Evaluating methods for setting catch limits in data-limited fisheries

Thomas R. Carruthers, André E. Punt, Carl J. Walters, Alec MacCall,Murdoch K. McAllister, Edward J. Dick, Jason Cope

SEDAR46-RD-04

February 2016

Fisheries Research 153 (2014) 48–68

Contents lists available at ScienceDirect

Fisheries Research

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

Evaluating methods for setting catch limits in data-limited fisheries

Thomas R. Carruthers[a,*], André E. Punt[b,1], Carl J. Walters[a], Alec MacCall[c,2], Murdoch K. McAllister[a], Edward J. Dick[c,2], Jason Cope[d,3]

[a] Fisheries Center, AERL, University of British Columbia, 2202 Main Mall, Vancouver, BC, Canada V6T 1Z4
[b] School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195, United States
[c] National Marine Fisheries Service, Southwest Fisheries Science Center, Fisheries Ecology Division, 110 Shaffer Road, Santa Cruz, CA 95060, United States
[d] National Marine Fisheries Service, Northwest Fisheries Science Centre, 2725 Montlake Boulevard East, Seattle, WA 98112, United States

ARTICLE INFO

Article history:
Received 27 June 2013
Received in revised form 9 December 2013
Accepted 15 December 2013
Handling Editor Dr. K.A. Rose
Available online 3 February 2014

Keywords:
Data-limited
Data-poor
Management strategy evaluation
Catch limits
Simulation evaluation
Stock assessment

ABSTRACT

The majority of global fish stocks lack adequate data to evaluate stock status using conventional stock assessment methods. This poses a challenge for the sustainable management of these stocks. Recent requirements to set scientifically based catch limits in several countries, and growing consumer demand for sustainably managed fish have spurred an emerging field of methods for estimating overfishing thresholds and setting catch limits for stocks with limited data. Using a management strategy evaluation framework we quantified the performance of a number of data-limited methods. For most life-histories, we found that methods that made use of only historical catches often performed worse than maintaining current fishing levels. Only those methods that dynamically accounted for changes in abundance and/or depletion performed well at low stock sizes. Stock assessments that make use of historical catch and effort data did not necessarily out-perform simpler data-limited methods that made use of fewer data. There is a high value of additional information regarding stock depletion, historical fishing effort and current abundance when only catch data are available. We discuss the implications of our results for other data-limited methods and identify future research priorities.

1. Introduction

The majority of global fish stocks lack adequate catch, survey, and other biological data to calculate current abundance and productivity using conventional stock assessment methods. In developed countries, the fraction of fish stocks that are assessed ranges between 10 and 50%. This fraction is generally lower in developing countries where it ranges between 5 and 20% (Costello et al., 2012). This poses a significant challenge for the sustainable management of these stocks. Recent requirements to set scientifically based catch limits in countries such as Australia, New Zealand, and the United States, along with growing consumer demand for sustainably managed fish, have spurred an emerging field of methods for estimating overfishing thresholds and setting catch limits for stocks with limited data.

In 2006, the U.S. Magnuson-Stevens Fishery Conservation and Management Act was amended to require annual catch limits (ACLs) to prevent overfishing for most federally managed fish stocks, including many data-limited stocks. According to the National Marine Fisheries Service's (NMFS's) National Standard 1 Guidelines (2009), setting ACLs is a three-step process that begins by identifying an overfishing limit (OFL). The OFL is the annual catch when fishing the stock's current abundance at the maximum sustainable fishing mortality rate (F_{MSY}). In the second step, a harvest control rule is used to determine the acceptable biological catch (ABC). The ABC is a catch level equal to or less than the OFL that accounts for the scientific uncertainty in the estimate of the OFL. Finally, fisheries managers use the ABC to establish an ACL. The ACL is set to a level equal to or below the ABC and accounts for various ecological, social and economic factors in addition to uncertainty in management controls.

The most established basis for estimating an OFL is by a conventional stock assessment, which typically uses fishery time series data to estimate current stock size and productivity. However, many populations have insufficient fishery catch data, survey data, or information about life-history characteristics to support a conventional stock assessment, requiring the use of alternative, data-limited methods. Most data-limited methods are designed to operate on a single time series of annual catches (generally no fishing effort or survey data are available) with additional

* Corresponding author. Tel.: +1 604 822 6903; fax: +1 604 822 8934.
E-mail address: t.carruthers@fisheries.ubc.ca (T.R. Carruthers).
[1] Tel.: +1 206 221 63193.
[2] Tel.: +1 831 420 3900.
[3] Tel.: +1 206 860 3200.

0165-7836
http://dx.doi.org/10.1016/j.fishres.2013.12.014

user-specified inputs for fisheries characteristics, demographic parameters, exploitation rate and/or stock status. Many of these methods are now being used in management, although they have not been thoroughly tested. Management strategy evaluation (MSE) is an appropriate tool to evaluate and compare the performance of existing methods across various types of fish stocks and relative population levels (see Section 2.2 for a detailed description of MSE). We use MSE in this research to test the performance of data-limited methods for various stock types and depletion levels (depletion is defined here as current biomass divided by unfished biomass).

It may be possible to make reasonable qualitative statements about the performance of various data-limited methods without undertaking an MSE. However detailed simulation evaluation enables the relative performance of methods to be quantified to support strategic decisions regarding data collection and selection of methods. Previous simulation evaluations of data-limited OFL-setting methods and ABC control rules have been conducted by Wetzel and Punt (2011) and Wilberg et al. (2011). Wetzel and Punt (2011) evaluated the performance of two methods (DB-SRA and DCAC) over a range of population and fishery dynamics. Limitations of their approach include the simulation of a relatively narrow range of fishery dynamics without simultaneously considering a realistic level of uncertainty and bias in all of the inputs to the methods under examination (e.g., natural mortality rate, *M*). Wilberg et al. (2011) simulation tested a more comprehensive range of data-limited methods. However, not all data-limited methods were applied to all stock types preventing a complete performance comparison (Vaughan et al., 2012). Their approach was also criticized on the basis of a relatively narrow range of simulated life-histories and discrete simulation of error and bias. We aim to address these criticisms by (1) simulating a wide range of fishery and population dynamics and (2) assigning probability distributions for bias and imprecision to more of the inputs to data-limited methods (e.g., depletion, *M*). Such an approach may better reveal the trade-offs among management objectives and provide a more detailed account of the performance characteristics of data-limited methods.

2. Materials and methods

This research is aimed at evaluating methods that determine an ABC as a basis for setting annual catch limits. Twenty-five methods for determining OFLs and modifying them using ABC control rules are evaluated, including nine that have been used in the management of U.S. fisheries (M1–M9), 12 alternative methods (A1–A12), and four reference methods that can be used to comparatively assess the performance of the other methods (R1–R4).

The methods are classified as follows: (1) those that rely on a time series of recent catch ("catch-based methods"); (2) those that adjust historical catches using assumptions about historical depletion and life history characteristics ("depletion-based methods"), and (3) those that rely on current estimates of absolute abundance ("abundance-based methods"). Methods within these classes can be further distinguished into those methods that dynamically update with current information on depletion and those that remain static. The following section describes the specific methods selected for evaluation (see Table 1 for a list of all methods). The data requirements of each method tested are summarized in Table 2, and their detailed description can be found in Appendix B.

These methods are subject to modification by two types of ABC control rule. The first is no downward adjustment. For example, methods M1–M3 are catch methods for which ABC equals the OFL. The second type of ABC control rule uses a simple scalar approach in which a point value produced by a method (e.g., the median outcome of DB-SRA or DCAC) is multiplied by a factor. These scalar factors differ depending on a broadly defined characterization of scientific uncertainty for different groups of stocks (e.g., alternative methods A1, A2 and A7–A12 make use of 75% and 100% scalars).

2.1. Methods evaluated in this study

2.1.1. Catch-based methods

Catch-based methods have generally been employed where insufficient data exist for determining an OFL using more sophisticated methods. For example, the U.S. Southeast and Mid-Atlantic Fishery Management Councils currently apply catch-based methods to dozens of stocks. The South Atlantic Fishery Management Council (SAFMC) has adopted two quantitative approaches to ACL-setting that are simulation tested: the OFL is set to the third highest landings over the last ten years or to the median landings over the last ten years (SAFMC, 2011). The Mid-Atlantic Fishery Management Council has adopted an OFL for Atlantic Mackerel that is the median catch from the last three years (MAFMC, 2010; NMFS, 2011). These approaches stem from the work of Restrepo et al. (1998) who suggested the use of average catches with a downward adjustment based on uncertainty about stock status, although these implementations do not include a downward adjustment. All three of these methods are tested: the median catch over the most recent three years (M1), the median catch over the most recent 10 years (M2), and the third-highest catch over the most recent 10 years (M3).

Other catch-based methods that have been proposed attempt to introduce dynamic updates of simple catch-based control rules based on generally subjective scoring systems, such as the Only Reliable Catch Stocks (ORCS, Berkson et al., 2011) method and Productivity–Susceptibility Analysis (PSA, Patrick et al., 2009). Both of these approaches use biological and fishery characteristics to calculate a single catch value. Berkson et al. (2011) identify a possible means of using the outcome from ORCS to categorize stocks into exploitation levels. Each level leads to a different multiplication of interquartile mean catch (the average of all catches greater than the 25th percentile and less than the 75th percentile) that is selected as a proxy for the OFL or ABC. PSA has been suggested as a basis for an ABC control rule that increases the precautionary buffer with increasing vulnerability of the stock (Berkson et al., 2011). Unfortunately, it proved difficult to test these approaches due to an inability to simulate the subjective scoring systems in a defensible way. The success of the methods is likely to be determined by how they are implemented, so we decided to omit them from the comparative performance analysis.

Instead of simulating these subjective methods we tested a control rule similar to that proposed by Berkson et al. (2011). This control rule dynamically scales a catch-based OFL according to periodic estimates of depletion. The OFL is set to half, equal or twice the interquartile mean catch when current biomass is considered to be less than 20% of unfished, greater than 20% and less than 65% of unfished, and greater than 65% of unfished levels, respectively. In lieu of a subjective scoring system to estimate depletion, we test the performance of the catch scalar methods using imperfect knowledge of simulated current depletion. An imperfect estimate of depletion was simulated by calculating the current level of stock depletion (current biomass divided by unfished biomass) and then adding error according to specified levels of bias and imprecision. This method (referred to as "Depletion Adjusted Catch Scalar", DACS) represents a very simple approach to modifying an OFL using coarse subjective information about current stock levels. We test the DACS method with two ABC control rules: 75% and 100% scalars (methods A1 and A2).

Table 1
A summary table of the methods tested in this management strategy evaluation, including nine methods currently in use in the management of stocks in U.S. fishery management plans (M1–M9), 12 alternative methods described in the peer-reviewed literature (A1–A12) and four reference methods (R1–R4).

Type	Code	Name	OFL setting	ABC control rule	Source
Static methods					
Catch-based (static)	M1	Median catch – 3 years	Median catch over last 3 years	None	MAFMC
	M2	Median catch – 10 years	Median catch over last 10 years	None	SAFMC
	M3	3rd highest catch	3rd highest catch over last 10 years	None	SAFMC
Depletion-based (static)	M4	DB-SRA (depletion fixed @ 40% B_0) – 69.4% scalar	Median of OFL distribution	69.4% scalar	PFMC (Dick and MacCall, 2011)
	M5	DB-SRA (depletion fixed @ 40% B_0) – 83.4% scalar	Median of OFL distribution	83.4% scalar	PFMC (Dick and MacCall, 2011)
	M6	DB-SRA (depletion fixed @ 40% B_0) – 91.3% scalar	Median of OFL distribution	91.3% scalar	PFMC (Dick and MacCall, 2011)
	M7	DCAC (depletion fixed @ 40% B_0) – 69.4% scalar	Median of OFL distribution	69.4% scalar	PFMC (Dick and MacCall, 2010)
	M8	DCAC (depletion fixed @ 40% B_0) – 83.4% scalar	Median of OFL distribution	83.4% scalar	PFMC (Dick and MacCall, 2010)
	M9	DCAC (depletion fixed @ 40% B_0) – 91.3% scalar	Median of OFL distribution	91.3% scalar	PFMC (Dick and MacCall, 2010)
Dynamic methods					
Catch-based (dynamic)	A1	Depletion adjusted catch scalar – 75% scalar	0.5, 1.0, or 2.0 × mean landings	75% scalar	Berkson et al. (2011)
	A2	Depletion adjusted catch scalar – 100% scalar	0.5, 1.0, or 2.0 × mean landings	100% scalar	Berkson et al. (2011)
Depletion-based (dynamic)	A3	DB-SRA (depletion adjusted) – 25% P^*	Stochastic model output	25% P^*	Dick and MacCall (2011)
	A4	DB-SRA (depletion adjusted) – 50% P^*	Stochastic model output	50% P^*	Dick and MacCall (2011)
	A5	DCAC (depletion adjusted) – 25% P^*	Stochastic model output	25% P^*	Dick and MacCall (2011)
	A6	DCAC (depletion adjusted) – 50% P^*	Stochastic model output	50% P^*	Dick and MacCall (2011)
Abundance-based (dynamic)	A7	Life history analysis – 75% scalar	F_{MSY} × abundance	75% scalar	Beddington and Kirkwood (2005)
	A8	Life history analysis – 100% scalar	F_{MSY} × abundance	100% scalar	Beddington and Kirkwood (2005)
	A9	F_{MSY}/M (low) – 75% scalar	F_{MSY} @ 0.5 M × abundance	75% scalar	Gulland (1971) and Walters and Martell (2002)
	A10	F_{MSY}/M (low) – 100% scalar	F_{MSY} @ 0.5 M × abundance	100% scalar	Gulland (1971) and Walters and Martell (2002)
	A11	F_{MSY}/M (hi) – 75% scalar	F_{MSY} @ 0.8 M × abundance	75% scalar	Gulland (1971) and Walters and Martell (2002)
	A12	F_{MSY}/M (hi) – 100% scalar	F_{MSY} @ 0.8 M × abundance	100% scalar	Gulland (1971) and Walters and Martell (2002)
Reference cases					
Stock assessment (dynamic)	R1	Delay-difference – 75% scalar	Delay-difference assessment	75% scalar	Deriso (1980) and Schnute (1985)
	R2	Delay-difference – 100% scalar	Delay-difference assessment	100% scalar	Deriso (1980) and Schnute (1985)
Status quo (static)	R3	Current catch	Catch in last simulated year	None	N/A
	R4	Current effort	Effort in last simulated year	None	N/A

Table 2
The data requirements or inputs of the data-limited methods tested in this evaluation. These include a time series of historical catches (Catch), current stock size relative to unfished condition (Depltn), the ratio of fishing mortality rate at maximum sustainable yield to the natural mortality rate (F_{MSY}/M), biomass at maximum sustainable yield relative to unfished biomass (B_{MSY}/B_0), natural mortality rate (M), median age at maturity, current biomass, the rate parameter K of the von Bertalanffy growth equation (Von Bert. K) and the mean length at first capture.

