA.

2.4. Adaptive importance sampling

Adaptive importance sampling was applied to both XDB-SRA and XSSS to update prior parameter distributions based upon the fit to the index data. Adaptive importance sampling grew out of the foundation of sampling importance resampling (SIR) (Ruben, 1987, 1988). SIR samples parameter vectors from a prior distribution taken from a sampling envelope and has been applied in fishery stock assessment for parameter estimation (e.g., McAllister et al., 1994). However, the adaptive importance sampling approach updates the sampling envelope based upon iterative SIR draws and can be beneficial when the best sampling envelope is unknown or not well understood a priori due to correlation among parameters (Givens and Raftery, 1996; Kinas, 1996). See supplementary material for technical details.

2.5. Scenarios

Multiple scenarios were created to explore the performance of each estimation method and the sensitivity to parameter misspecification (Tables 2 and 3):

- Scenario T_{at}: All prior distributions were centered about the true values (T) and the simulated stock was at the target relative biomass (at) as defined by the PFMC based on the life-history (flatfish 25%, rockfish 40% of virgin biomass) in year 50.
- Scenario T_{below}: All prior distributions were centered about the true values (T) and the true stock was below the target relative biomass (below) as defined by the PFMC based on the life-history (flatfish 10%, rockfish 20% of virgin biomass) in year 50.
- Scenario D_{below}: The prior distribution for the relative population size in year 50 (δ_{50}) was centered about an optimistic value relative to the true value (D: relative biomass misspecified) and all other prior distributions were centered about the true values. The true stock was below the target relative biomass level (below) defined by the PFMC based on the life-history (flatfish 10%, rockfish 20% of virgin biomass) in year 50.

- Scenario DP_{below}: The prior distributions for the productivity parameter (F_{MSY}/M or h) and the relative population size in year 50 (δ_{50}) were each centered about optimistic values relative to the true values (DP: relative biomass and productivity misspecified). The true stock was below the target relative biomass level (below) defined by the PFMC based on the life-history (flatfish 10%, rockfish 20% of virgin biomass) in year 50.

A challenge for any estimation method is determining which parameters need to be fixed or estimated, due to the lack of information in the data or correlations among the parameters. XDB-SRA and XSSS allow for the estimation of key biological parameters through the application of adaptive importance sampling, but it is uncertain what information the data (a biomass index) may contain to inform each estimation method to correctly estimate these parameters. Specific scenarios explore the performance of the estimation methods when the distribution for the relative stock status in the target year was misspecified and when this parameter was misspecified along with the productivity parameter. The final scenario, that evaluates the misspecification of the relative stock status and productivity parameter, was designed to investigate the impact of multiple misspecifications on estimation performance and whether precaution in setting the productivity parameter would offset the bias in the specification of relative biomass level. Similar model runs could have been conducted to explore the misspecification of the priors for M and B_{MSY}/B_0 (DCAC, DB-SRA, XDB-SRA), but they were beyond the scope of the current analysis.

2.6. Performance measures

The mean of the estimated parameter distribution for each of the model's input parameters was evaluated (DCAC, DB-SRA, and XDB-SRA: M, F_{MSY}/M, B_{MSY}/B_0 and δ_{50}; XSSS: M, h, and δ_{50}). The choice to summarize by the mean rather than the median for the parameter inputs was based upon the prior parameter distributions, which were specified by a mean value and were either assumed to be lognormal or beta distributed. If the data provided no information and the prior was not updated, resulting in a posterior equal to the prior, summarizing the distributions by the median would result in estimates that would differ from the input distribution values that generated the prior. The model-derived estimates (spawning biomass, current relative biomass, ABCs) were summarized by the medians of their distributions. The relative error of the estimates was calculated as:

$$RE = \frac{E - T}{T} \tag{4}$$

where E is the estimated value and T is the corresponding true value from the operating model.

The performance of each estimation method was evaluated using the following criteria:

1. Relative biomass at the end of the projection period relative to management targets.
2. The probability of overfishing (ABC > true OFL).
3. The probability of being overfished (flatfish: < $0.125B_0$, rockfish: < $0.25B_0$) at the end of the projection period.
4. The percent of the estimated catch realized over the projection period relative to the operating model OFL by each estimation. A perfect performance for the catch-only estimation methods would be 69% (ABC = 69% OFL) and 83% (ABC = 83% OFL) for the data-moderate estimation methods.
5. The average annual variability of ABC catches (abbreviation AAV) over the projection period defined as:

$$AAV = 100 \frac{\sum_{t=1}^{N} |C_t - C_{t+1}|}{\sum_{t=1}^{N} C_t} \tag{5}$$

where C_t is the catch during year t.

To evaluate estimation performance, simulations were conducted where the populations were managed with perfect knowledge and independent of the estimation methods. The "perfect knowledge" results provide information on the best possible performance given the current management guidelines. The true OFL and the resulting ABC, based on the operating model population status were determined and removed without error.

3. Results

3.1. Overview

All estimation methods resulted in overfishing when the δ_{50} prior was misspecified for both life-history types. This was most notable for DB-SRA. The catch-only methods (DCAC and DB-SRA) generally avoided overfishing when the stocks were correctly judged to be at the relative biomass target at the time of the first assessment in year 51, estimating OFLs that resulted in ABC values at or below the true OFLs. The data-moderate estimation methods (XDB-SRA and XSSS) resulted in very different results when the stock was at the target biomass level with XDB-SRA estimating OFLs and resulting ABCs that were well above the true OFLs. All methods, except DCAC, resulted in overfishing when the stock was below the target biomass level at the time of the first assessment when the δ_{50} prior was misspecified at an overly optimistic value. However, within 10–15 years, the index data used by XDB-SRA and XSSS were sufficient to result in updated posterior distributions for the δ_{50} prior that estimated OFLs that produced ABC values which allowed the stocks to rebuild.

3.2. Failed simulations

A small number of simulations 'failed' for XDB-SRA and XSSS because a sample was drawn which resulted in a singular covariance matrix. The failed simulations only occurred for the flatfish life-history. This behavior was commonly observed later in the projection period and was attributed to the parameters entering a narrowly defined parameter space which led to very few unique parameter combinations being supported by the data, resulting in repeated draws by adaptive importance sampling. The percent of simulations on which the results are based for XDB-SRA by scenario were; T_{at}: 99%, T_{below}: 90%, D_{below}: 94%, and DP_{below}: 100% and for XSSS by scenario were; T_{at}: 95%, T_{below}: 85%, D_{below}: 72%, and DP_{below}: 85%.

3.3. Flatfish life-history

3.3.1. Scenario T_{at}

The status of the simulated stocks at the end of the projection period were variable within and among estimation methods for the T_{at} scenario, where all priors were assumed correctly and the simulated relative biomass was at the management target at the start of the projection period (Table 2). A high percentage of the stocks where XDB-SRA was applied for estimation were below the relative biomass target value ($0.25B_0$) in year 75, with a large subset of those below the overfished level ($0.125B_0$) (Table 4). In contrast, at the end of the projection period, DCAC, DB-SRA, and XSSS resulted in either none or a low percentage of simulated stocks below the target level (Table 4). The median operating model relative biomasses at the end of the projection period for these three methods were above the target levels (Fig. 2a). The OFLs and the resulting ABCs set

Table 4
The percent of simulated flatfish stocks that are below the relative biomass target (25%), overfished (<12.5%), or extinct at the end of the projection period (year 75).

	Below target %				Overfished %				Extinct %			
	Estimation method				Estimation method				Estimation method			
	DCAC	DB-SRA	XDB-SRA	XSSS	DCAC	DB-SRA	XDB-SRA	XSSS	DCAC	DB-SRA	XDB-SRA	XSSS
Scenario T_{at}	0	17	81	21	0	9	59	4	0	4	15	0
Scenario T_{below}	41	60	51	10	37	55	30	2	31	48	14	0
Scenario D_{below}	44	88	55	49	40	86	32	33	36	83	14	14
Scenario DP_{below}	29	57	32	49	23	51	14	18	19	46	4	2

by DCAC, DB-SRA, and XSSS were well below the true OFLs, while XDB-SRA substantially overestimated the true OFLs resulting in a probability of overfishing that was > 0.50 for each year the stock was assessed (Fig. 3a, Table 4). The median AAV for XDB-SRA was highest in the first five years of the projection period (Fig. 4a). However, the average catch increased over the projection period (Fig. 4a), resulting in the observed median population decline (Fig. 2b).