Type	Code	Name	Catch	Depltn.	F_{MSY}/M	B_{MSY}/B_0	M	Age at 50% Maturity	Current biomass	Von Bert. K	Length-at-first capture
Static Methods											
Catch-Based (Static)	M1	Mean Catch - 3 Years	■								
	M2	Median Catch - 10 Years	■								
	M3	3rd Highest Catch	■								
Depletion-Based (Static)	M4	DB-SRA (Depletion Fixed @ 40%B_0) - 69.4% scalar	■		■	■	■	■			
	M5	DB-SRA (Depletion Fixed @ 40%B_0) - 83.4% scalar	■		■	■	■	■			
	M6	DB-SRA (Depletion Fixed @ 40%B_0) - 91.3% scalar	■		■	■	■	■			
	M7	DCAC (Depletion Fixed @ 40%B_0) - 69.4% scalar	■		■	■	■				
	M8	DCAC (Depletion Fixed @ 40%B_0) - 83.4% scalar	■		■	■	■				
	M9	DCAC (Fixed Depletion @ 40%B_0) - 91.3% scalar	■		■	■	■				
Dynamic Methods											
Catch-Based (Dynamic)	A1	Depletion Adjusted Catch Scalar - 75% scalar	■	■							
	A2	Depletion Adjusted Catch Scalar - 100% scalar	■	■							
Depletion-Based (Dynamic)	A3	DB-SRA (Depletion Adjusted) - 25% P*	■	■	■	■	■	■			
	A4	DB-SRA (Depletion Adjusted) - 50% P*	■	■	■	■	■	■			
	A5	DCAC (Depletion Adjusted) - 25% P*	■	■	■	■	■				
	A6	DCAC (Depletion Adjusted) - 50% P*	■	■	■	■	■				
Abundance-Based (Dynamic)	A7	Life History Analysis - 75% scalar							■	■	■
	A8	Life History Analysis - 100% scalar							■	■	■
	A9	F_{MSY}/M (Low) - 75% scalar					■		■		
	A10	F_{MSY}/M (Low) - 100% scalar					■		■		
	A11	F_{MSY}/M (Hi) - 75% scalar					■		■		
	A12	F_{MSY}/M (Hi) - 100% scalar					■		■		
Reference Cases											
Stock Assessment (Dynamic)	R1	Delay-Difference - 75% scalar	■					■		■	■
	R2	Delay-Difference - 100% scalar	■					■		■	■
Status Quo (Static)	R3	Current Catch	■					■		■	■
	R4	Current Effort	■					■		■	■

2.1.2. Depletion-based methods

These data-limited methods rely on estimates of depletion relative to an unfished population, combined with other inputs to estimate an OFL directly or to adjust historical catch with historical depletion to derive a catch level recommendation. Depletion-Based Stock Reduction Analysis (DB-SRA, Dick and MacCall, 2011) is a method for estimating an OFL based on a complete time series of historical catches and four key inputs: (1) the level of current depletion, (2) the ratio of F_{MSY} to the natural mortality rate (F_{MSY}/M), (3) the natural mortality rate (M) and (4) the most productive stock size relative to unfished (B_{MSY}/B_0). Given input values for M, F_{MSY}/M and B_{MSY}/B_0, DB-SRA finds a stock reconstruction that matches the input level of depletion and historical catch. DB-SRA then calculates the OFL by multiplying together F_{MSY}, depletion, and the reconstructed unfished biomass. The process is stochastic, and samples many values for all four inputs, each sample leading to an estimate of unfished biomass and therefore an OFL recommendation (see Appendix B.1 for details). DB-SRA also requires an estimate of the age at which fish become recruited to the fishery since it assumes delay-difference stock dynamics.

Depletion-Corrected Average Catch (DCAC, MacCall, 2009) provides an estimate of "sustainable catch" based on an estimate of average annual catch and the same four key inputs as DB-SRA (depletion, F_{MSY}/M, M and B_{MSY}/B_0). In essence, DCAC calculates average catches accounting for the removal of the "windfall harvest" of less productive biomass that may have occurred as the stock became depleted (the equations are included in the Appendix B.1). DCAC requires the same inputs as DB-SRA and is also stochastic, sampling many input values to produce numerous estimates of "sustainable catch."

Both DB-SRA and DCAC are currently being used to set OFLs and ABCs for data-limited stocks by the Pacific Fishery Management Council (PFMC, 2010). Different ABC control rules are applied depending on the degree of scientific uncertainty for different stocks. The Pacific Fishery Management Council's implementation of DB-SRA and DCAC assumes that current depletion is, on average, 40% of unfished biomass – for many stocks this may be considered a productive and healthy stock size (Dick and MacCall, 2010). These methods also do not make direct use of the stochastic OFL output of DB-SRA and DCAC. Instead, a downward adjustment is achieved by superimposing a distribution (with a pre-specified variance) over the median OFL estimate from DB-SRA and DCAC. It is a percentile of this superimposed distribution that is used as the ABC. Three versions of DB-SRA and DCAC are tested that rely on distributions for depletion which are centered on 40% of unfished biomass. The OFL for each method is then adjusted according to the same ABC control rules applied to different categories of data-limited stocks by the PFMC (M4–M9, Appendix B.1).

Two generic implementations of DB-SRA and DCAC were tested (A3–A6) that include dynamic updates in depletion (they are linked to the actual simulated level of stock depletion and do not rely on a fixed assumption of 40% unfished biomass). These implementations also make direct use of the stochastic output of DB-SRA and DCAC to derive the ABC based on pre-specified percentiles (25% and 50%).

2.1.3. Abundance-based methods

As an alternative to data-limited methods that rely solely or primarily on catch data and/or depletion estimates we tested a class of methods that rely on estimates of current abundance and F_{MSY}. While methods such as DB-SRA attempt to reconstruct historical stock levels, abundance-based methods rely only on current data. The methods that use current biomass are also not reliant on historical catch data and there is no positive feedback from previous management recommendations (the catch prescribed in one year does not directly inform the next catch recommendation). These methods also rely on weaker assumptions of stationary population and fishery dynamics.

We examine two methods of quantifying F_{MSY} based on growth and natural mortality rate. Beddington and Kirkwood (2005) describe a method for calculating F_{MSY} using length at first capture and information about maximum growth rate of individuals. Simpler still are methods that assume a fixed value for F_{MSY}/M. The originator of this concept, Gulland (1971), assumed $F_{MSY} = M$. Subsequent publications have recommended lower ratios of 0.8 (Thompson, 1993) and 0.5 (Walters and Martell, 2002). An estimate of current biomass is required to apply these approaches. The North Pacific Fishery Management Council (NPFMC) currently uses an F_{MSY}/M ratio method for managing stocks for which typical stock assessment reference points are not available ('Tier 5' stocks NPFMC, 2012,2013, referred to as 'data poor' by DiCosimo et al., 2010). Six variants of the abundance-based method are considered (A7–A12) depending on the assumed ratio of F_{MSY} to M, and the assumed ratio of the ABC to the OFL.

2.1.4. Reference cases

Four reference cases are included to provide a yardstick for the performance of the methods described above (R1–R4). We test a stock assessment method based on a delay-difference model (Deriso, 1980; Schnute, 1985) (R1–R2), which may be applied in instances where catch age- and length-composition data are not available (similar population dynamics are assumed by DB-SRA). The delay-difference assessment also requires auxiliary information regarding the form of the stock-recruit function, the fraction of mature fish-at-age, body growth rate, natural mortality rate, and the vulnerability-at-age curve. It calculates the OFL directly from estimates of current biomass and F_{MSY}. The performance of 100% and 75% scalar ABC control rules is evaluated. Similar to the data-limited methods, the delay-difference stock assessment method has inputs that are subject to imperfect information regarding historical catches. The delay-difference reference cases may be expected to perform better than the data-limited methods that only make use of catch data. Two "status quo" reference cases are simulated to frame the results of the data-limited methods in terms of two non-adaptive methods: (R3) a constant current catch scenario and (R4) a constant current effort scenario.

2.2. Management strategy evaluation

Experimental evaluation of methods for setting OFLs and ABCs through manipulation and monitoring of wild populations is impractical. Previous research has sought to compare the outputs of data-limited methods with those of data-rich assessments given the same data (e.g., Dick and MacCall, 2011). The principal limitation of this approach is the difficulty in assessing risks, and the inability to quantify bias. For example, relatively large differences in predicted fishing mortality rate (F) between an assessment and a data-limited method may not translate to commensurate differences in the risks of certain events occurring (e.g., the probability of reduction in biomass below B_{MSY}). Stock assessment models typically make use of common assumptions that may bias their results in similar ways (e.g., not accounting for habitat degradation, spatial expansion of fishing, or increases in fishing efficiency), and may therefore provide a limited basis for comparative performance evaluation. Equally, the stocks that are subject to assessment may not be representative of those with limited data; perhaps due to economic value they are heavily exploited or conversely subject to stringent management. Fundamentally, it is not possible to evaluate the accuracy of a data-limited method without knowledge of the quantity which is to be estimated (e.g., actual abundance or simulated abundance). For these reasons simulation evaluation is recommended as an important first step in testing data-limited methods (Butterworth et al., 2010).

Management Strategy Evaluation (MSE, Cochrane et al., 1998; Butterworth and Punt, 1999) is a simulation approach which generates many realizations of a real fishery system encompassing a credible range of population and exploitation scenarios. The simulated reality, commonly referred to as the "operating model," is then projected forward in time and updated according to the ACL recommendations generated by a particular management method (the ACL is assumed to be the ABC in this study). The relative performance of each management strategy can then be evaluated relative to defined management objectives. MSE also provides an opportunity to better understand the trade-offs among management objectives for any given management method and to quantify the value of various types of information and data. The core requirements of the MSE approach are the operating model that describes the "true" simulated population (Section 2.3), a range of candidate management methods (Section 2.1), and criteria for evaluating the performance of management methods (Section 2.7). Fig. 1 describes the components of the MSE design as it related to this research.

2.3. Operating model

The operating model is parameterized for six life-history types (also referred to as "stocks" or "simulated stocks"): mackerel (*Scombridae*), butterfish (*Stromateidae*), snapper (*Lutjanidae*), porgy (*Sparidae*), sole (*Pleuronectidae*) and rockfish (*Sebastidae*).[4] In addition to providing diversity in life-history, these stocks also represent generic versions of real-world stocks that appear in various geographic regions. Populations were first simulated for 50 years using random selections for various parameters. This duration was sufficiently long to develop a range of exploitation patterns over a length of time similar to industrial fishing in US waters. Management reference points such as maximum sustainable yield (MSY),

[4] The results of this research should not be interpreted as empirical support for the status of real-world fish stocks.

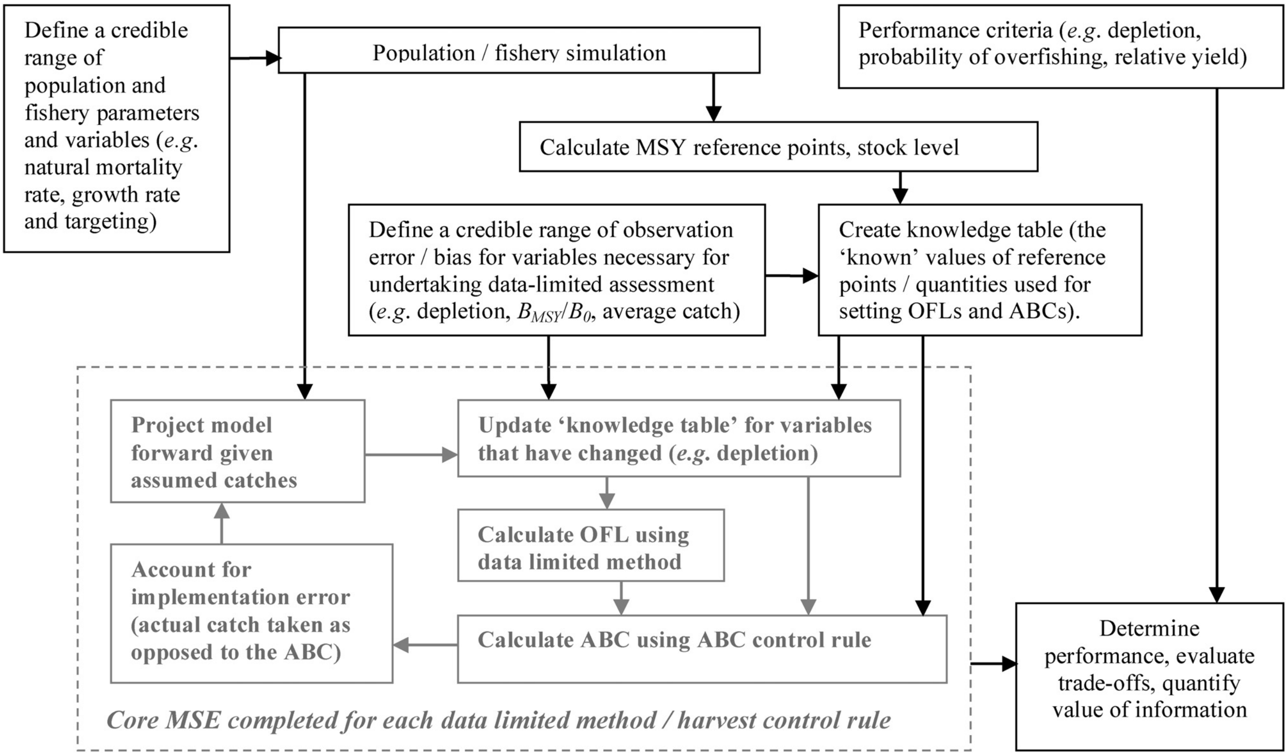

Fig. 1. A flow diagram of the components of the MSE for any given stock. The dashed box represents the projection of the model and update according to a particular combination of data-limited OFL setting method (e.g., DB-SRA) and ABC control rule (e.g., the P^* approach).

B_{MSY}, and F_{MSY} were then calculated for each simulation. Bias and imprecision in the knowledge of the simulated system were generated for all variables and parameters used by the management methods (e.g., M, current biomass, etc.). Each simulation was then projected forward subject to the ABC recommendations from each of the management methods. This update of information and setting of a new ABC was simulated every three years of the projection period to approximate a typical assessment cycle. To provide meaningful advice over a time-scale relevant to each stock, generation time was used as a basis for setting the number of projected years. Simulations were projected for a maximum of either 30 years or twice the mean generation time. The rockfish stock, with a generation time of 25 years, was projected for 50 years.

For each of the six stock types we carried out 10,000 simulations for each data-limited method. A much lower level of replication was required to obtain stability in aggregate performance metrics (the difference was less than 2% between 2000 and 3000 simulations for such metrics). However a larger degree of replication was required to provide plots of trends in performance with changing simulation parameters. The simulation evaluation framework was programmed in the statistical environment R (2.15.0 64bit, R Development Core Team, 2012) using the "Snowfall" package for parallel computing.