XDB-SRA and XSSS resulted in dissimilar median posterior estimates for several of the parameters. The medians of the posterior means for the δ_{50} parameter for XDB-SRA were consistently greater

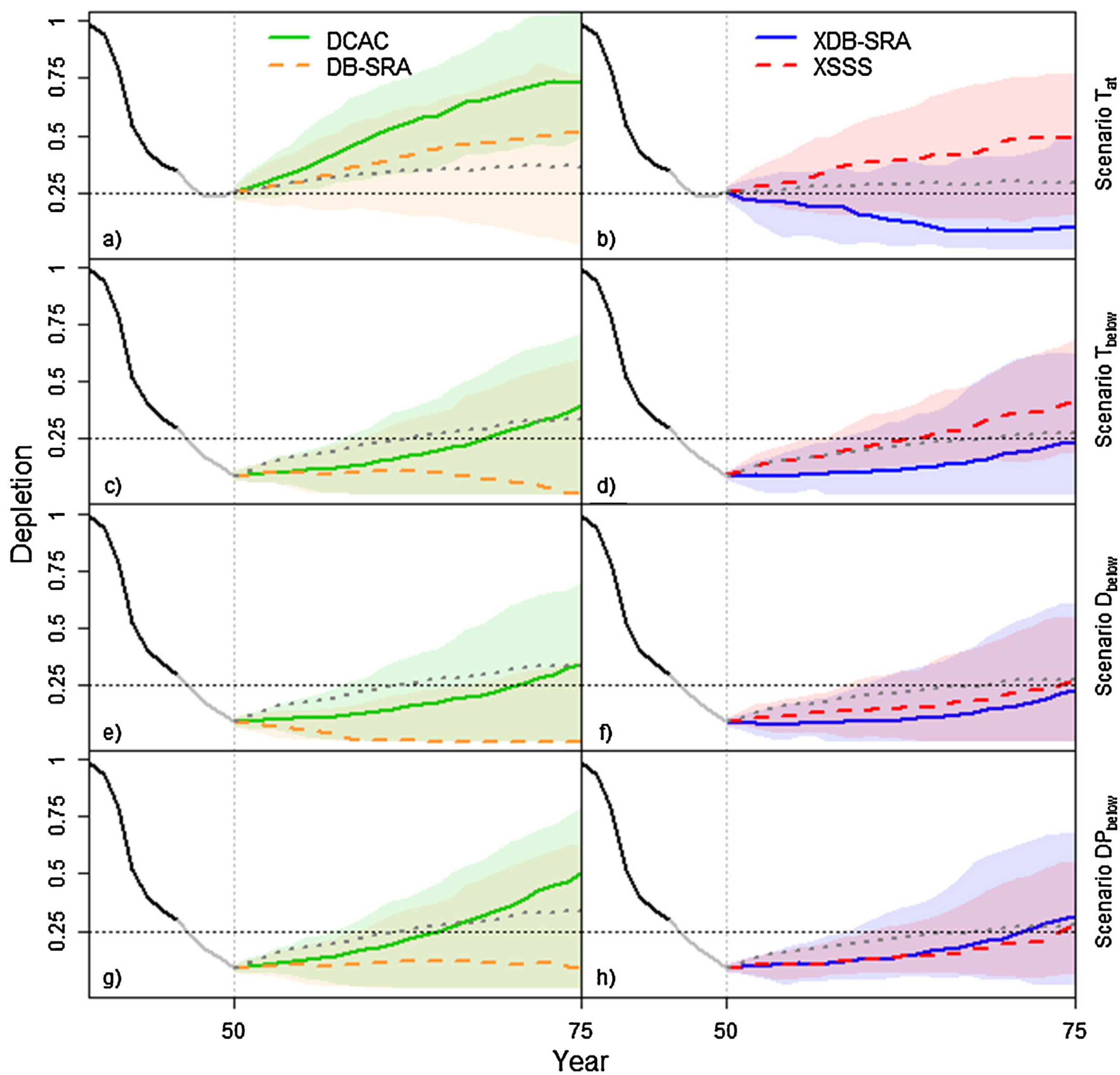

Fig. 2. Time-trajectories of relative biomass for the flatfish population with 90% simulation intervals when the OFLs and ABCs are provided by: DCAC (green line and interval) and DB-SRA (orange dashed line and interval), shown in the left panels, and XDB-SRA (blue line and interval) and XSSS (red dashed line and interval), shown in the right panels for each of the four scenarios. The median relative biomass over the simulations if the stock was managed with perfect information from the operating model with the OFL adjusted by the appropriate buffer being removed without error is shown in each panel (dotted gray line). The years for which a biomass index was available for the first assessment in year 50 is shown by the light gray line of the time-trajectories of the simulated stocks prior to start of the projection period. The vertical dotted line indicates the start of the projection period and the horizontal dotted line indicates the target value for flatfish stocks set by the PFMC. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

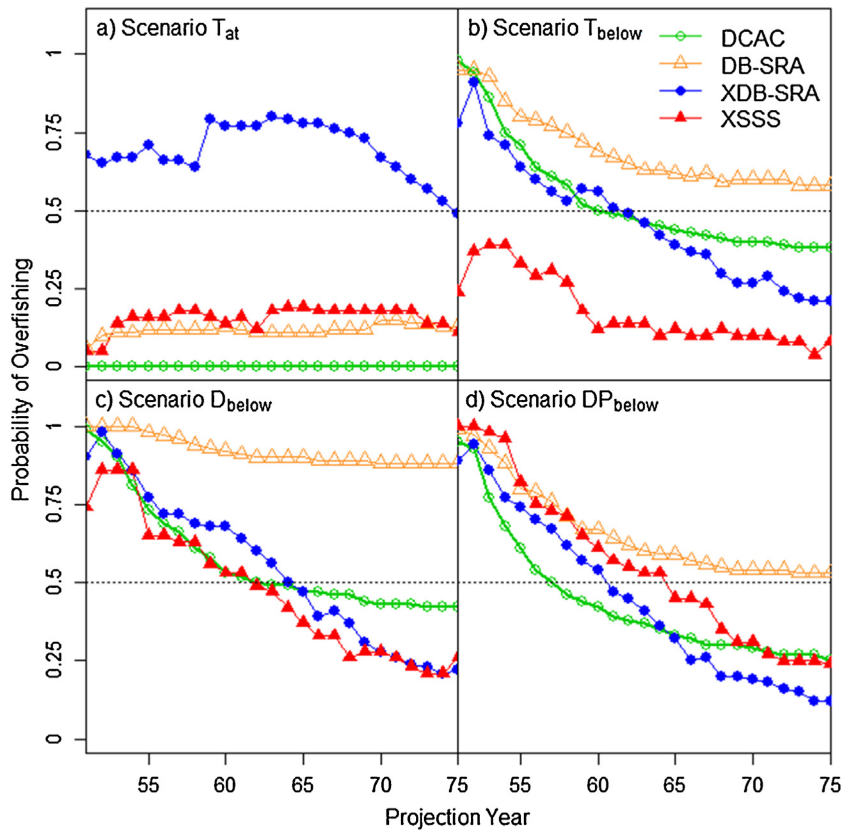

Fig. 3. The probability of overfishing (ABC > true OFL) for the flatfish life-history during the projection period for each assessment method (DCAC, DB-SRA, XDB-SRA, and XSSS) and scenario.

than the true value throughout the projection period (Fig. 5c), although this tendency declined over time. There was little contrast in the biomass index for the flatfish life-history for years 31–50, thus the index data were insufficient to inform stock status, causing a bias for the relative biomass at the end of year 50. The medians of the posterior means for *M* and F_{MSY}/M from XDB-SRA were unbiased for the first two assessments, but for later assessments the among-simulation variance increased and most of the posterior means were biased low (Fig. 5a and b). Overall, the median of the posterior medians for the estimated assessment year relative biomass for XDB-SRA were greater than the true values and the among-simulation variance increased over the projection period despite the increased amount of data (Fig. 5e). The posterior distribution means for *M* and *h* from XSSS were generally centered about the true values for all years (Fig. 6a and b). The posterior distributions for relative stock status at the end of year 50 from XSSS emphasized values larger than the true value at the start of the projection period but the posterior medians were less than the true value by the time of the last assessment in year 74 (Fig. 6c). The distribution of posterior medians for the assessment year estimated relative biomass was unbiased for the first two assessments, but distributions became biased low relative to true stock status over the projection period (Fig. 6d).

3.3.2. *Scenario* T_{below}

The median relative biomass was below the target level at the end of year 50, and all prior distributions were unbiased for the T_{below} scenario (Table 2). This scenario resulted in a high percentage of stocks that were still below the management target level at the end of the projection period, with a portion of them below the overfished threshold or even extinct depending upon the estimation method (Table 4). The median time-trajectory of relative biomass varied among estimation methods for this scenario, with the estimated OFLs and the resulting ABCs set using DCAC, XDB-SRA, and XSSS leading to increases in relative biomass, while the median relative biomass of the populations where DB-SRA was applied declined toward zero by year 75 (Fig. 2c and d). DCAC and XDB-SRA set ABCs that were greater than the true OFLs during the early part of the projection period, resulting in a high probability of overfishing (approximately 1 and 0.75) (Fig. 3b). However, the probability of overfishing for DCAC declined to < 0.50 by approximately year 60 (Fig. 3b). DB-SRA estimated OFLs that resulted in ABCs that were greater than two times the true value OFLs resulting in a declining population (Fig. 2c, Table 5), and in the highest probability of overfishing among the estimation methods (Fig. 3b). XDB-SRA and XSSS resulted in similar median AAVs, but XSSS had less among-simulation variation in both average catch and AAV compared to XDB-SRA (Fig. 4b).