The "branched" form of experimental design (Fig. 2) allows management methods to be compared side-by-side because projections are made from the same set of historical simulations and the same future recruitment patterns. An additional benefit of this design is that the performance of any management method can be phrased in terms of a "best case" reference method based on identical conditions. For example, we standardized the predicted yield of a particular management method for any given simulation by dividing it by the "best case" yield that could be obtained with perfect knowledge of F_{MSY}.

The operating model was an age-structured, spatial model (a detailed description can be found in Appendix A). Simulating

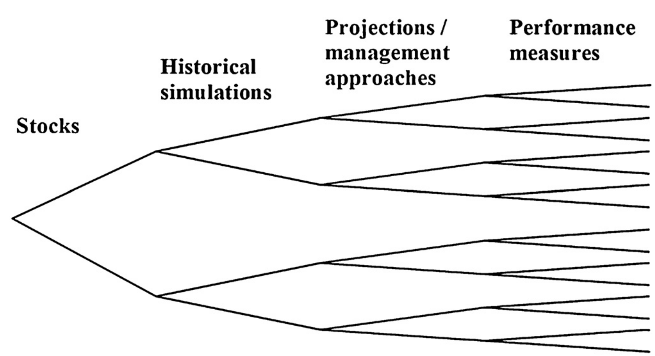

Fig. 2. The "branched" design of the simulation evaluation including six stock types, 50 historical years, 30–50 projected years, 25 data-limited and reference methods, and 7 performance measures.

spatial dynamics provided the basis to account for differences among life-history types that may be considered important, such as low mixing among areas and refuges from fishing. All stocks are assumed to have density-dependent recruitment that does not decrease with increasing stock size, and maximum surplus recruitment is achieved when spawning output is less than half of unfished (Beverton and Holt, 1957). For the purposes of simulation, variability among simulations and where applicable, inter-annual variability within simulations, were generated in a number of biological parameters such as M, stock-recruitment parameters and recruitment deviations. Auto-correlation in recruitment was not simulated. The location and slope of the age-at-maturity curve, weight-at-length curve and scale parameters such as unfished stock size and maximum length did not vary among simulations for the same stock.

Five discrete areas were modeled for each population. The operating model generated both directed and diffusive movement among areas by adjusting regional gravity parameters and a stock

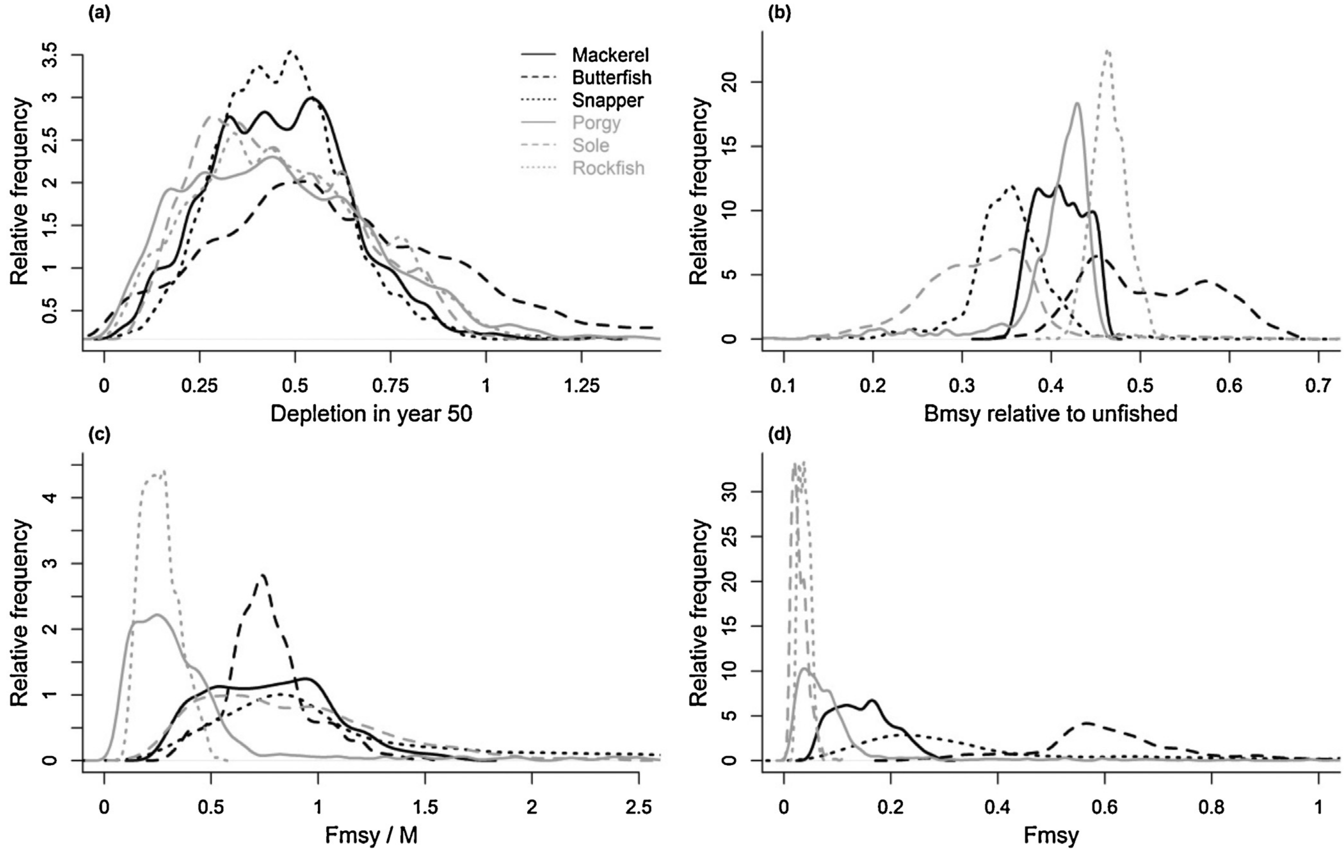

Fig. 3. The historical simulation conditions (10,000 simulations). Plotted in panel a are the relative frequencies of sampled depletion (the biomass in year 50, the final historical year, divided by unfished biomass). Panel b describes the sampled ratio of B_{MSY}/B_0. Plotted in panel c are the relative frequencies of the sampled ratio of F_{MSY}/M. Panel d describes the sampled distribution of F_{MSY}.

mixing ("viscosity") parameter (Eqs. (App.A.27) and (App.A.28)). With the exception of recruitment deviations, all population dynamics parameters were assumed to be time-invariant. Simulations were also conducted without spatial structure to evaluate the sensitivity of results to spatial dynamics.

For each simulation a single trend in fishing effort was generated. This time series represents the total effort on the stock from all sources of fishing. Among simulations both the mean trend and inter-annual variability in effort was allowed to vary (see Appendix App.A.2 for full details). For all simulations mean trends always increased during the first 25 years. Subsequently fishing effort could range from a strong decline to a steep increase over the last 25 historical years. The same inter-annual variation in fishing effort was simulated for each stock with a coefficient of variation (*CV*) ranging from 0.2 and 0.4. For all stocks, catch observation error was sampled over a range for the *CV* of 0.1–0.5. Some species-specific fishery characteristics were specified, including vulnerability-at-age, spatial targeting (or avoidance) and spatial refuges from fishing. While fishing effort, targeting and fishing efficiency could change temporally, all other fishery characteristics were assumed to remain constant over time.

2.4. Defining simulations for specific stocks

The operating model inputs for each stock are summarized in Table App.A.1. Some of these inputs describe a range from which a value is sampled (e.g., M uniformly sampled between 0.2 and 0.4 yr^{-1}). The number of areas (5), historical simulation years (50), the level of unfished recruitment, the rate of catch observation error and the variability in the simulated trend in effort are the same for each stock.

Fifty years of historical projection prior to first application of the management methods (Fig. 3) led to a wide range of depletions that were nevertheless comparable among stocks so that conclusions were not confounded by stock-specific depletion levels. All stocks had mean depletion values close to 45% at the end of the historical simulation period (Fig. 3). The exception is butterfish which, due to a short life-span and high recruitment variability, could not be made comparable to the depletion distributions of the other stocks. The six life-history types span a reasonably wide range of values for B_{MSY}/B_0 (mean simulated values in the range of 0.33 for sole to 0.52 for butterfish). The range for F_{MSY}/M among stocks was greater, with mean values between 0.27 (rockfish) and 1.4 (snapper). F_{MSY} varied widely among stocks, with mean rates of 0.05 for sole and 0.6 for butterfish.

2.5. Calculating MSY reference points

B_{MSY} and F_{MSY} are required to evaluate the performance of data-limited methods (Section 2.7). These quantities were computed for each simulation by projecting the operating model forward for 100 years, numerically optimizing for the fishing effort that provided the maximum yield. Optimizations were undertaken assuming that future recruitment is deterministically related to the stock-recruit relationship, and that there are no changes in fisheries targeting and catchability. Optimizations to find F_{MSY} were conducted via successive parabolic interpolation using the function 'optimize' of the R stats package.

2.6. Simulating imperfect knowledge

There may be considerable uncertainty regarding the inputs to the management methods. Imperfect knowledge of these quantities was simulated by adding error to the "true" simulated values of the operating model. Since these inputs are likely to control the relative performance of the methods they are assigned ranges that are considered to be representative of the magnitude of uncertainty in a data-limited setting. An additional purpose for generating imperfect information is to determine the effect of the misspecification of inputs on the performance of a particular management method. A related objective is quantifying the value of more precise and/or accurate information regarding population variables (e.g., current stock depletion) and parameters (e.g., M).

Table 3 describes how bias (and in some cases imprecision) was introduced to operating model parameters that are used by the management methods. All such variables have the subscript "*obs*" to denote an observed quantity. For example, M_{obs} is the simulated value of M, subject to variable bias determined by a coefficient of variation parameter CV_M. In each simulation the same biased level of M_{obs} is used by the methods throughout the projection to determine OFLs and ABCs. In some cases, data-limited methods require inputs that are updated annually as the population is projected (e.g., current biomass $Bcur_{obs}$, current depletion, and current fishing mortality rate). Both bias and imprecision are simulated in such instances. For example, $Bcur_{obs}$ is the simulated "true" current biomass ($Bcur$), subject to error sampled in each projected year according to a bias (μ_{Bcur}) and imprecision (σ_{Bcur}) that are perpetuated over the whole projection (on average, inputs were allowed to be positively or negatively biased and precise or imprecise over the whole projection). The rationale for the values of these inputs is explained further in Appendix A.5.

2.7. Evaluating performance

Performance of the data-limited and reference methods were evaluated against the legal standards implied by the Magnuson-Stevens Fishery Conservation and Management Act ("MSA"): preventing overfishing, avoiding becoming overfished, and producing maximum sustainable yield. The MSA's National Standard 1 (NSG, 2009) requires that "[c]onservation and management measures shall prevent overfishing while achieving, on a continuing basis, the optimum yield from each fishery." 16 U.S.C. § 1851(a)(1). The National Standard 1 Guidelines (50 C.F.R. § 600.310(f)(4)) specify that the probability of overfishing cannot exceed 50%, but should be lower based on the degree of scientific uncertainty in the estimate of the OFL. The MSA requires that overfished stocks, which are often defined as $Bcur/B_{MSY} < 50\%$, be rebuilt as fast as possible.

Performance was measured in terms of preventing overfishing, avoiding becoming overfished, and producing long-term yield in light of these management objectives. The probability of overfishing is recorded for each simulation by calculating the fraction of projected years in which $F > F_{MSY}$. This was averaged over multiple simulations to create a probability of overfishing metric (POF) that is the expected probability of overfishing in a projected year using a particular management method. We use B_{MSY} as a management reference point for overfished stock status. Similarly to the POF metric, the future stock biomass relative to B_{MSY} (B/B_{MSY}) was averaged over projected years and simulations to provide the expectation of stock status using a particular management method. Absolute yield of any projection is difficult to interpret because it depends on the specific conditions of each projection (i.e., starting depletion, future productivity, etc.). A standardized measure of yield was calculated by dividing the total projected yield for each simulation by the catch under F_{ref}, the constant F that maximizes catch over the projected time period with perfect knowledge of future recruitment deviations. In this way, yields are standardized as a percentage of an "upper bound." In some cases it is possible for a method to obtain relatively high total yields over the whole projection by depleting the stock (a "mining" strategy). The yield metric was therefore calculated based on the last five years of each projection (e.g., the yield from a method in projected years 26–30 divided by the yield of the F_{ref} strategy in projected years 26–30) since it is of more interest to identify methods that can achieve sustainable long-term yields. This was averaged over multiple simulations of each stock to provide the expected relative yield (herein referred to as 'Yield') of a management method.

The metrics POF, B/B_{MSY} and Yield relate to the central reference points for overfishing, overfished status and sustainable yield, but cannot be readily interpreted in terms of the average trajectory of biomass using a particular management method. To address this, we derive four additional metrics that relate to stock status in the final three years of the projections. The probability of biomass increasing, P_{inc}, is the fraction of projected simulations for which average biomass in the last three years of the projection is larger than average biomass for the last three years of the historical simulation. B_{end} is the mean biomass over the final three years of the projection divided by B_{MSY} averaged over simulations. The probability of ending below 50% B_{MSY}, $P_{<50}$ is the fraction of simulations for which the mean biomass of the last three projected years is below 50% B_{MSY}. Similarly, $P_{<10}$ is the fraction of simulations ending below 10% B_{MSY}.

Each performance metric was calculated for each simulation allowing performance to be averaged over various subsets of the simulations. For example, of the 10,000 simulations that were conducted for mackerel, approximately 1700 corresponded to stocks that were below 50% B_{MSY} levels at the end of the historical projection. The mean performance metrics were calculated for this subset of 1700 simulations to reveal how the expected performance when starting from low population levels. We used a similar approach to quantify the value of different sources of information (Section 2.8 below).

2.8. Quantifying value of information

We evaluated how long-term yield can be expected to vary with the uncertainty in each input. This was used to assess the value of various sources of information for each method. To do this we took each input variable/parameter in turn and subdivided the simulations into ten equally sized blocks relating to the 10th percentiles of the sampled input. For example, less than the 10th percentile of sampled bias in depletion, greater than or equal to the 10th percentile but less than the 20th percentile of bias in depletion, and so on. Since the 10,000 simulations of each stock type were subdivided according to percentiles in the input parameters, these subsets were approximately equal in size at around 1000 simulations. The mean relative yield for each of the ten subdivisions was calculated for each method. The standard deviation of these relative yield scores can be interpreted as the marginal effect of an input variable on expected yield. These results are unit-less because they are standardized according to the level of simulated uncertainty for each of the input parameters/variables.

3. Results

3.1. Performance

The general results statements below refer to the mackerel, snapper, porgy, sole and rockfish simulations. The results for butterfish are discussed in Section 3.2 because the simulations for butterfish behaved very differently from those for the other stocks.

Table 3
Summary of the bias/error parameters and related distributions that control the accuracy and imprecision of knowledge of the simulated system that is subsequently used by the data-limited methods and harvest control rules. The log-normal distribution described in the table below (~LN(μ,CV)) is the exponent of the normal distribution with mean and standard deviation ($sd = CV \times$ mean) parameters: $N(-0.5\log(1+sd^2/\mu^2), \sqrt{\log(1+sd^2/\mu^2)})$.

Variable	Symbol	Related functions	All stocks
The coefficient of variation of the log-normally distributed bias in natural mortality rate M	CV_M	$M_{obs} = M \times \mu_M$ $\mu_M \sim dlnorm(\mu = 1, CV_M)$	0.5
The coefficient of variation of the log-normally distributed bias in von Bertalanffy growth rate parameter K	CV_K	$K_{obs} = K \times \mu_K$ $\mu_K \sim dlnorm(\mu = 1, CV_K)$	0.2
The coefficient of variation of the log-normally distributed bias in length at first capture, L_c	CV_{Lc}	$Lc_{obs} = Lc \times \mu_{Lc}$ $\mu_{Lc} \sim dlnorm(\mu = 1, CV_{Lc})$	0.5
The coefficient of variation of the log-normally distributed bias in biomass at maximum sustainable yield relative to unfished B_{peak} (B_{MSY}/B_0)	CV_{Bpeak}	$B_{peak_{obs}} = B_{peak} \times \mu_{Bpeak}$ $\mu_{B_{peak}} \sim dlnorm\ (\mu = 1, CV_{B_{peak}})$	0.2
The coefficient of variation of the log-normally distributed bias in the ratio of maximum sustainable fishing mortality rate to natural mortality rate c	CV_c	$c_{obs} = c \times \mu_c$ $\mu_c \sim dlnorm(\mu = 1, CV_c)$	0.2
The coefficient of variation of the log-normally distributed bias in the age at first maturity Am	CV_{Am}	$Am_{obs} = Am \times \mu_{Am}$ $\mu_{Am} \sim dlnorm(\mu = 1, CV_{Am})$	0.2
The coefficient of variation of the log-normally distributed bias in the intrinsic rate of increase parameter r	CV_r	$r_{obs} = r \times \mu_r$ $\mu_r \sim dlnorm(\mu = 1, CV_r)$	0.5
The coefficient of variation of the log-normally distributed bias in the current level of stock depletion D ($Bcur/B_0$)	CV_D	$D_{obs} = D \times j_D$ $j_D \sim dlnorm(\mu_D, \sigma_D)$ $\mu_D \sim dlnorm(\mu = 1, CV_D)$	1
The maximum coefficient of variation for log-normal error around bias in current stock depletion μ_D for projected years	σ_{maxD}	$D_{obs} = D \times j_D$ $j_D \sim dlnorm(\mu_D, \sigma_D)$ $\sigma_D \sim U(0, \sigma_{maxD})$	2
The coefficient of variation of the log-normally distributed bias in the current stock level $Bcur$	CV_{Bcur}	$Bcur_{obs} = Bcur \times j_{Bcur}$ $j_{Bcur} \sim dlnorm(\mu_{Bcur}, \sigma_{Bcur})$ $\mu_{Bcur} \sim dlnorm(\mu = 1, CV_{Bcur})$	1
The maximum coefficient of variation for log-normal error around bias μ_{Bcur} for projected years	$\sigma_{maxBcur}$	$Bcur_{obs} = Bcur \times j_{Bcur}$ $j_{Bcur} \sim dlnorm(\mu_{Bcur}, \sigma_{Bcur})$ $\sigma_{Bcur} \sim U(0, \sigma_{maxBcur})$	2

It was instructive to separate the simulations according to the depletion at the start of the projection. Four categories were chosen relating to projections starting (1) below 50% of B_{MSY}, (2) between 50% and 100% B_{MSY}, (3) between 100% and 150% of B_{MSY} and (4) above 150% B_{MSY}. The largest discrepancies in performance were found among the first three categories and for the benefit of brevity the table for projections starting above 150% B_{MSY} is included in the Appendix (Table App.C.1)

3.1.1. Catch-based methods

Methods that set the ABC to average historical catches or a percentile of recent catch (M1–M3) led to the worst performance of the methods tested by a large margin. When starting below 50% B_{MSY}, the probability of overfishing was high – typically above 80% ("P_{OF}", Table 4). While some catch-based methods performed better at moderate levels of depletion (above 50% of B_{MSY}) particularly in regard to yield, they still led to relatively high probabilities of overfishing-in most cases exceeding 60% of the simulations (Tables 5 and 6). These static catch-based methods failed to rebuild stocks initially below 50% B_{MSY} to above 50% B_{MSY} in the majority of simulations (between 60% and 95%; on most occasions the failure rate was over 85% ("$P_{<10}$", Table App.C.2). The static catch based methods could lead to very high probabilities of dropping below 10% of B_{MSY} (generally 40–60%) when applied to stocks starting below B_{MSY} (Table App.C.3). Relative to other methods, $P_{<10}$ remained high even when stock levels were above B_{MSY} (between 12% and 26% for M1–M3 compared with less than 2% for M4–M9, Table App.C.4). Methods M1–M3 also led to amongst the lowest yields in simulations starting below B_{MSY} (Figs. 4 and 5). The performance of these methods was poor for all stocks except butterfish (see Section 3.2), and was not as strongly related to life-history type compared to the other methods. Methods M1–M3 performed worse than the "status quo" current catch and effort scenarios (R3–R4) in several instances. This was particularly the case for method M3 (ABC set at the third highest historical catch) which drove 19 out of 20 stocks that were already below 50% of B_{MSY} at the start of the projection to below 10% of B_{MSY} by the end of the projection (Table 4). This was only somewhat reduced to 7 out of 10 stocks in those simulations starting between 50% and 100% of B_{MSY} (Table 5).

The dynamic catch-based methods A1 and A2 led to intermediate performance at low stock sizes (i.e., less than 50% B_{MSY}) in terms of the probability of overfishing and yield relative to the other methods. At moderate stock sizes they performed much better, leading to reasonably high yields (approximately 50–80% of those corresponding to F_{ref}), with moderate probabilities of overfishing (approximately 30–40%) (Tables 4 and 5, Figs. 5 and App.C.6). Methods A1 and A2 reduced catches by multiplying historical mean catch by 50% when the stock declines below 20% of unfished levels. This does not appear to be sufficiently responsive to prevent these methods from frequently depleting the stock below the overfished threshold of 50% B_{MSY}, even in simulations that start above 50% B_{MSY} (Tables App.C.3 and App.C.4).

3.1.2. Depletion-based methods

The static implementation of DB-SRA that assumes that stock depletion is, on average, 40% of unfished levels (equivalent to ~100% of B_{MSY}) performed well when this assumption was reasonably close to actual depletion (e.g., 50–150% of B_{MSY}, Tables 5 and 6). At these stock levels, the probability of overfishing, projected stock status (B/B_{MSY}) and yield were among the best of any method. The probabilities of stocks falling below 50% B_{MSY} were also relatively small, with the majority of cases exhibiting an increasing biomass trend on average ("P_{inc}", Table App.C.3). However, these methods prescribed OFLs that were too high and stocks suffered from high probabilities of overfishing, depletion and consequently reduced yields when starting biomass was much below that assumed (Table 4). Since the PFMC DB-SRA methods do not introduce feedback between stock status and the OFL recommendation, these methods suffer from a similar, but less pronounced phenomenon as the average catch methods. DB-SRA performed relatively poorly,

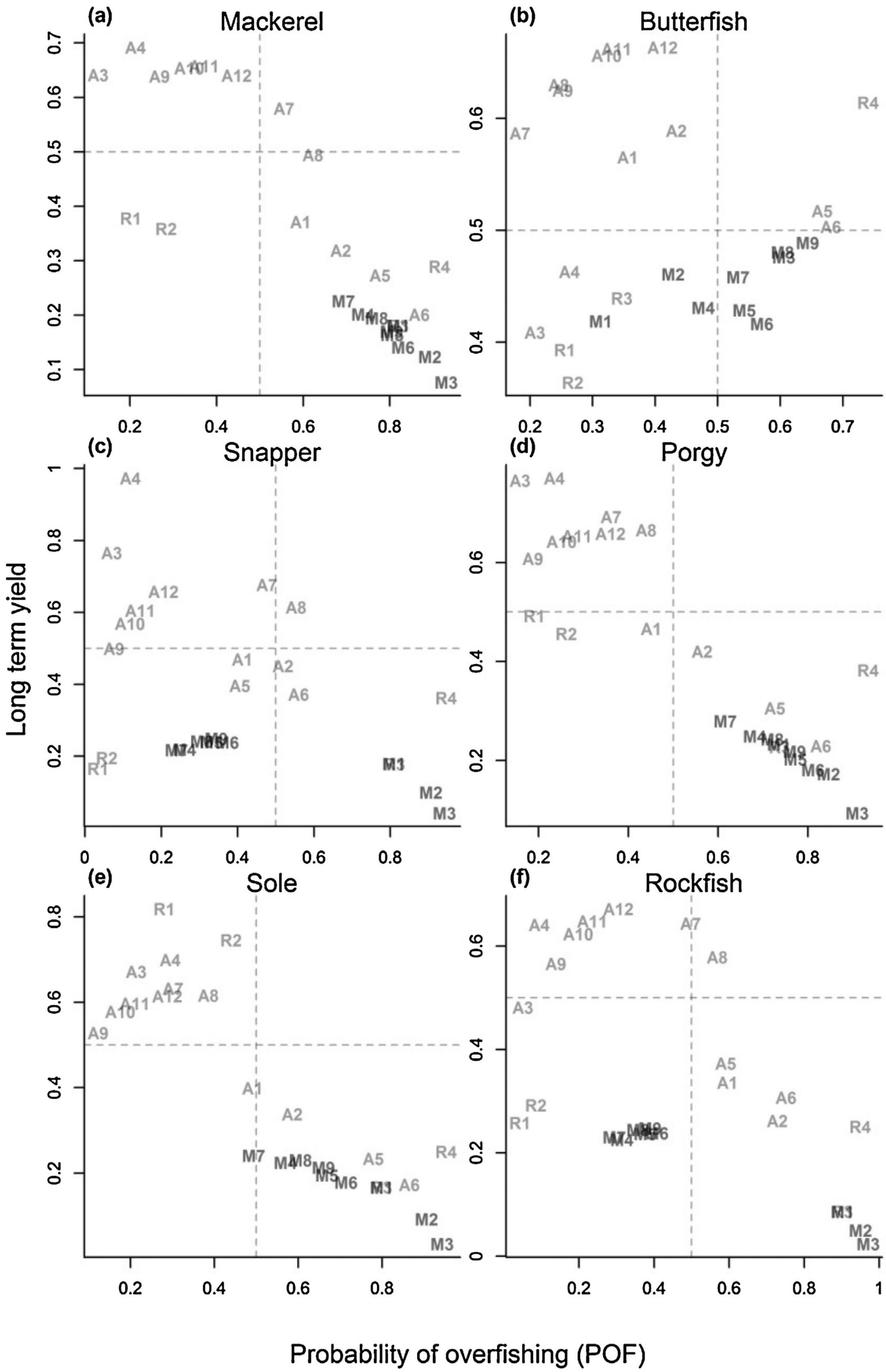

Fig. 4. The trade-off between of long term yield (yield over last 5 projected years divided by that of the F_{ref} strategy) and the probability of overfishing (fraction of projected years for which fishing mortality rate exceeded F_{MSY}) for projections starting below 50% B_{MSY}.

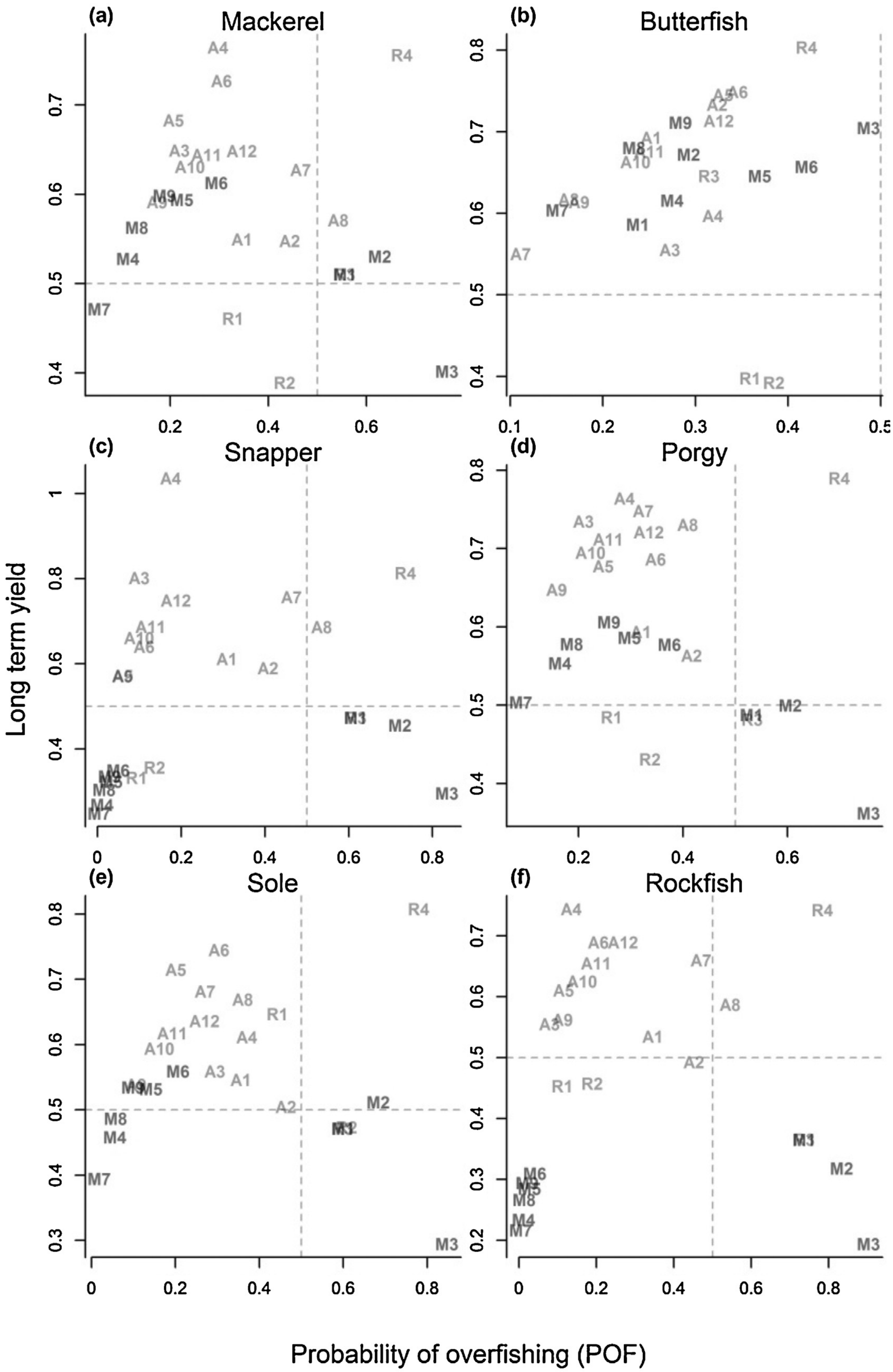

Fig. 5. The trade-off between of long term yield (yield over last 5 projected years divided by that of the F_{ref} strategy) and the probability of overfishing (fraction of projected years for which fishing mortality rate exceeded F_{MSY}) for projections starting between 50% and 100% B_{MSY}.