The medians of the posterior means were more variable and further from the true values for XDB-SRA compared to XSSS. The medians of the posterior means for *M*, F_{MSY}/M, and B_{MSY}/B_0 from XDB-SRA were less than the true values and error generally increased over time (Fig. 5f, g, and i). The first assessment resulted in the largest among-simulation variation and medians of the posterior means that were well below the true value for the δ_{50} parameter from XDB-SRA (Fig. 5h). The posterior distributions for estimated assessment year stock status from XDB-SRA were greater than the true value at the start of the projection period, but approached the true value with each subsequent assessment (Fig. 5j). The XSSS posterior distributions for *M* were less than the true value for all assessments (Fig. 6e). The posterior distributions for *h* had most of their mass less than the true value for the first assessment with increasing inter-simulation variation in the posterior medians over time (Fig. 6f). The medians of

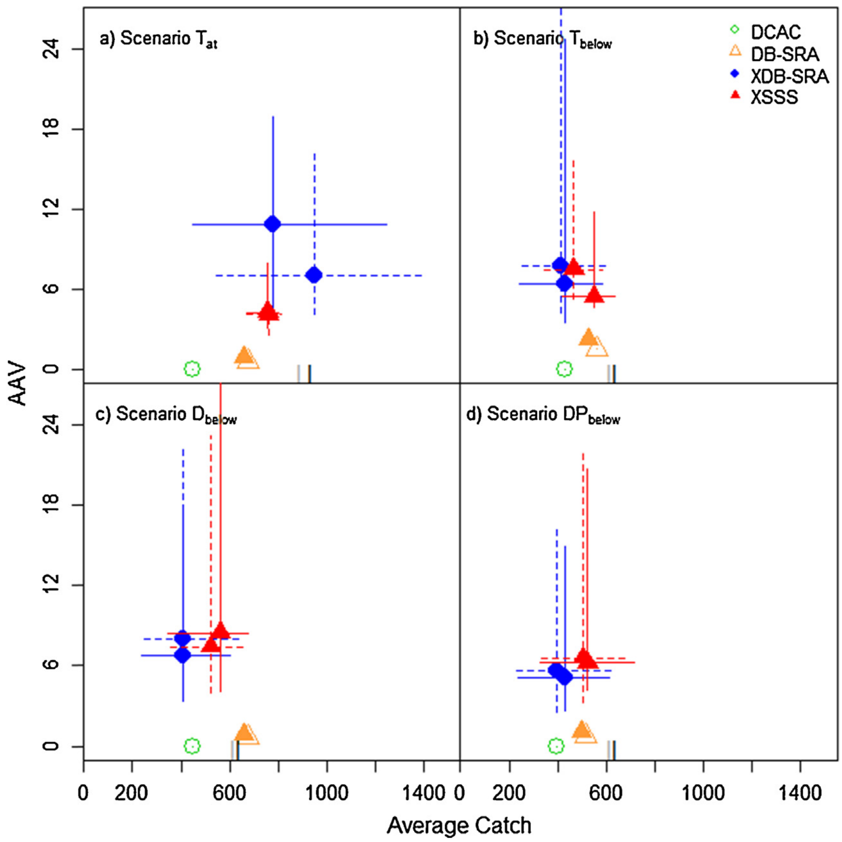

Fig. 4. The average annual variation in catch vs the average catch (with 90% simulation intervals) after five years (solid line), and 25 years (dotted line) by assessment method; DCAC, DB-SRA (5 yrs: open diamond, 25 yrs: filled diamond), XDB-SRA, and XSSS, and each scenario. There is no inter-simulation variation in the AAV and average catch for DCAC and DB-SRA since each method estimates a time-series of harvest levels independent of data. The average catch over 25 years when the true maximum catch from the operating model was removed without error from the stock is shown on the *x*-axis for each buffer level (data-moderate category two [XDB-SRA and XSSS]: black and data-limited category three [DCAC and DB-SRA]: gray).

Table 5
The simulated flatfish stock median percent and 95% simulation interval of catch realized by each method relative to the operating model estimated OFL over the projection period. A perfect performance for the catch-only estimation methods would be 69% (ABC = 69% OFL) and 83% (ABC = 83% OFL) for the data-moderate estimation methods.

	DCAC		DB-SRA		XDB-SRA		XSSS	
	Median	95% SI	Median	95% SI	Median	95% SI	Median	95% SI
Scenario T_{at}	23.3%	(17.7–45.3%)	46.6%	(36.0–61.8%)	167.7%	(48.2–329.2%)	52.3%	(28.1–151.8%)
Scenario T_{below}	70.6%	(33.6–186.2%)	133.1%	(75.1–194.0%)	107.1%	(39.4–1347.2%)	60.1%	(38.0–143.1%)
Scenario D_{below}	80.2%	(37.4–194.9%)	391.6%	(248.3–558.4%)	122.0%	(40.3–1024.1%	100.6%	(46.7–752.4%)
Scenario DP_{below}	53.9%	(26.2–167.7%)	115.8%	(58.5–204.3%)	79.2%	(33.1–644.6%)	108.5%	(50.0–353.2%)

posterior distributions for the relative biomass in year 50 varied among assessments (Fig. 6g). The distribution of the posterior medians for the estimated assessment year stock status were roughly centered about the true value until the final two assessments (years 70 and 74) at which point the posterior medians were less than the true value (Fig. 6h).

3.3.3. *Scenario D_{below}*

The population was below the relative biomass target, and the prior for this parameter was misspecified at an overly optimistic value for scenario D_{below} (Table 2). The results for this scenario were qualitatively similar to scenario T_{below}, with the following noteworthy exceptions. Specifically, DB-SRA resulted in a high percentage of extinct stocks by the end of the projection period due to a high rate of overfishing (Fig. 3c, Tables 4 and 5) while XSSS led to overfishing with > 0.5 probability until year 65 (Fig. 3c) resulting in almost half of stocks being below the target level at the end of the projection period (Table 4). The posterior means from XSSS are notably more variable among simulations for scenario D_{below} than for scenario T_{below} (Fig. 6i–l vs e–h), and in contrast to scenario T_{below} the posterior means for the relative biomass in year 50 were estimated greater than the true values for each assessment year (Fig. 6k vs g). The variability observed in the distributions for XSSS is due in part to the smaller simulation size for this scenario (43) relative to the other scenarios (See Section 3.2 for details).

3.3.4. *Scenario DP_{below}*

The prior distributions for the productivity parameter and for the relative stock status in year 50 are centered on incorrect values, and the stock was below the relative biomass target for scenario DP_{below} (Table 2). The results for this scenario in terms of relative stock status and probability of overfishing for DCAC, DB-SRA, and XDB-SRA were generally similar those for scenario T_{below} (Figs. 2c, d and 3b). In contrast, XSSS performed poorer compared to scenario T_{below} (Fig. 2h vs d) with estimated ABCs exceeding the true OFLs over the projection period (Table 5) resulting in 49% of the simulated populations below the relative biomass target at the end of the projection period (Table 4).

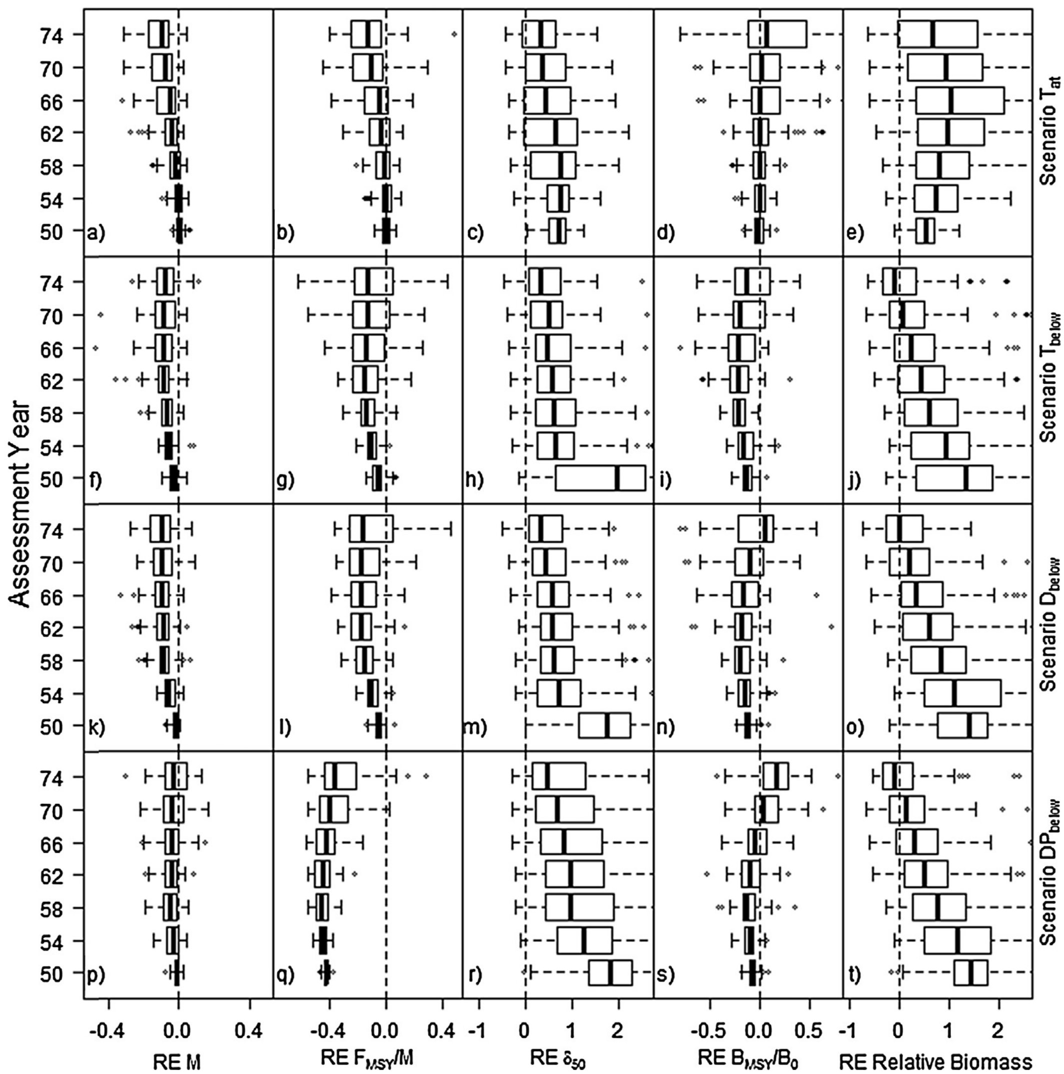

Fig. 5. The distribution of the relative error between the true values and the means of the posterior distributions for the four leading parameters (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) and relative error between the median of the posterior distribution for the assessment year stock status for XDB-SRA for the flatfish life history. The thick black line indicates the median of the posterior distribution means (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

XDB-SRA and XSSS resulted in varying updating in the misspecified prior distributions. The medians of the posterior means for F_{MSY}/M from XDB-SRA were less than the true value for all years, showing little evidence for updating (Fig. 5q). The median of the posterior means for the δ_{50} parameter, which was also misspecified, was furthest from the true value for the first assessment and improved with each subsequent assessment (Fig. 5r). Although there was little updating for each of the misspecified priors, the medians of the posterior distributions for estimated assessment year relative stock status were similar to those observed in scenario T_{below} (Fig. 5j vs t), which was achieved through shifting the distributions of the other estimated model parameters. The posterior distribution for h, one of the misspecified parameters, from XSSS showed no evidence of updating toward the correct value (Fig. 6n). The median of the posterior means for the δ_{50} parameter in year 50, the prior which was misspecified, was furthest from the true value in the first assessment, but declined with subsequent assessments toward the true value (Fig. 6o). In contrast to XDB-SRA, the medians of the posterior distributions for the estimated assessment year relative stock status were much more variable among simulations and shifted to values greater than the true value compared to that observed in scenario T_{below} (Fig. 6p vs h).

3.4. Rockfish life-history

3.4.1. Scenario T_{at}

The median final relative biomass was above the target level for DCAC and DB-SRA, when the assumed prior distributions were centered on the true values and the stock was at the target level at the start of the projection period (Scenario T_{at}) (Fig. 7a, Table 3). The median final relative biomass was below the target level when management was based on XDB-SRA and XSSS (Fig. 7b), with a majority of the simulations below the target level at the end of the projection period (Table 6). The probability of overfishing occurring over the projection period was low for all estimation methods except for XDB-SRA, for which this probability increased to more than 0.50 by the end of the projection period (Fig. 8a, Table 7). The median AAV was comparable between XDB-SRA and XSSS, and was consistent over time (Fig. 9a). XDB-SRA had the highest average catch, which resulted in overfishing and a decline in the median population trajectory, while DCAC had the lowest average catch (Fig. 9a).

The medians of the posterior distributions for parameters were similar for both XDB-SRA and XSSS. The medians of the posteriors means for M, F_{MSY}/M, and B_{MSY}/B_0 were generally centered about

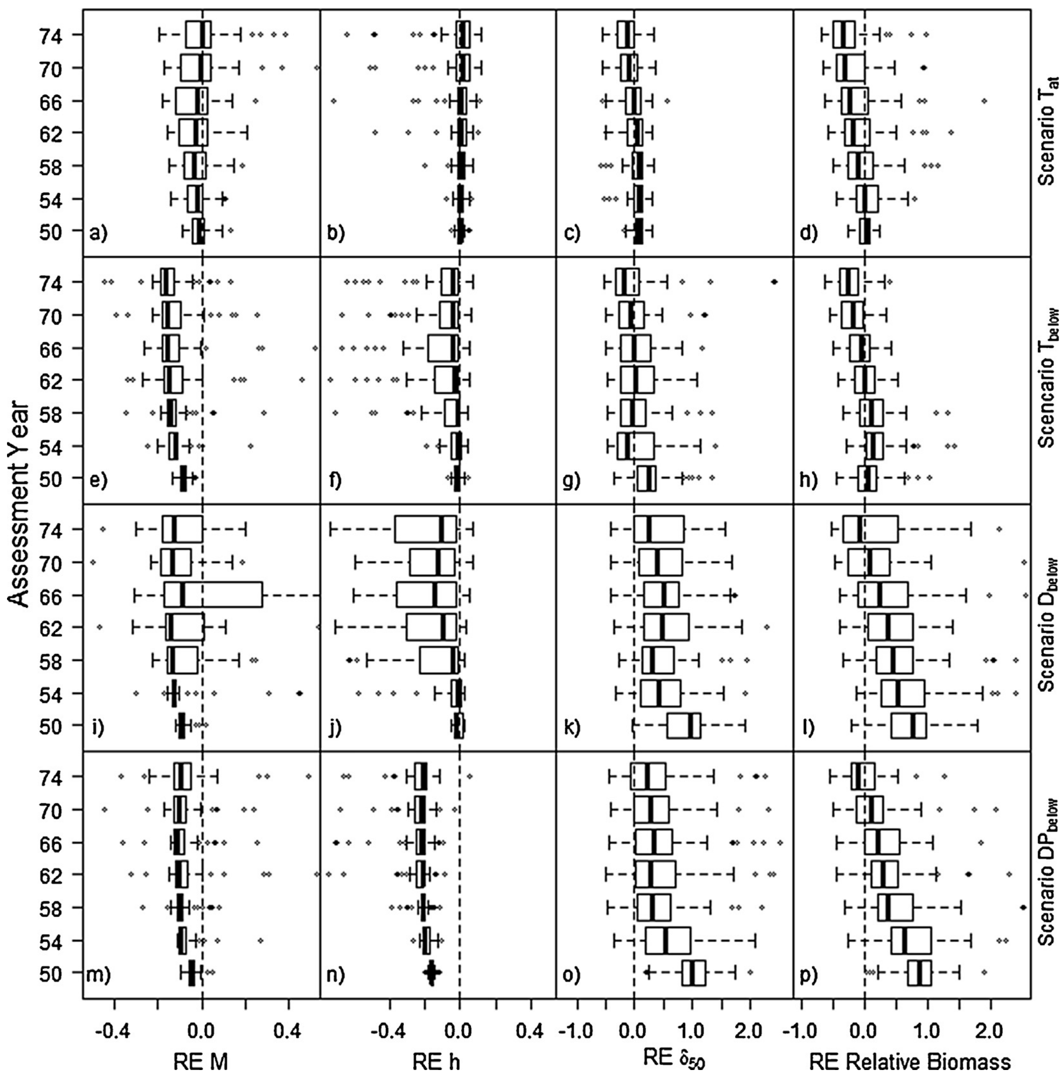

Fig. 6. Relative error between the means of the posterior distributions for the three leading parameters (M, h, δ_{50}) and the relative error between the median of the posterior distributions for assessment year stock status and their true values for XSSS for the flatfish life history, expressed as relative to the true values. The thick black line indicates the median of the posterior distribution means (M, h, δ_{50}) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

Table 6
The percent of simulated rockfish stocks that are below the relative biomass target (40%), overfished (<25%), or extinct at the end of the projection period.

	Below target %				Overfished %				Extinct %			
	Estimation method				Estimation method				Estimation method			
	DCAC	DB-SRA	XDB-SRA	XSSS	DCAC	DB-SRA	XDB-SRA	XSSS	DCAC	DB-SRA	XDB-SRA	XSSS
Scenario T_{at}	0	31	93	78	0	1	11	3	0	0	1	0
Scenario T_{below}	100	100	100	100	56	58	54	28	0	0	1	0
Scenario D_{below}	100	100	100	100	78	100	76	67	0	25	0	0
Scenario DP_{below}	100	100	100	100	50	93	41	83	0	1	0	2

Table 7
The simulated rockfish stock median percent and 95% simulation interval of catch realized by each method relative to the operating model estimated OFL over the projection period. A perfect performance for the catch-only estimation methods would be 69% (ABC = 69% OFL) and 83% (ABC = 83% OFL) for the data-moderate estimation methods.