Table 4
Overfishing, stock status and yield performance metrics for simulations starting below 50% of *BMSY*. All of the numbers represent a percentage. The probability of overfishing (P_{OF}) is the fraction of years (across all simulations and all of their projection years) for which fishing mortality rate exceeds F_{MSY}. 'B/B_{MSY}' is the mean biomass (across all simulations and all of their projection years) divided by biomass at maximum sustainable yield. 'Yield' is the mean relative yield over the last five years of the projection (the yield of a simulation over the last five years of the projection divided by that of the F_{ref} policy). Dark gray shading reflects poor scores (P_{OF} greater than 50%, B/B_{MSY} less than 50%, yield less than 25%). Light gray shading reflects intermediate scores (P_{OF} greater than 25%, B/B_{MSY} less than 100%, yield less than 50%).

Type	Code	Name	Mackerel			Butterfish			Snapper			Porgy			Sole			Rockfish		
			P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield
Catch-Based (Static)	M1	Median Catch - 3 Years	82	22	18	31	103	42	81	29	18	74	39	23	80	31	17	90	14	9
	M2	Median Catch - 10 Years	89	14	12	43	88	46	91	16	10	85	26	17	91	17	9	95	8	5
	M3	3rd Highest Catch	93	10	8	61	67	48	94	9	4	91	16	9	94	9	3	97	5	2
Depletion-Based (Static)	M4	DB-SRA (Depletion Fixed @ 40%B0) - 69.4% scalar	74	32	20	48	78	43	26	98	22	68	47	25	57	63	22	31	69	23
	M5	DB-SRA (Depletion Fixed @ 40%B0) - 83.4% scalar	81	25	16	54	71	43	33	88	24	77	35	20	67	49	20	38	63	24
	M6	DB-SRA (Depletion Fixed @ 40%B0) - 91.3% scalar	83	22	14	57	67	42	37	83	24	81	30	18	71	42	18	41	60	24
	M7	DCAC (Depletion Fixed @ 40%B0) - 69.4% scalar	69	38	23	53	75	46	24	102	22	62	55	28	49	73	24	29	71	23
	M8	DCAC (Depletion Fixed @ 40%B0) - 83.4% scalar	77	29	19	60	66	48	31	92	24	72	42	24	61	58	23	36	65	24
	M9	DCAC (Fixed Depletion @ 40%B0) - 91.3% scalar	80	26	17	64	61	49	34	86	25	77	36	22	66	50	21	39	62	25
Catch-Based (Dynamic)	A1	Depletion Adjusted Catch Scalar - 75% scalar	59	39	37	36	92	57	41	67	47	45	61	47	49	64	40	60	36	34
	A2	Depletion Adjusted Catch Scalar - 100% scalar	69	32	32	43	83	59	52	55	45	56	50	42	59	52	34	73	27	26
Depletion-Based (Dynamic)	A3	DB-SRA (Depletion Adjusted) - 25% P*	13	67	64	21	105	41	7	122	77	16	90	77	21	99	67	5	85	48
	A4	DB-SRA (Depletion Adjusted) - 50% P*	21	60	69	26	98	46	12	110	97	24	81	77	29	88	70	9	75	64
	A5	DCAC (Depletion Adjusted) - 25% P*	78	26	27	67	58	52	41	74	40	73	40	31	78	34	23	59	42	37
	A6	DCAC (Depletion Adjusted) - 50% P*	87	18	20	68	57	50	56	56	37	83	29	23	86	23	17	75	30	31
Abundance-Based (Dynamic)	A7	Life History Analysis - 75% scalar	56	38	58	18	110	59	48	59	68	36	74	69	30	89	63	50	43	64
	A8	Life History Analysis - 100% scalar	62	31	49	25	102	63	55	49	61	44	64	67	39	76	62	57	36	58
	A9	FMSY/M (Low) - 75% scalar	27	64	64	25	102	63	8	120	50	19	96	61	12	117	53	14	77	57
	A10	FMSY/M (Low) - 100% scalar	34	58	65	32	94	66	12	112	57	25	87	64	18	107	58	20	71	62
	A11	FMSY/M (Hi) - 75% scalar	37	55	66	34	92	66	14	107	61	29	83	65	21	102	60	24	68	65
	A12	FMSY/M (Hi) - 100% scalar	45	48	64	41	84	66	21	97	66	36	73	66	29	91	61	31	61	67
Stock Assessment	R1	Delay-Difference - 75% scalar	20	69	38	26	100	39	3	142	17	19	100	49	28	99	82	4	92	26
	R2	Delay-Difference - 100% scalar	28	63	36	27	97	36	6	138	20	26	92	46	44	81	75	8	88	29
Status Quo (Static)	R3	Current Catch	82	22	18	35	99	44	81	29	18	74	39	23	80	31	17	90	14	9
	R4	Current Effort	91	16	29	74	54	61	95	19	36	93	23	38	95	17	25	95	14	25

leading to a low probability of recovery from biomass below 50% B_{MSY} regardless of the ABC control rule (scalar multipliers between 69% and 91%) ('$P_{<50}$', Table App.C.2). This was particularly the case for the mackerel and porgy stocks, where the probability of projections ending below half of B_{MSY} was between 50% and 80% when starting below half of B_{MSY} (Table App.C.2).

DB-SRA and DCAC performed somewhat better for long-lived life history types such as snapper and rockfish compared with other methods. This result is a product of the greater "windfall" biomass of older age classes, that is deliberately accounted for by DCAC and is approximated by the delay-difference stock dynamics of DB-SRA.

Performance of DB-SRA is improved for stocks starting below 50% B_{MSY} when stock depletion is updated dynamically (methods A3 and A4), leading to less than 20% probability of overfishing on average. Methods A3 and A4 lead to increasing biomass from low levels in over 70% of simulations regardless of life-history type (Table App.C.2). Rebuilding performance was considerably worse for the mackerel, and while these methods managed better performance than any other method, between 36% and 42% of stocks did not rebuild above 50% B_{MSY}. The performance of methods A3 and A4 became much worse at higher stock levels in comparison to the other data-limited methods largely due to the high level

Table 5
As for Table 4, except the simulations start between 50% and 100% of B_{MSY}.

Type	Code	Name	Mackerel			Butterfish			Snapper			Porgy			Sole			Rockfish		
			P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield
Catch-Based (Static)	M1	Median Catch - 3 Years	56	76	51	24	126	59	62	72	47	53	84	49	60	76	47	74	54	37
	M2	Median Catch - 10 Years	63	68	53	29	119	67	72	60	46	61	75	50	68	67	51	83	43	32
	M3	3rd Highest Catch	76	51	40	49	97	70	83	43	30	76	54	36	85	45	29	90	31	19
Depletion-Based (Static)	M4	DB-SRA (Depletion Fixed @ 40%B0) - 69.4% scalar	11	128	53	27	122	62	1	174	27	16	132	55	6	152	46	1	150	23
	M5	DB-SRA (Depletion Fixed @ 40%B0) - 83.4% scalar	22	115	59	37	111	65	3	167	32	30	115	59	14	137	53	3	145	28
	M6	DB-SRA (Depletion Fixed @ 40%B0) - 91.3% scalar	29	107	61	42	105	66	5	162	35	37	105	58	21	128	56	4	143	31
	M7	DCAC (Depletion Fixed @ 40%B0) - 69.4% scalar	6	135	47	15	135	60	0	177	25	9	143	50	2	161	39	0	152	22
	M8	DCAC (Depletion Fixed @ 40%B0) - 83.4% scalar	13	125	56	23	124	68	2	170	30	19	128	58	6	149	49	1	148	27
	M9	DCAC (Fixed Depletion @ 40%B0) - 91.3% scalar	19	118	60	28	118	71	3	166	34	26	119	61	10	142	53	2	145	29
Catch-Based (Dynamic)	A1	Depletion Adjusted Catch Scalar - 75% scalar	35	92	55	25	125	69	31	106	61	32	106	59	36	102	55	35	82	53
	A2	Depletion Adjusted Catch Scalar - 100% scalar	44	78	55	32	115	73	41	89	59	42	91	56	46	84	50	45	68	49
Depletion-Based (Dynamic)	A3	DB-SRA (Depletion Adjusted) - 25% P*	22	108	65	27	124	56	10	155	80	21	122	73	29	117	56	8	134	56
	A4	DB-SRA (Depletion Adjusted) - 50% P*	30	98	76	32	117	60	18	138	104	29	111	76	37	105	61	14	114	74
	A5	DCAC (Depletion Adjusted) - 25% P*	21	110	68	33	113	75	6	146	57	25	117	68	20	118	72	12	117	61
	A6	DCAC (Depletion Adjusted) - 50% P*	30	100	73	35	111	75	11	133	64	35	104	69	30	107	75	21	105	69
Abundance-Based (Dynamic)	A7	Life History Analysis - 75% scalar	47	80	63	11	143	55	46	84	76	32	111	75	27	121	68	47	73	66
	A8	Life History Analysis - 100% scalar	54	67	57	16	135	62	54	70	69	41	97	73	36	106	67	55	61	59
	A9	FMSY/M (Low) - 75% scalar	17	128	59	17	134	61	6	165	57	16	141	65	11	156	54	11	131	56
	A10	FMSY/M (Low) - 100% scalar	24	117	63	24	125	66	10	155	66	22	129	69	16	144	59	16	121	63
	A11	FMSY/M (Hi) - 75% scalar	27	111	64	25	123	68	13	149	69	26	123	71	19	137	62	20	116	66
	A12	FMSY/M (Hi) - 100% scalar	35	99	65	33	114	71	19	136	75	34	109	72	27	123	64	27	105	69
Stock Assessment	R1	Delay-Difference - 75% scalar	33	104	46	36	115	40	9	166	33	26	127	49	44	98	65	11	131	45
	R2	Delay-Difference - 100% scalar	43	91	39	38	111	39	14	158	36	34	114	43	61	77	47	19	121	46
Status Quo (Static)	R3	Current Catch	56	76	51	31	118	65	62	72	47	53	84	48	60	76	47	74	54	37
	R4	Current Effort	67	70	76	42	101	80	74	68	81	70	72	79	78	69	81	78	61	74

Table 6
As for Table 4, except the simulations start between 100% and 150% of B_{MSY}.

Type	Code	Name	Mackerel			Butterfish			Snapper			Porgy			Sole			Rockfish		
			P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield	P_{OF}	B/B_{MSY}	Yield
Catch-Based (Static)	M1	Median Catch - 3 Years	26	130	65	26	129	61	34	122	77	29	130	63	26	130	70	43	109	67
	M2	Median Catch - 10 Years	25	128	76	27	127	69	34	116	86	29	127	73	22	128	85	47	103	76
	M3	3rd Highest Catch	41	109	72	46	104	72	52	96	77	45	104	66	44	104	78	62	85	63
Depletion-Based (Static)	M4	DB-SRA (Depletion Fixed @ 40%B0) - 69.4% scalar	1	176	43	22	135	64	0	209	24	2	178	54	0	190	41	0	193	17
	M5	DB-SRA (Depletion Fixed @ 40%B0) - 83.4% scalar	2	168	53	31	124	67	0	204	29	5	166	65	0	180	51	0	190	21
	M6	DB-SRA (Depletion Fixed @ 40%B0) - 91.3% scalar	4	163	58	36	118	68	0	201	32	9	159	70	1	174	57	0	188	23
	M7	DCAC (Depletion Fixed @ 40%B0) - 69.4% scalar	0	181	37	12	146	61	0	211	22	0	186	46	0	196	35	0	195	16
	M8	DCAC (Depletion Fixed @ 40%B0) - 83.4% scalar	1	174	46	19	137	69	0	206	27	2	176	57	0	187	44	0	192	20
	M9	DCAC (Fixed Depletion @ 40%B0) - 91.3% scalar	1	170	51	23	131	73	0	204	30	3	169	63	0	182	49	0	190	22
Catch-Based (Dynamic)	A1	Depletion Adjusted Catch Scalar - 75% scalar	28	128	61	21	134	68	24	139	71	26	136	64	28	135	64	27	116	63
	A2	Depletion Adjusted Catch Scalar - 100% scalar	36	110	60	30	123	74	30	123	73	34	119	62	35	116	62	38	99	59
Depletion-Based (Dynamic)	A3	DB-SRA (Depletion Adjusted) - 25% P*	26	120	54	25	132	56	11	174	66	22	133	63	33	126	53	9	159	55
	A4	DB-SRA (Depletion Adjusted) - 50% P*	35	107	58	29	125	62	18	152	89	30	121	61	40	115	59	14	132	73
	A5	DCAC (Depletion Adjusted) - 25% P*	3	158	65	27	126	77	1	182	57	5	163	70	1	162	69	3	162	55
	A6	DCAC (Depletion Adjusted) - 50% P*	4	152	71	28	124	77	1	174	65	7	155	76	2	156	74	4	154	64
Abundance-Based (Dynamic)	A7	Life History Analysis - 75% scalar	46	97	59	9	151	52	48	97	81	30	131	72	28	134	64	47	91	69
	A8	Life History Analysis - 100% scalar	53	83	53	15	143	59	55	83	76	38	115	71	36	117	62	55	76	61
	A9	FMSY/M (Low) - 75% scalar	15	155	57	17	141	59	8	185	65	15	163	62	11	172	53	11	160	56
	A10	FMSY/M (Low) - 100% scalar	22	142	61	23	132	65	12	174	73	21	150	66	17	159	57	16	149	63
	A11	FMSY/M (Hi) - 75% scalar	25	136	62	25	130	66	15	167	76	24	143	67	20	152	59	19	142	67
	A12	FMSY/M (Hi) - 100% scalar	33	122	63	32	121	71	20	154	82	31	128	68	28	136	60	26	129	71
Stock Assessment	R1	Delay-Difference - 75% scalar	32	121	39	37	118	44	13	169	36	24	140	42	38	110	50	19	145	37
	R2	Delay-Difference - 100% scalar	40	107	38	41	114	41	15	158	36	30	128	43	49	91	47	27	129	39
Status Quo (Static)	R3	Current Catch	26	130	65	37	118	69	34	122	78	29	130	63	26	130	70	43	109	67
	R4	Current Effort	22	130	81	33	117	75	27	122	96	27	127	86	21	128	89	34	118	85

of uncertainty regarding depletion. This led to many occasions of inflated OFL recommendations and therefore stock declines when depletion was assumed to be too high.

MacCall (2009) notes that DCAC is "not directly suitable for specifying catches in a stock-rebuilding program." This is because it returns an estimate of an MSY proxy ("sustainable catch" which is particular to a productive stock size) and not an estimate of the OFL (which changes with depletion level). It is not surprising, therefore, that DCAC performs relatively poorly at low starting levels (below 50% B_{MSY}, Tables 4 and App.C.2) regardless of whether or not depletion is dynamically updated. The static DCAC provides yields and probabilities of overfishing comparable to the best performing methods at intermediate levels of depletion when the stock is closer to MSY levels (Tables 5 and App.C.3). As is the case with the dynamic update in DB-SRA, the high level of uncertainty in current depletion that was simulated led to relatively poor performance at moderate depletion levels (50–150% depletion).