	DCAC		DB-SRA		XDB-SRA		XSSS	
	Median	95% SI	Median	95% SI	Median	95% SI	Median	95% SI
Scenario T_{at}	30.7%	(26.4–37.8%)	57.5%	(52.1–64.8%)	97.7%	(60.0–165.7%)	80.1%	(52.4–147.1%)
Scenario T_{below}	97.5%	(77.2–129.0%)	99.3%	(86.7–117.0%)	89.9%	(53.9–149.6%)	78.1%	(46.9–153.6%)
Scenario D_{below}	131.6%	(112.2–154.8%)	517.3%	(241.4–2325.3%)	116.6%	(65.7–197.5%)	99.4%	(60.3–208.6%)
Scenario DP_{below}	87.4%	(66.7–120.1%)	201.1%	(174.3–274.5%)	85.7%	(47.3–142.7%)	124.1%	(76.9–251.8%)

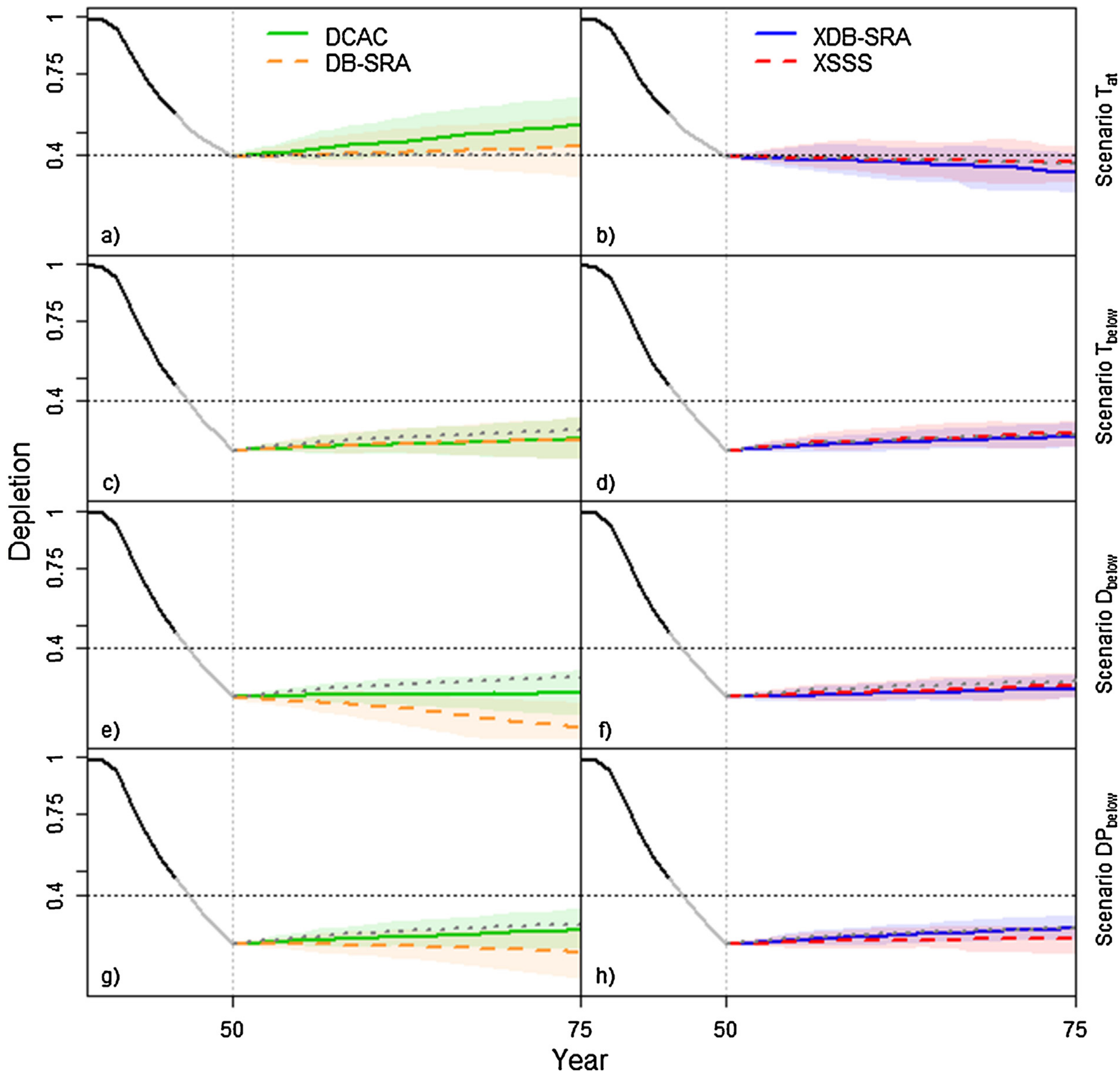

Fig. 7. Time-trajectories of relative biomass for the rockfish population with 90% simulation intervals when the OFLs are provided by: DCAC (green line and interval) and DB-SRA (orange dashed line and interval), shown in the left panels, and XDB-SRA (blue line and interval) and XSSS (red dashed line and interval), shown in the right panels for each of the four scenarios. The median relative biomass over the simulations if the stock was managed with perfect information from the operating model with the OFL adjusted by the appropriate buffer being removed without error is shown in each panel (dotted gray line). The years for which a biomass index was available for the first assessment in year 50 is shown by the light gray line of the time-trajectories of the simulated stocks prior to start of the projection period. The vertical dotted line indicates the start of the projection period and the horizontal dotted line indicates the target value for rockfish stocks set by the PFMC. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the true value (Fig. 10a, b, and d). The among-simulation variance of the parameters increased over time, a counter-intuitive result given that the estimation method has more data as time goes on. This could, however, be attributed to the inconsistencies between the model assumptions and the way the data are generated. The medians of the posterior means for the δ_{50} parameter for XDB-SRA were greater than the true value (Fig. 10c) so the estimation method often assumed a more optimistic relative biomass level for the population when setting OFLs. The variability in the posterior means also increased over time for XSSS, with the posterior means being closest to the true values for the first assessment (Fig. 11a–c). Similar to XDB-SRA, XSSS estimates an overly optimistic relative biomass level in year 50, and the outlook for year 50 relative biomass becomes more optimistic as the projection continues (Fig. 11c). The result of this is an increase in the among-simulation variance for the estimates of estimated assessment year stock status with the median of the posterior medians being greater than the true value (Fig. 11d).

3.4.2. *Scenario* T_{below}

The median final relative biomass was below the target level at the start of the projection period and all prior distributions were centered about the correct values for scenario T_{below} (Table 3). The time-trajectory of median biomass recovered slowly toward the target level for all estimation methods in this scenario (Fig. 7c and d). However, all of the simulated stocks were still below the target level at the end of the projection period due to the slow dynamics of the long-lived rockfish with a low steepness value governing population growth, where many simulations led to populations below the overfished threshold by the end of the projection period

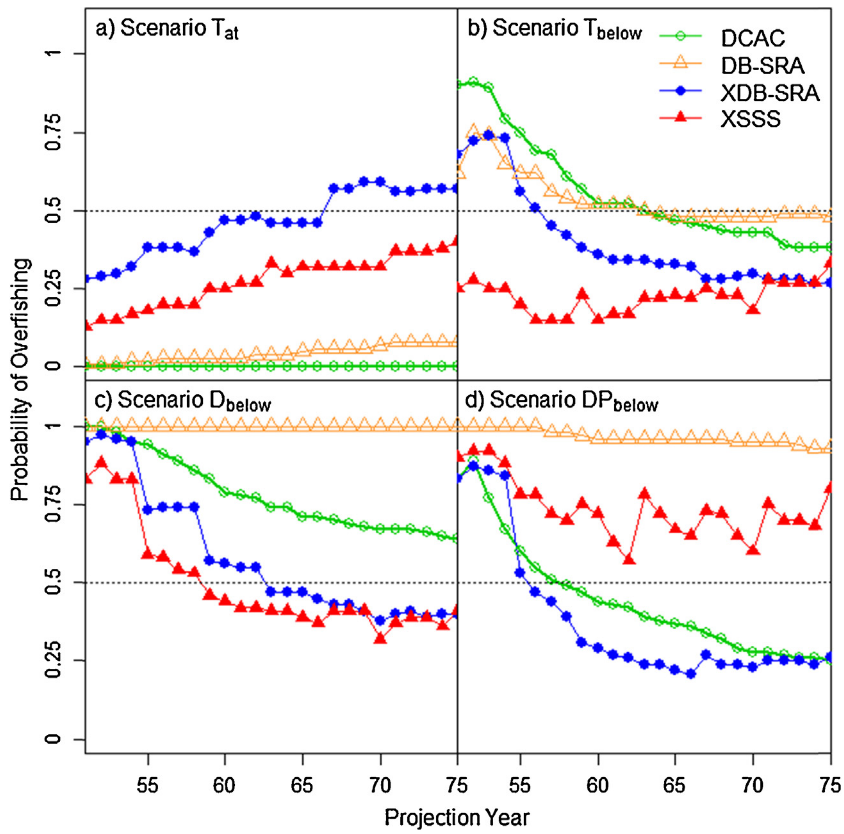

Fig. 8. The probability of overfishing (ABC > true OFL) for the rockfish life-history during the projection period for each assessment method (DCAC, DB-SRA, XDB-SRA, and XSSS) and scenario.

(Table 6). DCAC, DB-SRA, and XDB-SRA resulted in high probabilities of overfishing at the start of the projection period (Fig. 8b). The input assumption for DCAC and DB-SRA was that the stock was at or near a target value and not in need of rebuilding, which was not the case here and consequently resulted in overly optimistic OFLs and hence ABCs (Table 7). DCAC and DB-SRA resulted in slightly higher average catches over the projection period, while XDB-SRA and XSSS had slightly lower but more variable catches (Fig. 9b, Table 7).

The posterior means for the estimated parameters for XDB-SRA were generally less than the true values and the magnitude of the errors increased over time, except for δ_{50} (Fig. 10f–i) The posterior means for δ_{50} were above the true value at the time of the first assessment but decreased toward the true values over the projection period (Fig. 10h). The posterior medians for the estimated assessment year stock status from XDB-SRA were above the true value (Fig. 10j). However, the mass of the distribution shifted to encompass the true value by the third assessment in year 58. The posterior means for the productivity parameter, *h*, from XSSS were consistently less than the true value following the first assessment (Fig. 11f). A similar pattern was observed for the distribution of the posterior means for *M*, which were near the true value in year 50, but shifted to lower values in all subsequent years (Fig. 11e). However, the median of the posterior medians for the estimated assessment year stock status from XSSS were only slightly above the true values, with the bulk of the distribution encompassing the true value for all years (Fig. 11h).

3.4.3. Scenario D$_{below}$

The population was initially below the target level and the mean of the prior distribution for the stock status parameter in year 50 was assumed incorrectly in scenario D_{below} (Table 3). Qualitatively, the results for XDB-SRA and XSSS were similar to those for scenario T_{below}, with median relative biomass trajectories that were slowing increasing (Fig. 7f), although each method estimated OFLs that resulted in ABCs above the true values for the first few assessments, resulting in probabilities of overfishing > 0.50 for the corresponding period (Fig. 8c). In contrast, DCAC and DB-SRA performed poorer than for scenario T_{below}, with a high probability of overfishing (>0.60) (Fig. 8c, Table 7), and a majority of the simulated populations below the overfished threshold by the end of the projection period (Table 6).

XDB-SRA resulted in overfishing and a high proportion of simulation populations below the overfished threshold at the end of the projection period because the posterior means of the δ_{50} parameter were too high during the early part of the projection period, likely due to the misspecification of this parameter and the data not being informative enough for the model to update it to the correct value (Fig. 10m). The posterior means for this parameter did move closer to the true value over time, although the majority of the posterior means were above the true value. Similar to XDB-SRA, the median of the posterior means for δ_{50} from XSSS were above the true value by the greatest amount in the first assessment and updated closer to the true value over time, although most of the posterior means remained above the true values (Fig. 11k).

3.4.4. Scenario DP$_{below}$

The population was below the target level, and the prior distributions for the productivity and the relative biomass priors in year 50 were assumed incorrectly for scenario DP$_{below}$ (Table 3). Both DB-SRA and XSSS resulted in high percentages of the simulated stocks below the overfishing threshold at the end of the projection

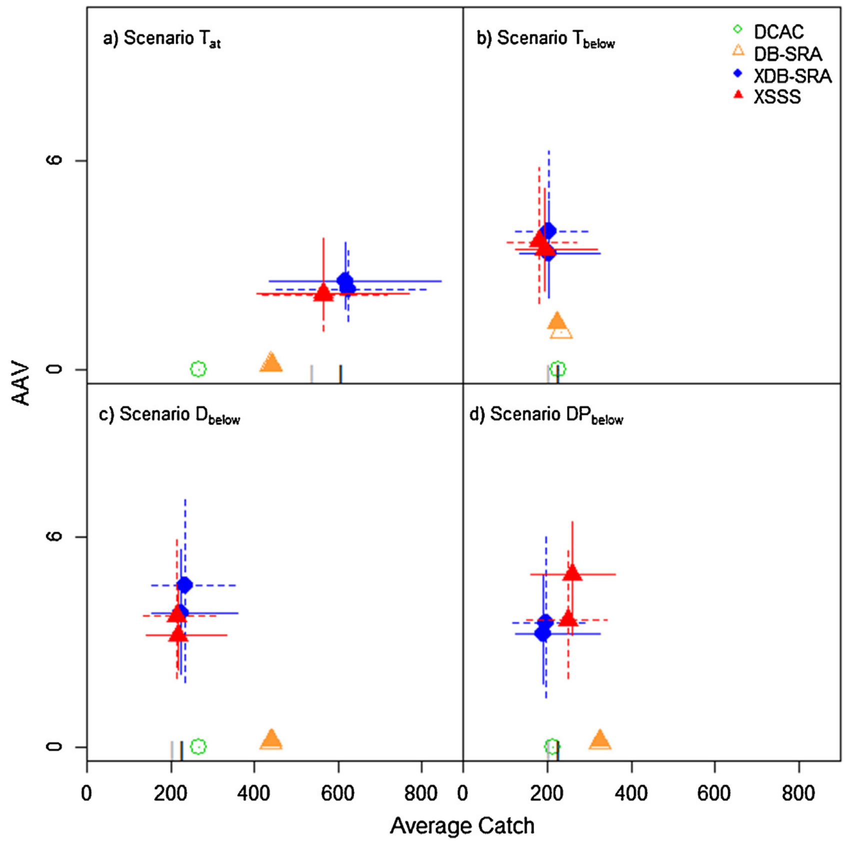

Fig. 9. The average annual variation in catch vs the average catch (with 90% simulation intervals) after five years (solid line), and 25 years (dotted line) by assessment method; DCAC, DB-SRA (5 yrs: open diamond, 25 yrs: filled diamond), XDB-SRA, and XSSS, and each scenario. There is no inter-simulation variation in the AAV and average catch for DCAC and DB-SRA since each method estimates a time-series of harvest levels independent of data. The average catch over 25 years when the true maximum catch from the operating model was removed without error from the stock is shown on the *x*-axis for each buffer level (data-moderate category two [XDB-SRA and XSSS]: black and data-limited category three [DCAC and DB-SRA]: gray).

period (Table 6) due to a high probability of overfishing (Fig. 8d), and large average catches (Fig. 9d, Table 7).

There was again little evidence of strong updating in either of the misspecified parameters, F_{MSY}/M and δ_{50} (Fig. 10q and r) for XDB-SRA. The estimates of *h*, the misspecified parameter, from XSSS updated away from the true value to favor smaller values, similar to scenario T_{below}, although to a greater extent (Fig. 11n vs f). The distribution of the posterior medians for δ_{50} in year 50 was above the true value, and moved closer to the true value in years 54 and 58, but then drifted away in the final assessments (Fig. 11o).

4. Discussion

The performance of XDB-SRA was poor for the flatfish life-history, most markedly when the simulated stock was initially at the relative biomass target and all parameter distributions were specified correctly. The fast flatfish dynamics, along with the historical exploitation pattern, generally resulted in a biomass index with little contrast when the first assessment was conducted. The non-informative index data led to an overestimation of the spawning biomass, which resulted in both an overly optimistic estimate of relative biomass and OFLs. The reduction between the OFLs and the ABCs by the buffer was not large enough to prevent overfishing. This behavior continued for the subsequent assessments, with estimation performance often not improving until the true stock was sufficiently depleted to offer informative index data. The pattern of overly optimistic estimates of relative biomass for the first assessment was also evident in the scenarios where the stock was initially below the target level regardless if the prior for relative biomass was specified correctly or not, although the subsequent estimates improved fairly rapidly as the index data became more informative.

In contrast, overall XSSS resulted in median population trajectories that either maintained above or rebuilt the stock to the target level over the projection period for all scenarios for the flatfish life-history. XDB-SRA and XSSS apply varying assumptions regarding the productivity of the stock and age-structure. XDB-SRA is a delay-difference model while XSSS has the advantage of full age-structured dynamics similar to the operating model. Also, while the operating model stock-recruitment relationship did not match the form assumed by either of the estimation methods, the rigid form of Shepherd stock-recruitment curve applied in the operating model is more akin to the Beverton–Holt form assumed by XSSS compared to the hybrid Pella–Tomlinson–Fletcher and Schaefer surplus production model assumed by (XDB-SRA). The combined impact of an age-structured population and the stock-recruitment form for a highly productive stock (flatfish high steepness) is the likely cause for the varying performance between these estimation methods. However, XSSS also resulted in poor estimates of assessment year stock status for the first assessment in the scenarios where the relative biomass prior was misspecified (D_{below} and DP_{below}), although the posterior means were closer to the true value compared to XDB-SRA. Each of these methods failed to estimate posterior distributions that matched the true value about relative biomass when misspecified. The largest overestimation occurred at the time of the first assessment, resulting in overfishing despite the reductions applied to the estimated OFLs by the category 2 data-moderate buffer.