3.1.3. Abundance-based methods

The method of Beddington and Kirkwood (2005; A7 and A8) that estimates F_{MSY} based on size at first recapture and age at 50% maturity appears to offer intermediate performance overall. Often providing relatively high yields, the method tended to overfish more than the best performing approaches (see trade-off plots, Figs. 4 and 5). The propensity to overfish was not reduced substantially for simulations at intermediate depletion levels (between 50% and 150% B_{MSY}, Table 5) unlike other methods that make use of current information regarding stock level. Methods A7 and A8 appeared to perform particularly poorly for mackerel, snapper and rockfish in terms of the probability of ending below the 50% B_{MSY} threshold, even when biomass is initially above this threshold (Table App.C.3).

In general, F_{MSY}/M methods A9–A12 were among the best performers regardless of life-history and initial depletion level. Along with methods A3 and A4, methods A9 and A10 were unique in their ability to rebuild stocks in a substantial number of simulations while achieving relatively high yields. Overall, F_{MSY}/M method A9 performed somewhat worse than DB-SRA method A3 at low stock sizes, with the exception of higher yields for rockfish and a lower probability of overfishing for porgy. At intermediate stock depletion levels, method A9 compared favorably with method A3 and led to similar yields with lower probabilities of overfishing for all stocks, with the exception of rockfish (Tables 5 and 6).

3.1.4. Reference case methods

The delay-difference assessment had mixed performance despite having unbiased information regarding vulnerability at age, median age at maturity, growth rate and natural mortality rate. The probability of overfishing was generally low, but yields were unremarkable compared with the other methods, particularly when starting from moderate stock sizes (i.e., between 50% and 150% B_{MSY}). Projected biomass increased from low stock sizes in most cases, but the probability of remaining below the overfished threshold was still high for mackerel. As expected, the current catch and effort methods performed poorly due to their lack of feedback between the OFL and stock depletion. It follows that simulations that did not lead to stock collapses coincided with those for which the final historical fishing mortality rate happened to be sustainable.

3.1.5. Trade-offs among ABC control rules

ABC control rules, incorporating varying downward adjustments, were considered for each OFL-setting method. As expected, the reduction in the ABC led to a reduced probability of overfishing and increases in expected population size (e.g., B/B_{MSY}, Figs. 6, 7 and App.C.4). The pattern in long-term yield was less clear, with the largest downward adjustments leading to relatively small reductions in yield. For example: a 75% scalar applied to method A9 led to a 27% probability of overfishing and 64% yield for mackerel starting below 50% B_{MSY} compared with the unmodified rule (method A10) that achieved a 34% probability of overfishing and 65% yield. In methods where the probability of overfishing is generally higher, greater downward adjustment increases the long term expectation of yield. For example, a 75% scalar for methods A7 and A8 leads a lower probability of overfishing, higher expected biomass and higher long-term yield for the snapper stock.

3.1.6. Inter-method performance trade-offs

There is a relatively well-defined inverse relationship between the expected probability of overfishing and expected stock status

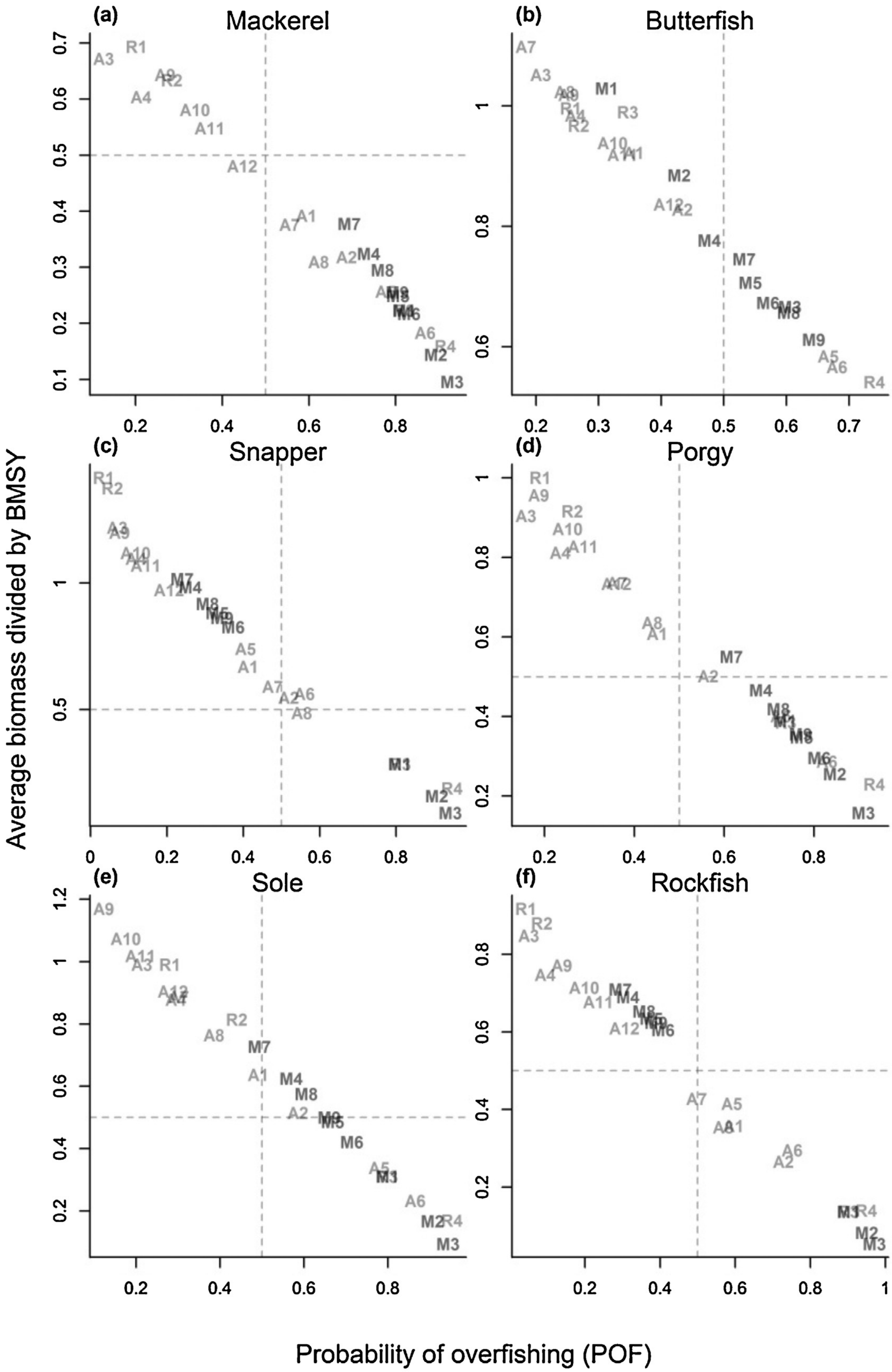

Fig. 6. The trade-off between average stock depletion (projected biomass divided by B_{MSY}) and the probability of overfishing (fraction of projected years for which fishing mortality rate exceeded F_{MSY}) for projections starting below 50% B_{MSY}.

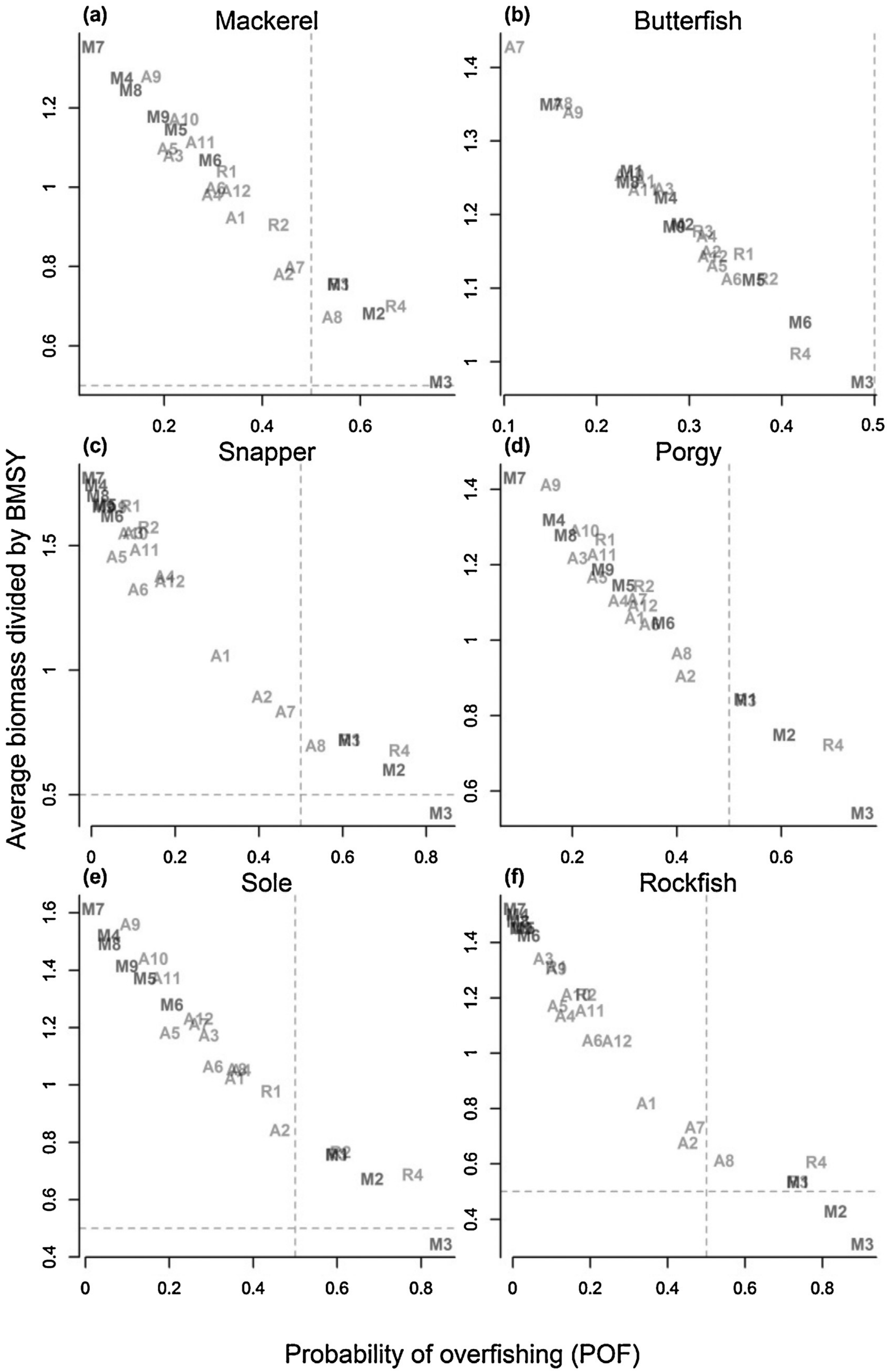

Fig. 7. The trade-off between average stock depletion (projected biomass divided by B_{MSY}) and the probability of overfishing (fraction of projected years for which fishing mortality rate exceeded F_{MSY}) for projections starting between 50% and 100% B_{MSY}.

(B/B_{MSY}) across all methods (Figs. 6 and 7). The ranking of methods in terms of these criteria is relatively clear. It is not surprising that a method that provides the lowest propensity to overfish leads to the highest abundance levels. The relationship between the probability of overfishing and long-term yield is less clear (Figs. 4, 5 and App.C.6). When simulations start from low stock sizes, the methods are either scattered in this trade-off space (snapper, butterfish and rockfish stocks) or show a weak negative relationship, where higher yields are achieved at lower probabilities of overfishing (mackerel, porgy and sole stocks). This is intuitive since stock recovery to productive biomass levels increases longer term yields. This pattern in this trade-off becomes weakly positive from intermediate starting depletion (50–150% B_{MSY}). The scatter in the trade-off plots indicates opportunities to select methods that can achieve both lower probabilities of overfishing and higher yields than other methods. As identified from Tables 4–6, methods A3, A4, A9 and A10 lead to high yields and low probabilities of overfishing across several starting depletions.

3.2. Performance for butterfish

Butterfish proved to be the most challenging test of the data-limited methods. We include the results of DCAC and DB-SRA even though these methods are not appropriate for stocks such as butterfish that have natural mortality rates higher than the guideline of 0.2 yr^{-1} (MacCall, 2009; Dick and MacCall, 2011). The relative performance of the methods for butterfish was unique among the species considered. In general, all methods led to moderate probabilities of overfishing without commensurate stock depletion (Table 4). Similarly, expected yield for butterfish was relatively high compared with other stocks even when applying the worst performing methods. Methods that led to the likely collapse of other stocks (e.g., average catch methods M1–M3) achieved a relatively high rate of rebuilding for butterfish when projections were started from below 50% B_{MSY} (Table App.C.2). This result emphasizes the larger role of temporal changes in stock productivity in determining abundance for species such as butterfish, which are short-lived and exhibit highly variable recruitment.

3.3. Value of different sources of information for each data-limited method

Current abundance, historical fishing effort, and stock depletion have the highest information content; only those methods that incorporated these sources of data had good performance across all depletion levels (e.g., could recover stocks from low stock sizes and did not lead to declines below 50% B_{MSY} in a high fraction of simulations). This additional value can be expressed in either the difference in the expected long-term yield or the probability of overfishing. Butterfish aside, benefits in yield and the probability of overfishing were very large at very low stock sizes (<50% B_{MSY}), but negligible or non-existent at more intermediate stock sizes (50–150% B_{MSY}). For example, methods A3, A4, A9–A12 lead to expected probabilities of overfishing that are between 70% and 35% lower than the other methods when biomass is initially below 50% B_{MSY}, while offering expected yields that are between 2 and 6 times higher. Overfishing may occur with higher frequency than other methods at moderate stock sizes, but yields generally remained between 10% and 30% higher for these dynamic approaches.

The yield and probability of overfishing varied more strongly with consistent bias in depletion and current biomass, indicating that accuracy in these inputs is a critical determinant of the performance of the associated methods (Tables 7 and App.C.6). This is particularly important as the methods that make use of these inputs are those that appear to perform best (e.g., methods A3 and A9). This sensitivity is to be expected since these inputs provide the dynamic link to changes in stock size, which is the central reason these methods perform well. Since M is a factor in the calculation of the OFL, it follows that the F_{MSY}/M methods are sensitive to uncertainty in this input. It may not be immediately clear why yields should vary to a larger extent across the bias in current biomass in comparison to M. The simple explanation is that twice the level of potential bias was prescribed for current biomass (a CV of 1 compared with 0.5 for M). While bias in depletion and current biomass led to large changes in yield for some methods, the precision of these inputs was much less important.

There is evidence that methods offering intermediate performance may be somewhat less sensitive to inputs. For example, the DACS methods (A1 and A2) appeared relatively robust to bias in depletion although they did not perform well at low stock levels. This result points to a possible problem in the interpretation of the performance metrics which aggregate across factors, that they do not convey the extent to which the performance of the methods degrades under misspecification of inputs. On average, bias in inputs was sampled with a mean of 1 (unbiased on average). It follows that it may be possible for a method to lead to a mean probability of overfishing of 20% but this performance is only representative of a small set of unbiased simulations. Examining the sensitivity of the methods A3, A4, and A9–A12 reveals this problem. This phenomenon is illustrated in Figs. App.D1–D4 where the slope in expected probability of overfishing is very steep at zero bias (a value of 1) in depletion and current biomass, respectively. Methods A3 and A4 that allow for dynamic update of depletion also exhibit considerably more sensitivity to M for snapper and rockfish.

3.4. Sensitivity of performance to population and fishing dynamics

Mackerel and porgy were the most difficult to rebuild. Snapper has the highest probability of increasing stock trends (P_{inc}) and of ending above the rebuilding threshold for all methods, with the notable exception of the average catch methods (Table 4).

There were relatively few interactions between the performance of methods and life-history type; while the absolute performance of most methods changed markedly among stocks, within each stock the ranking of methods was consistent. There are a few notable exceptions. For example, the average catch methods (M1–M3) have similarly poor absolute performance across the life history types, with the exception of butterfish. Methods M4–M9 also led to relatively low yields for the more long-lived stocks, such as snapper and rockfish when projections started at intermediate biomass levels (Tables 5 and 6). Mackerel and sole showed unexpectedly a high likelihood of dropping below 50% B_{MSY} for intermediate initial depletion levels for methods A3 and A4. Methods A7 and A8 also led to markedly better performance for the butterfish.

The most important characteristics determining the probability of overfishing for those methods that do not include dynamic updates in depletion or current biomass are the steepness of the Beverton–Holt stock recruitment curve and the annual increase in fishing efficiency (Table App.C.7). The success of these methods coincides with productive stocks (high steepness) subject to low historical fishing mortality rates due to their lack of feedback between the ABC and stock status. This difference is demonstrated by dynamic abundance-based methods A9–A10, for which probability of overfishing is much less affected by variability in the simulated population and fishery parameters.

Overall, the performance of methods was unaffected by different input values for inter-annual recruitment variability ("Proc. Err"), inter-annual variability in fishing effort ("Eff. CV"), spatial targeting ("Targeting"), the von Bertalanffy growth coefficient ("Von B K"), stock viscosity and the degree of overlap among vulnerability

Table 7

The sensitivity in the yield metric to imperfect knowledge. The variables are *CV* in observation error (Obs err), bias in depletion (Dep bias), *CV* in depletion error (Dep *CV*), bias in the ratio of F_{MSY}/M (FMSY/M), bias in the ratio of B_{MSY} relative to unfished (BMSY/B0), bias in natural mortality rate (M), bias in the age at 50% maturity (50% Mat), bias in the current biomass (B bias), *CV* of error in current biomass (B CV), bias in the von Bertalanffy growth coefficient K (Von B K) and bias in the length at first recapture (L 1st Cap). All numbers are the standard deviation in probability of overfishing across ten divisions of each variable (10 percentile ranges). Sensitivity scores over 10 are shaded light gray, scores over 20 are shaded dark gray.

Type	Code	Methods	Mackerel											Butterfish										
			Obs err	Dep bias	Dep CV	FMSY/M	BMSY/B0	M	50% Mat	B bias	B CV	Von B K	L 1st Cap	Obs err	Dep bias	Dep CV	FMSY/M	BMSY/B0	M	50% Mat	B bias	B CV	Von B K	L 1st Cap
Catch-Based (Static)	M1	Median Catch - 3 Years	1											1										
	M2	Median Catch - 10 Years	1											2										
	M3	3rd Highest Catch	2											3										
Depletion-Based (Static)	M4	DB-SRA (Depletion Fixed @ 40%B0) - 69.4% scalar	1			1	1	8	1					3			2	1	1	2				
	M5	DB-SRA (Depletion Fixed @ 40%B0) - 83.4% scalar	1			1	1	8	1					2			2	1	3	2				
	M6	DB-SRA (Depletion Fixed @ 40%B0) - 91.3% scalar	1			1	1	8	1					2			2	2	3	2				
	M7	DCAC (Depletion Fixed @ 40%B0) - 69.4% scalar	1			0	1	7	1					2			2	1	1	4				
	M8	DCAC (Depletion Fixed @ 40%B0) - 83.4% scalar	0			1	1	7	1					2			1	1	2	3				
	M9	DCAC (Fixed Depletion @ 40%B0) - 91.3% scalar	1			1	1	7	1					2			1	1	3	3				
Catch-Based (Dynamic)	A1	Depletion Adjusted Catch Scalar - 75% scalar	1	9	4									3	9	2								
	A2	Depletion Adjusted Catch Scalar - 100% scalar	1	10	1									2	7	2								
Depletion-Based (Dynamic)	A3	DB-SRA (Depletion Adjusted) - 25% P*	1	22	5	4	3	5	1					2	22	5	2	4	2	2				
	A4	DB-SRA (Depletion Adjusted) - 50% P*	3	22	3	4	4	3	2					3	19	5	2	3	1	2				
	A5	DCAC (Depletion Adjusted) - 25% P*	0	7	1	1	1	3						2	4	1	1	1	4					
	A6	DCAC (Depletion Adjusted) - 50% P*	0	5	1	1	1	2						2	4	2	1	1	4					
Abundance-Based (Dynamic)	A7	Life History Analysis - 75% scalar								31	1	5	2								23	2	5	2
	A8	Life History Analysis - 100% scalar								34	3	6	2								21	3	4	2
	A9	FMSY/M (Low) - 75% scalar						8		19	0								6		20	3		
	A10	FMSY/M (Low) - 100% scalar						5		17	1								5		17	2		
	A11	FMSY/M (Hi) - 75% scalar						4		17	1								4		17	3		
	A12	FMSY/M (Hi) - 100% scalar						5		19	1								3		14	2		
Stock Assessment	R1	Delay-Difference - 75% scalar	2											1										
	R2	Delay-Difference - 100% scalar	1											3										

Type	Code	Methods	Snapper											Porgy										
			Obs err	Dep bias	Dep CV	FMSY/M	BMSY/B0	M	50% Mat	B bias	B CV	Von B K	L 1st Cap	Obs err	Dep bias	Dep CV	FMSY/M	BMSY/B0	M	50% Mat	B bias	B CV	Von B K	L 1st Cap
Catch-Based (Static)	M1	Median Catch - 3 Years	1											1										
	M2	Median Catch - 10 Years	1											1										
	M3	3rd Highest Catch	1											1										
Depletion-Based (Static)	M4	DB-SRA (Depletion Fixed @ 40%B0) - 69.4% scalar	0			0	0	12	1					1			1	1	10	1				
	M5	DB-SRA (Depletion Fixed @ 40%B0) - 83.4% scalar	0			0	0	14	1					0			1	1	10	1				
	M6	DB-SRA (Depletion Fixed @ 40%B0) - 91.3% scalar	1			0	0	15	1					0			1	1	9	1				
	M7	DCAC (Depletion Fixed @ 40%B0) - 69.4% scalar	0			0	0	11	1					0			1	1	8	1				
	M8	DCAC (Depletion Fixed @ 40%B0) - 83.4% scalar	0			0	0	13	1					1			1	1	9	1				
	M9	DCAC (Fixed Depletion @ 40%B0) - 91.3% scalar	0			0	0	14	1					1			1	1	9	1				
Catch-Based (Dynamic)	A1	Depletion Adjusted Catch Scalar - 75% scalar	1	12	3									2	9	4								
	A2	Depletion Adjusted Catch Scalar - 100% scalar	2	11	3									3	7	3								
Depletion-Based (Dynamic)	A3	DB-SRA (Depletion Adjusted) - 25% P*	4	35	14	2	7	19	3					3	21	6	4	4	2	6				
	A4	DB-SRA (Depletion Adjusted) - 50% P*	5	29	19	4	7	10	2					2	18	5	2	3	3	5				
	A5	DCAC (Depletion Adjusted) - 25% P*	0	15	1	3	3	7						1	8	2	2	2	4					
	A6	DCAC (Depletion Adjusted) - 50% P*	0	12	1	3	2	6						1	5	1	2	1	2					
Abundance-Based (Dynamic)	A7	Life History Analysis - 75% scalar								32	4	3	3								23	3	1	3
	A8	Life History Analysis - 100% scalar								37	4	4	2								27	3	2	4
	A9	FMSY/M (Low) - 75% scalar						16		26	3								11		22	2		
	A10	FMSY/M (Low) - 100% scalar						15		24	3								8		20	2		
	A11	FMSY/M (Hi) - 75% scalar						14		22	3								6		20	2		
	A12	FMSY/M (Hi) - 100% scalar						11		19	4								4		21	2		
Stock Assessment	R1	Delay-Difference - 75% scalar	2											2										
	R2	Delay-Difference - 100% scalar	2											2										

Type	Code	Methods	Sole											Rockfish										
			Obs err	Dep bias	Dep CV	FMSY/M	BMSY/B0	M	50% Mat	B bias	B CV	Von B K	L 1st Cap	Obs err	Dep bias	Dep CV	FMSY/M	BMSY/B0	M	50% Mat	B bias	B CV	Von B K	L 1st Cap
Catch-Based (Static)	M1	Median Catch - 3 Years	1											1										
	M2	Median Catch - 10 Years	2											1										
	M3	3rd Highest Catch	1											1										
Depletion-Based (Static)	M4	DB-SRA (Depletion Fixed @ 40%B0) - 69.4% scalar	0			1	0	12	1					0			1	0	10	1				
	M5	DB-SRA (Depletion Fixed @ 40%B0) - 83.4% scalar	0			1	0	12	1					1			1	0	11	1				
	M6	DB-SRA (Depletion Fixed @ 40%B0) - 91.3% scalar	0			2	0	12	1					1			1	0	12	1				
	M7	DCAC (Depletion Fixed @ 40%B0) - 69.4% scalar	0			1	0	9	0					0			0	0	9	1				
	M8	DCAC (Depletion Fixed @ 40%B0) - 83.4% scalar	0			1	0	11	0					0			1	0	10	1				
	M9	DCAC (Fixed Depletion @ 40%B0) - 91.3% scalar	0			1	0	11	0					0			1	0	11	1				
Catch-Based (Dynamic)	A1	Depletion Adjusted Catch Scalar - 75% scalar	0	12	2									2	8	2								
	A2	Depletion Adjusted Catch Scalar - 100% scalar	1	11	1									2	6	2								
Depletion-Based (Dynamic)	A3	DB-SRA (Depletion Adjusted) - 25% P*	2	18	5	2	3	3	1					2	30	9	3	4	13	6				
	A4	DB-SRA (Depletion Adjusted) - 50% P*	3	20	6	2	3	2	1					4	20	11	5	3	5	6				
	A5	DCAC (Depletion Adjusted) - 25% P*	1	7	2	2	1	3						0	13	1	3	2	4					
	A6	DCAC (Depletion Adjusted) - 50% P*	1	5	1	1	1	2						0	11	1	2	2	3					
Abundance-Based (Dynamic)	A7	Life History Analysis - 75% scalar								20	3	1	1								29	4	4	4
	A8	Life History Analysis - 100% scalar								23	3	2	1								33	5	5	4
	A9	FMSY/M (Low) - 75% scalar						12		24	2								12		25	2		
	A10	FMSY/M (Low) - 100% scalar						9		21	3								10		23	3		
	A11	FMSY/M (Hi) - 75% scalar						7		19	3								9		21	3		
	A12	FMSY/M (Hi) - 100% scalar						4		17	3								6		19	3		
Stock Assessment	R1	Delay-Difference - 75% scalar	1											1										
	R2	Delay-Difference - 100% scalar	2											1										

and maturity curves ("50%V–50%M") (Table App.C.7). The lack of sensitivity to different spatial parameterizations is supported further by a set of simulations that was conducted without any spatial structure (Appendix E). Spatial phenomenon such as refugia and stock viscosity lead to small reductions in the probability of overfishing (typically between 1 and 3%). In general the results of the spatially aggregated simulations were within 2% of those of the spatially disaggregated simulations, and did not provide any meaningful differences in the ranking of the methods. Only snapper were simulated with refuges, and these averaged only 5% of the population. Much larger differences in the performance results arising from spatially explicit and spatially aggregated operating models may be expected where refugia are larger.

4. Discussion

4.1. Performance of data-limited methods

Setting an ABC at average historical catch levels (methods M1–M3) is likely to lead to poor performance in cases where stocks are below their most productive levels. Generally, the performance of such methods was comparable to the status quo reference methods that simulated current catch or current fishing effort. Method M3, third-highest catch, generally performed worse than maintaining current fishing levels. The main reason for the poor performance of methods M1–M3 is the lack of feedback between stock depletion and the ABC. Recent historical catches rates were often higher than those associated with F_{MSY}, ensuring that using their average as an ABC perpetuated overfishing. These methods lead to positive feedback between past and future ABC recommendations; future ABCs are based on previous ABCs and therefore tend toward a stable value over time. If the initial ABC is too high, exploitation rates become exponentially larger over time. In contrast, if this value is too low the stock tends toward some biomass above B_{MSY}. Consequently, these methods are often divergent and move the stock away from B_{MSY}.

Other static management methods that do not include feedback between the ABC recommendation and stock status can provide good performance, but only when stocks are at intermediate levels of depletion (e.g., the PFMC DB-SRA and DCAC methods M4–M9). While the performance of the static methods was generally poor at low stock levels, the static DB-SRA method still led to lower probabilities of overfishing and higher yields than the average catch methods (M1–M3). Unsurprisingly, methods that dynamically account for population changes achieved better performance when the stock is not near B_{MSY}. This was not the case for DCAC, which is designed to return a proxy for MSY, which is not an appropriate basis for OFLs for stocks at low population levels (as acknowledged MacCall (2009)). The dynamic DB-SRA and F_{MSY}/M ratio methods (A3 and A9) generally led to the best performance by some margin. While the aggregate performance of these methods may appear satisfactory, it is strongly affected by bias in two key inputs: depletion (DB-SRA) and current stock biomass (F_{MSY}/M methods). Methods which involve estimates of biomass or current depletion (rather than assumptions about them) would, however, generally not be considered to be data-poor, but rather data-moderate (PFMC, 2010; NPFMC, 2012).

The simulation testing of ABC control rules (e.g., 75% and 100% scalar multipliers) revealed that the largest downward adjustments in the OFL often led to higher expected long-term yields and lower probabilities of overfishing (e.g., F_{MSY}/M ratio methods A9 and A10). This was particularly the case for simulations starting below 50% B_{MSY} where lower exploitation rates could allow rebuilding to more productive stock sizes. However, the range of downward adjustments was not sufficient in some instances to achieve high probabilities of rebuilding. For example, the three ABC control rules based on methods M4–M9 ranged from a 9% to a 30% reduction in the OFL. The results of all three multipliers were similar, and did not span a sufficiently wide range of adjustment to allow stocks to recover from low levels, when depletion is assumed a priori to 40% (e.g., methods M4–M9).