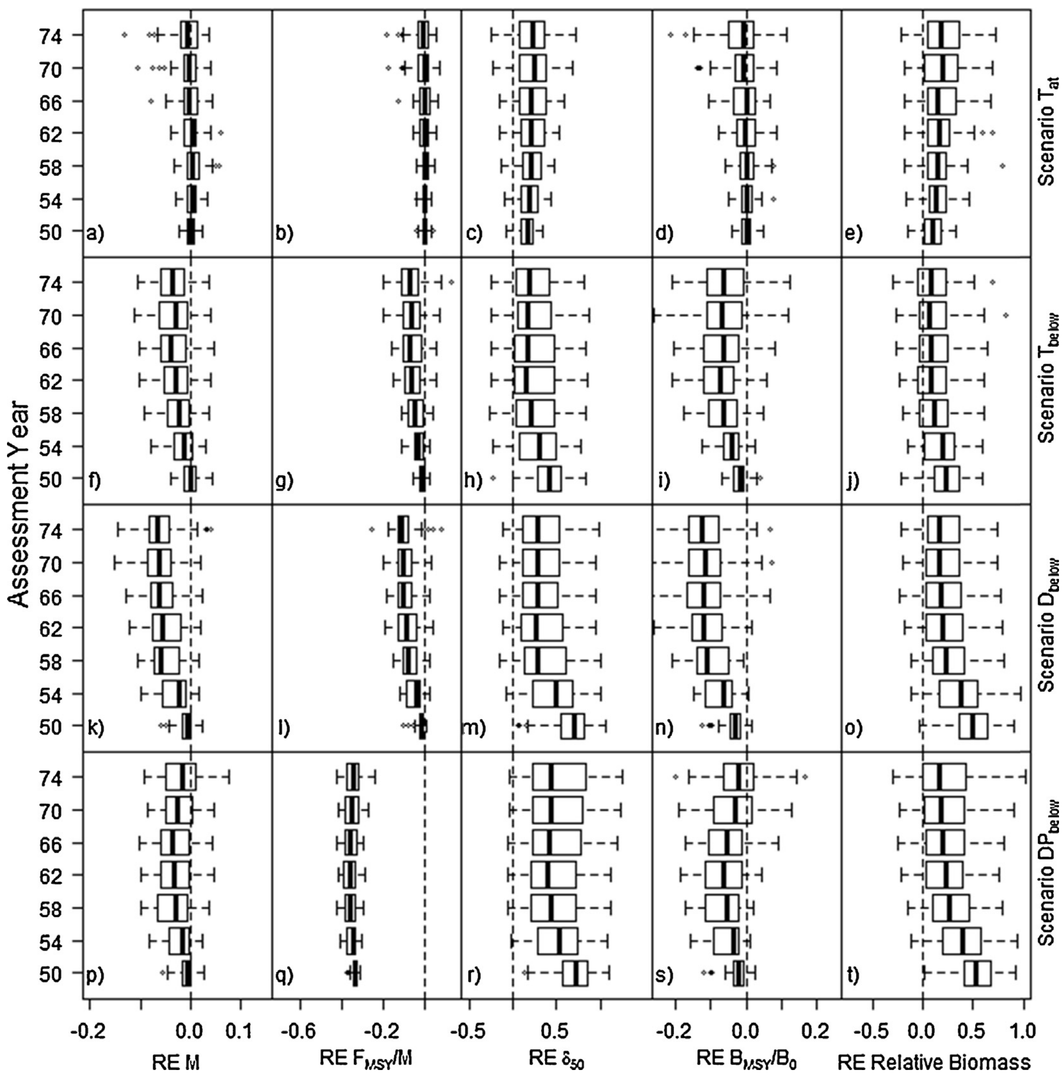

Fig. 10. The distribution of the relative error between the true values and the means of the posterior distributions for the four leading parameters (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) and relative error between the median of the posterior distribution for the assessment year stock status for XDB-SRA for the rockfish life history. The thick black line indicates the median of the posterior distribution means (M, F_{MSY}/M, δ_{50}, B_{MSY}/B_0) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

The data-moderate estimation methods that used biomass index data, XDB-SRA and XSSS, performed similarly for the rockfish life-history. There was little between-simulation variance for all methods for this life-history, especially when compared to the flatfish, due to the slow rockfish dynamics and the inherent inertia of a long-lived species. The median population trajectories for the two data-moderate methods were similar to those when the stock was managed with perfect information (dotted gray lines in Fig. 7). The average catch attained by each of these methods was often close to that under "perfect information" for all scenarios. However, harvest estimates by XDB-SRA and XSSS each resulted in overfishing in specific scenarios. The estimated OFLs and the resulting ABCs for XSSS resulted in overfishing when the productivity and relative biomass were both misspecified (DP_{below}). Estimates for XDB-SRA resulted in overfishing when the stock was at the relative biomass target and all distributions were specified correctly (T_{at}), which resulted in an increasing probability of overfishing over the projection period (Fig. 8a). In this scenario for the rockfish life-history the ABCs (OFLs reduced by the buffers) exceed true OFLs based upon the PFMC harvest control rule by only a small fraction resulting in very slight stock decline over the projection period. However, exceeding the OFL by any amount constitutes overfishing, a practice that fishery managers are mandated to prevent.

The catch-only methods, DCAC and DB-SRA, resulted in precautionary estimates for both life-histories when they were at the target relative biomass level. These methods generally estimated lower OFL values compared to the data-moderate estimation methods, and were subject to further reductions by the category 3 buffer to determine the ABC values. However, the catch-only methods, most notably DB-SRA, generally did not perform well when the prior for relative biomass was misspecified when the stock was below the management target (D_{below}), a result consistent with previous findings (Wetzel and Punt, 2011). In this scenario the probability of overfishing for the catch-only methods was often > 0.50 despite the large reduction between the OFL and ABC. Comparing the two life-histories, DCAC and DB-SRA performed somewhat better for the rockfish scenarios. Carruthers et al. (2014) drew a similar conclusion, noting that these estimation methods tended to better estimate the "windfall" biomass of the older ages classes associated with long-lived species. Although, for both life-histories, DCAC generally resulted in median population trajectories that either grew or

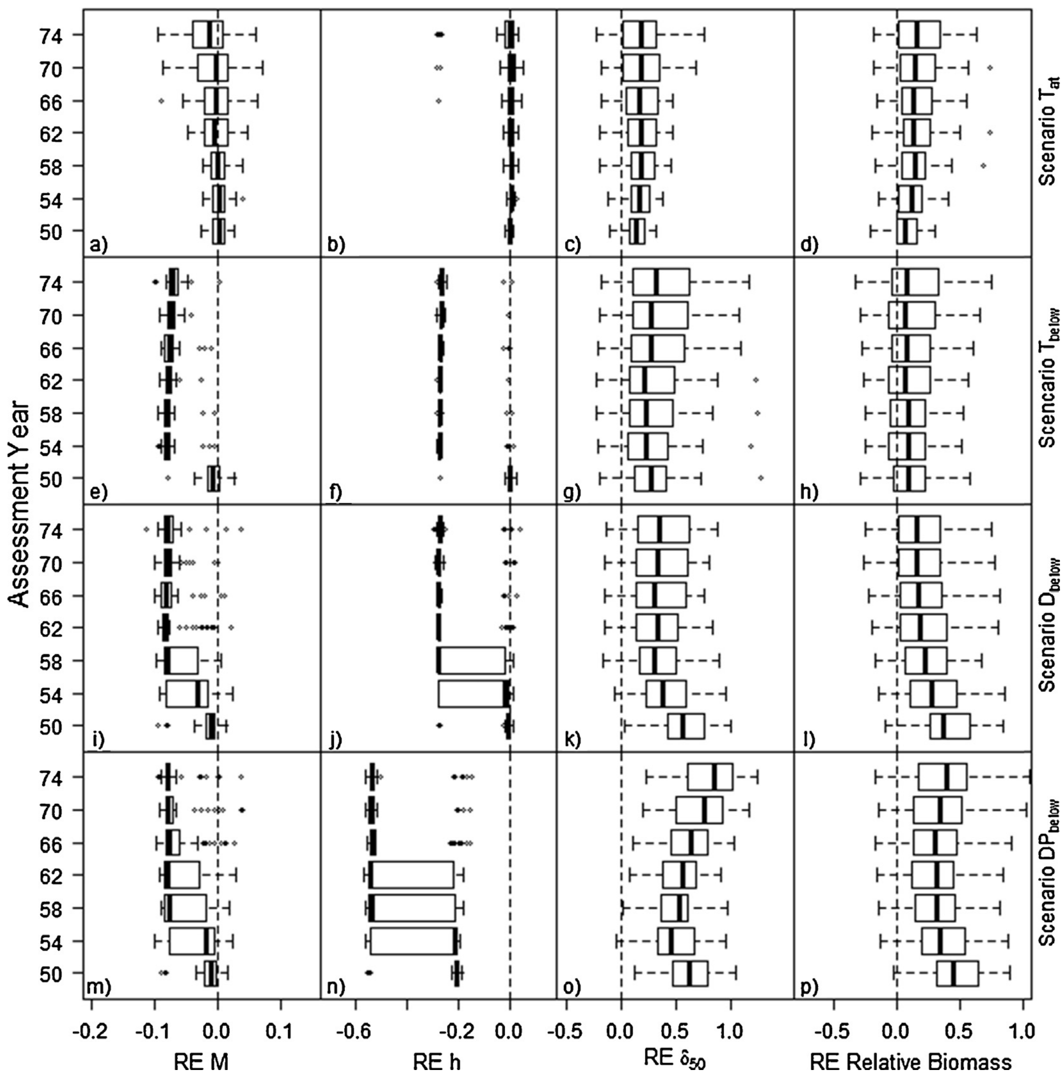

Fig. 11. Relative error between the means of the posterior distributions for the three leading parameters (M, h, δ_{50}) and the relative error between the median of the posterior distributions for assessment year stock status and their true values for XSSS for the rockfish life history, expressed as relative to the true values. The thick black line indicates the median of the posterior distribution means (M, h, δ_{50}) or medians (assessment year stock status). Results are shown for when each assessment is undertaken for each scenario.

remained flat, overfishing still occurred for portions of the projection period when harvest estimates were derived based upon misspecified parameters. This method is considered a way to estimate a yield that will likely be sustainable. However, this was not always realized when the relative biomass prior was misspecified, even when the yield was reduced to account for uncertainty.

The catch-only methods were initially created as a means to improve short-term management over yield only (average catch) based harvest estimates and were not intended for long-term management. Given the general conclusion that both catch-only methods were sensitive to misspecification, especially DB-SRA, applying either for long-term management, could have undesirable properties, especially compared to the data-moderate options which have some ability to update priors based upon data. However, on the short-term (e.g., 10 years or less) each of the methods could be very useful for management. DB-SRA resulted in OFLs that when reduced to account for uncertainty avoided overfishing for each life-history when the stocks were at the higher relative biomass levels. Based upon these results, this catch-only estimation method could be applied to determine short-term harvest estimates for stocks that have low historical exploitation. Also, DCAC generally resulted in OFLs that once reduced by the category 3 buffer (0.69) avoided overfishing over the long-term even when prior parameters were misspecified. Unfortunately, DCAC can only be applied for a one-time harvest calculation for a likely sustainable yield and alternative estimation methods would need to be applied in the future for updated harvest estimates, but once again this method could be applied for short-term estimates until more data can be incorporated for assessment. However, over the long-term, this work shows that it is beneficial to apply a method that can incorporate index data if available. The one notable exception to this general conclusion is the use of XDB-SRA for a life-history with fast dynamics and for which there is little contrast in the index data.

The posterior distributions of the model parameters for both XDB-SRA and XSSS displayed some unexpected behaviors. Specifically there were times when the posterior distributions for one, some, or all of the parameters shifted away from the true values over time. However, the resulting posterior distribution of the estimated assessment year stock status was approximately centered about the true value (e.g., rockfish T_{below}). An advantage of

XDB-SRA and XSSS is that they each have only a few parameters that have biological significance for which the user may have some a priori general idea of acceptable priors. However, the results of this analysis suggest that great care should be taken when attempting to draw conclusions from the individual posterior distributions from each of the methods for all parameters. All model misspecification is incorporated into the parameter estimates and interpreting them individually may be misleading. For example, often in data-rich assessments that contain index and composition data (e.g., length and age), the steepness parameter is fixed due to there being little information in the data (Lee et al., 2012). There should be no expectation that these data-moderate methods, applying only a biomass index, would be able to offer more information compared to data-rich methods. A proposed diagnostic for these estimation methods is to evaluate the posterior distributions for M, h or F_{MSY}/M, and B_{MSY}/B_0 for updating (Cope et al., 2015). Movement in these parameters indicates that there is information in the indices data, which is unlikely, and could be an indication of parameter misspecification.

Generally the highest probabilities of overfishing for both life-histories occurred when the simulated stock was initially below the target relative biomass level, even when all parameter priors were specified correctly, resulting in probabilities > 0.50 during the early part of the projection period (except for XSSS). Each of these methods tended to estimate OFLs that were too high, and the buffer values were insufficient to prevent overfishing for at least the first assessment. The performance of the data-moderate methods (XDB-SRA and XSSS) slowly improved as the amount of index data increased and became informative about a continuing population decline. Although the probability of overfishing decreased for the catch-only methods (DCAC and DB-SRA), this was not due to updating of the OFLs during the projection period. The estimated harvest exceeded the true OFL, which under the harvest control rule is designed to rebuild the stock, but the set ABCs were less than the corresponding productivity at the stock size allowing for population growth. If a stock is below the target level, specifying this information through the parameter distributions may not result in harvest estimates that avoid overfishing.

The results presented here highlight situations where specific estimation methods may be better suited given the data and the life-history, or when alternative methods could perform equally well, and how to interpret estimates. However, these results are dependent upon the modeling assumptions applied and additional work should be conducted to explore how robust these results are to alternative operating model structures. A strategic choice was made to incorporate only limited model misspecification between the operating model and estimation methods (e.g., stock recruitment relationship, selectivity) with process error through autocorrelated recruitment deviations and observation error about the biomass index in the operating model. This work is the first simulation test of these data-moderate methods and it was deemed critical to determine their general performance under ideal conditions to allow interpretability of the results. If a method performed poorly it was important to be able identify if that was a property of the estimation method or a particular misspecification that was assumed.

Future research should extend this work by adding more complexity to the operating model and should explore additional assumptions implemented by data-moderate and catch-only estimation methods. Additionally, the data-moderate index based estimation methods evaluated here are very time intensive to simulation test, which limited the number of scenarios that could be explored. Future work should be conducted to evaluate ways to improve the speed of these estimation methods and alternative methods for simulation testing.

The estimation methods examined here were selected because they have been applied for management of West Coast groundfish. However, there are many additional methods that have been developed to address data-limited and data-moderate stocks that range from simpler methods that apply life-history invariants to estimate maximum sustainable yield (e.g., Beddington and Kirkwood, 2005), or mean lengths to estimate spawning potential ratios (e.g., Ault et al., 2008), or methods that can incorporate catch and recent year fishing effort estimates (e.g., Walters et al., 2006) or apply a catch curve and limited recent year composition data to estimate fishing mortality (e.g., Thorson and Cope, 2015). Evaluating a collection of data-moderate and catch-only methods under a wide range of simulation scenarios will help identify when methods perform adequately and when they may lead to precautionary or risky advice for decision makers.

Acknowledgements

We would like to thank Alec MacCall, E.J. Dick, and Jason Cope for useful discussions and feedback that assisted in the formation of this work. Also, thank you to the multiple reviewers of early versions of this paper for their thoughtful comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fishres.2015.06.005

References

Ault, J.S., Smith, S.G., Lou, J., Monaco, M.E., Appeldoorn, R.S., 2008. Length-based assessment of sustainability benchmarks for coral reef fisheries in Puerto Rico. Environ. Conserv. 35, 221–231.

Beddington, J.R., Kirkwood, G.P., 2005. The estimation of potential yield and stock status using life history parameter. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 360, 163–170.

Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J., Cope, J., 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fish. Res. 153, 48–68.

Cope, J.M., 2013. Implementing a statistical catch-at-age model (stock synthesis) as a tool for deriving overfishing limits in data-limited situations. Fish. Res. 142, 3–14.

Cope, J., Dick, E.J., MacCall, A., Monk, M., Soper, B., Wetzel, C., 2015. Data-Moderate Stock Assessments for Brown, China Copper, Sharpchin, Stripetail, and Yellowtail Rockfishes and English and Rex Soles in 2013. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220, pp. 283.

Cope, J.M., Thorson, J.T., Wetzel, C.R., 2015. Evaluating a prior on relative stock status using simplified age-structured models. Fish. Res. 171, 101–109.

Dick, E.J., MacCall, A.D., 2010. Estimates of Sustainable Yield for 50 Data-Poor Stocks in the Pacific Coast Groundfish Fishery Management Plan. NOAA Technical Memorandum NMFS. NOAA-TM-NMFS-SWFSC-460., pp. 208.

Dick, E.J., MacCall, A.D., 2011. Depletion-Based Stock Reduction Analysis: a catch-based method for determining sustainable yields for data-poor fish stocks. Fish. Res. 110, 331–341.

Givens, G.H., Raftery, A.E., 1996. Local adaptive importance sampling for multivariate densities with strong nonlinear relationships. J. Am. Stat. Assoc. 91, 132–141.

Kimura, D.K., Balsinger, J.W., Ito, D.H., 1984. Generalized stock reduction analysis. Can. J. Fish. Aquat. Sci. 41, 1325–1333.

Kinas, P.G., 1996. Bayesian fishery stock assessment and decision making using adaptive importance sampling. Can. J. Fish. Aquat. Sci. 53, 414–423.

Lee, H.H., Maunder, M.N., Piner, K.R., Methot, R.D., 2012. Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models? Fish. Res. 125–126, 254–261.

MacCall, A.D., 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES J. Mar. Sci. 66, 2267–2271.

McAllister, M.K., Babcock, E.A., Pikitch, E.K., Prager, M.H., 1994. A Bayesian approach to stock assessment and harvest decision using sampling/importance resampling algorithm. Can. J. Fish. Aquat. Sci. 51, 2673–2687.

Methot, R.D., Wetzel, C.R., 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99.

Pacific Fishery Management Council (PFMC), 2010. Amendment 23: Considerations for a New Harvest Specification Framework that Incorporates Revised National Standard 1 Guidelines to Prevent Overfishing. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland OR 97220, pp. 41.

Pacific Fishery Management Council (PFMC), 2014a. Pacific Coast Groundfish Fishery Management Plan. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220, pp. 158.

Pacific Fishery Management Council (PFMC), 2014b. Status of the Pacific Coast Groundfish Fishery: Stock Assessment and Fishery Evaluation. Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220, pp. 280.

Ruben, D.B., 1987. Comment on "The calculation of posterior distributions by data augmentation.". JASA 82, 543–554.

Ruben, D.B., 1988. Using the SIR algorithm to simulate posterior distributions. In: Bayesian Statistics 3: Proceedings of the Third Valencia International Meeting, June 1–5, 1987. Clarendon Press, Oxford.

Thorson, J.T., Cope, J.M., 2015. Catch curve stock-reduction analysis: an alternative solution to the catch equation. Fish. Res. 171, 33–41.

Walters, C.J., Martell, S.J.D., Korman, J., 2006. A stochastic approach to stock reduction analysis. Can. J. Fish. Aquat. Sci. 63, 212–223.

Wetzel, C.R., Punt, A.E., 2011. Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fish. Res. 110, 342–355.