4.2. Sensitivity of performance to inputs and value of information

In general, the performance differences were much greater across methods than across life-history types. The exception to this was butterfish. All methods led to relatively high rates of overfishing for butterfish without necessarily leading to stock declines or reductions in long-term expected yield because of the short life span and high recruitment variability of this stock. The biomass of butterfish can easily depart from the mean by a factor of 2 in the absence of fishing, making natural variability in productivity a much stronger determinate of stock status than exploitation rate. The results for butterfish demonstrate the challenge of developing management systems for short-lived species. MSE for prawn species that examine both input (effort) and output (catch quota) controls (Dichmont et al., 2006, 2012) conclude that the effective use of quotas in such cases is dependent on the ability to predict and monitor recruitment. It may be beneficial to track current abundance and maintain close control of exploitation levels to prevent forgone yields and/or problematic stock declines for short-lived species. It follows that methods that rely on current information and aim for fixed exploitation rates such as the F_{MSY}/M ratio methods may be particularly suitable for species of short life history.

Previous simulation evaluations of DB-SRA and DCAC found sensitivity to misspecification in natural mortality rate for long-lived stocks (Wetzel and Punt, 2011), a result which is corroborated here for snapper and rockfish. This is due to propagating this error over a larger number of age classes and hence a larger fraction of the population.

The simulation of spatial population and fishing dynamics had very little impact on performance. All methods showed relatively weak sensitivity to variability in simulated spatial targeting, stock viscosity or spatial heterogeneity; a MSE with no spatial dynamics led to very similar results. Spatial phenomena such as refugia from fishing and stock viscosity led to very small reductions in the probability of overfishing relative to the differences among methods and simulated life-histories. This suggests that the subtleties of spatial stock dynamics are comprehensively overwhelmed by general problems associated with the inaccuracy and imprecision of the principal inputs such as natural mortality rate and stock size for the stocks simulated in this research. It is conceivable that spatial effects may be more critical for other stocks, for example sessile species or those that experience greater refuge from fishing.

All of the methods were most sensitive to imperfect information regarding either current stock depletion or current biomass. Consistent bias in these inputs strongly affected the expected probability of overfishing and long-term yield. On the other hand, relatively high imprecision in these estimates had little effect on performance: year on year, the estimates could vary strongly from the “true” underlying value of depletion or biomass. The dynamic DB-SRA method could lead to high probabilities of declining below 50% B_{MSY} when starting above B_{MSY}. This was due to the specification of OFLs much higher than MSY due to a positively biased input for depletion. An alternative ABC control rule which applies a downward adjustment to the smaller of the OFL or MSY may help to combat this problem and substantially improve the performance of the dynamic DB-SRA method in such instances.

4.3. Quantifying inputs

The inputs to these data-limited methods focus on those that can be developed quickly from existing sources, as opposed to those that require future data collection efforts. Given that the intent of the data-poor assessment is to provide information for immediate use, the latter category of inputs is less relevant to this discussion. However, additional or improved inputs may be needed if an attempt at assessment falls short due to lack of information, or if the results engender an urgent desire for a “more complete” assessment. A wide range of alternatives exist for supplementary data collection, depending on available labor and funding, and the time horizon for data delivery, but the result is to move toward a more data-rich approach that falls outside the scope of this study.

4.3.1. Depletion

The assessment methods that perform best included estimates of current depletion or abundance so it is instructive to discuss how these inputs may be obtained. Of these, depletion is perhaps the most difficult to obtain for data-poor stocks. Depletion is a data-rich quantity in many respects; it requires broad knowledge of stock trend, which in turn defines a data-rich stock in this paper and elsewhere (e.g., Punt et al., 2011). However, a case may be made that expert knowledge about depletion could be derived from anecdotal information such as changes in the spatial range of fishing. Expert judgment is especially useful when assessments have been carried out for other local stocks, and the similarity of fishing operations for the data-poor stock is suspected or known. For example, based on a calibration to 30 data-rich stock assessments, Productivity Susceptibility Analysis (Patrick et al., 2009) has been used by the PFMC to determine the mean of the prior for depletion when applying DB-SRA.

In some cases, a time series of fishery-independent surveys exists for other species, and the data-poor species may be caught occasionally. Although the data may contain an excessive number of “zeroes” it is often possible to derive an abundance index or estimate of depletion from a remarkably small number of positive samples, even if the time series has to be collapsed into a few multi-year time blocks. Examples of fishery-independent surveys include the Triennial trawl survey and slope surveys of the US West Coast (NMFS, 2013) and the MARMAP (2013) survey of the South Atlantic.

Trends in abundance inferred from catch and effort data can be included in methods such DB-SRA to update the depletion prior (Cope et al., 2013). Although historical effort is usually not known, it may be possible to “borrow” a time series of fishing rate estimates from assessments of other species in the region. Punt et al. (2011) have explored simultaneous assessments of multiple species using this “Robin Hood” approach. Other ways to construct estimates of depletion include recreational fishing databases (e.g. RecFIN, 2013) or the use of scientific observer data (NMFS (2013) includes a discussion of these sources of depletion information).

Our analysis of the value of information indicates that considerable *imprecision* in depletion estimates does not lead to dramatic loss of yield or increase in the probability of overfishing. *Bias* in depletion, on the other hand, strongly determines performance. This is potentially problematic because of difficulties in acquiring new information about past abundance trends.

4.3.2. Natural mortality rate

The DB-SRA, DCAC and F_{MSY}/M ratio methods all rely on an estimate of M, a common input in most stock assessments (the main exception being surplus production models). Although M is an uncertain parameter, stock assessments require only an approximate value. If tentative ages can be determined, covariates such as maximum age and von Bertalanffy growth parameters are estimable from quite small samples; tropical fishes lacking clear age indicators are more difficult. Useful meta-analyses have been published by Pauly (1980), Hoenig (1983), Hewitt and Hoenig (2005), and Gislason et al. (2010), among many others. If uncertainty in the value of M remains problematic, it may suffice to choose a most likely value of M from a simple list of candidate values (e.g., 0.2, 0.1, 0.05, 0.025 yr^{-1}). Note that many of these data-poor methods fail if $M > 0.2 yr^{-1}$, and values below 0.025 yr^{-1} for M are rare in fish.

While DB-SRA and DCAC have low fishery data requirements (historical catches), the remaining inputs are parameters and variables that strongly determine the methods' outcomes. Although direct estimation of these quantities requires conventional approaches used in data-rich assessments or meta-analyses (e.g., Punt et al., 2011; Zhou et al., 2012; Thorson et al., 2012b), data-poor assessments often require us to postulate values of key parameters by analogy to data-rich cases. Development of appropriate meta-analyses is an active area of fishery research that has gained impetus from the requirements of data-poor assessment methodologies.

4.3.3. Current abundance

In instances where it is not possible to estimate current depletion, future data-gathering efforts may focus on the estimation of current abundance which is an input to the F_{MSY}/M and life-history methods.

There are several possible ways to estimate current biomass that differ by cost and the assumptions on which they rely. The most conventional is a "fishery independent" research survey that uses a variety of fishing gears to sample the population from which total biomass may be extrapolated (Doubleday and Rivard, 1981; Gunderson, 1993). In the Gulf of Alaska and the Bering Sea, estimates of abundance from fishery-independent surveys are used in the F_{MSY}/M method to set ACLs for several stock complexes such as skates, sculpins, crab, and rockfishes (NPFMC, 2012). The principal limitation of surveys is their considerable cost which may not be justified in many data-limited situations, for example where the primary source of exploitation is bycatch. In addition, many species are unlikely to be fully selected by the survey gear or estimates from density in areas which can be surveyed may be extrapolated incorrectly to areas that cannot be surveyed leading to persistent bias in estimates of abundance. Such bias may dramatically affect the reliability of data-limited methods using these data.

An alternative approach to current abundance is to divide current catch by an estimate of current exploitation rate. If assessments have been carried out for other species, it may be possible to "borrow" their estimated fishing mortality rates. Punt et al. (2011) use this "Robin Hood approach" in simultaneous assessments of multiple species. Two possible direct means of estimating current exploitation rate are a tagging experiment or a catch curve analysis. The concept of mark-recapture analysis has a long history in fisheries science and was discussed at length by Beverton and Holt (1957). Tagging may be expensive, but can provide a relatively precise estimate of current fishing mortality rate and abundance. There are often challenges to the ready interpretation of these data, including tag mortality, shedding, reporting and detection rates, and a program may take many years especially if exploitation rate is low. To obtain exploitation estimates that can be generalized to the population requires knowledge of spatial distribution that may not be available in many data-limited situations. Perhaps the most important limitations of mark recapture analysis is that many species of fish are difficult to tag in sufficient numbers or not suitable candidates due to high post-release mortality rate or tag-induced mortality rate.

Catch-curve analysis can also provide estimates of current mortality rates, and is likely to be most successful in cases where fishing mortality rate, recruitment strength and age-vulnerability to fishing can be assumed to be relatively constant over recent years. Catch curve analysis (Ricker, 1975) assumes that after a certain age, individuals experience the same fishing mortality rate, allowing the descending proportion of catch-at-age (or catch-at-length) to be interpreted in terms of total mortality. An estimate of natural mortality rate is needed to separate fishing mortality from the total mortality rate estimated by catch-curve analysis. In a data-limited setting the primary advantage of catch-curve analysis is that it does not require historical data and relies only on catch composition data that can be collected today. Catch curves can be based on age- or length-composition data and can be used to form the basis for control rules for data-limited species (e.g., Klaer et al., 2012). There are a number of methods to account for temporal variability in recruitment and selectivity if multiple years of age-composition data are available (e.g., Schnute and Haigh, 2007). Despite the limitations of catch-curve analysis, it might produce estimates of current biomass that are no more biased or uncertain than the imperfect knowledge of biomass simulated in this analysis. This should be the focus of future simulation evaluation.

4.4. Methods that could not be simulation tested

There are data-limited assessment methods for setting catch limits that could not be simulation tested. These methods either did not provide estimates for OFLs (the methods of Patrick et al., 2009; Martell and Froese, 2012; Thorson et al., 2012a; Costello et al., 2012; Cope and Punt, 2009) or involved expert judgment that could not be simulated (the methods of Berkson et al., 2011; Punt et al., 2011).

The method of Martell and Froese (2012) aims to estimate MSY by reconstructing a stock history according to catches and discarding those simulations that cross certain thresholds (e.g., that fall out of a range of current stock depletion such as 5–95% of unfished biomass). This "MSY depletion method" is theoretically similar to DCAC. A central finding of Martell and Froese (2012) is that MSY may be well defined despite only weak prior information about maximum stock size, stock productivity and current depletion. However, this finding also explains our inability to include this approach in our analysis. While MSY is a theoretical quantity relating to the most productive level of depletion, the OFL is determined by current stock depletion (e.g., it tends to zero as the stock declines). It follows that MSY does not provide a means of setting the OFL without a control rule. Since the OFL can range from much higher than MSY to zero, the success of the method would rely on the control rule. It could be argued that a control rule should also be applied to DCAC since it is also an approximation of MSY. However in line with the recommendations of the PFMC (PFMC, 2010) we tested DCAC as a method of determining the OFL without such a control rule.

Thorson et al. (2012a) and Costello et al. (2012) use covariate information, such as life history characteristics and landings data to inform a predictive model of current stock depletion. These approaches use correlations between assessed stock status and other covariates to extrapolate the stock status of fisheries that are not assessed. It is possible that these methods could be adapted to provide OFL recommendations. However, doing so would require assumptions about the productivity of the stock with declining biomass (i.e., the shape of the productivity curve). It may be possible to combine these methods or DCAC or the method of Martell and Froese (2012).

Punt et al. (2011) propose a "Robin Hood" method in which data-rich assessments are used to inform the spawning stock biomass and exploitation history of data-limited stocks that are subject to fishing by the same fleets. A central assumption of this method is that the different stocks have comparable trends in exploitation rate. As such, the method relies on the existence of a contingent data-rich stock and a process to assess whether exploitation rates are similar. The choice of which fleets have the same trends in exploitation rate is based on expert judgment, which prevented a full evaluation of the method.

Cope and Punt (2009) outline a length-based approach that relates the observed fractions of fish of different classes (e.g., fraction mature) to stock status. While length-based reference points could provide a basis for designing control rules that provide OFL recommendations, these rules have yet to be established (Cope and Punt, 2009).

4.5. Limitations

Assumptions about how accurately and precisely the inputs to the data-limited methods may be quantified determines performance. It should be emphasized that the results are a product of the specific conditions of the simulation. For example, we may have found that methods which rely on M performed substantially better had the extent of error associated with M been assumed to be unrealistically low. This points to a fundamental circularity in this analysis, one of simulating knowledge in inputs to methods that are to be applied in instances where these inputs are not known. All of the methods evaluated performed poorly when their fundamental assumptions were invalid or inputs were strongly mis-specified. We recommend that when reviewing the performance of the data-limited methods, the reader should take care to consider the sensitivity of the performance to misspecification in inputs (as presented in Table 7 for example).

The objective of this research was to evaluate the impact of the data-limited methods regardless of the rate of compliance. In all of the simulations we assumed that the ABC recommendations were taken as catch and no implementation error was simulated. In practice, there are often overages or shortfalls that affect the level of future catch limits. It is possible that implementation error may interact with some data-limited methods and alter their relative performance. However, since all methods provide the same type of advice (i.e., catch limits) it is probable that this additional source of error would have had a comparable impact across methods and would limit the generality of the results while reducing the clarity of the inter-method comparisons.

4.6. Conclusions and recommendations

- In circumstances where only fishery catch data are available, this simulation evaluation indicates average catch methods such as median catch over the most recent 10 years or third highest catch cannot be expected to provide a better basis for management than maintaining current catch or effort levels. These methods often perform even worse than the status quo methods of current catch or current effort when biomass starts below B_{MSY}. However, the catch-based methods appear to provide performance more comparable to that of the other methods if it can be established that a stock is above B_{MSY}.
- Additional information regarding depletion, historical effort, or current abundance can be very valuable. Our analysis points to large expected gains in yield for all stock types (except high-M stocks such as butterfish) when stocks are heavily depleted given information about depletion or trend in relative abundance, with more modest gains for less depleted stocks. When considering how to obtain data in addition to historical catch, perhaps the most cost-effective avenue for investigation is the availability of unprocessed data. For example, fishing effort data that may be used to calculate an index of historical abundance or for estimating current depletion. Multispecies surveys may also be available from which a time-series of abundance could be constructed (e.g., MARMAP, 2013; West Coast trawl surveys NMFS, 2013). A research priority is summarizing these data sources and characterizing stocks according to uncertainty regarding stock status and the potential benefits of obtaining additional data. Where historical abundance trends or effort data are not available there is an onus on the collection of current abundance information, for example using fishery independent surveys, catch curve analysis or tagging studies. Simulation evaluation may offer a basis for determining the cost-benefit of new data-collection programs by quantifying the potential for additional long-term yields.
- The mixed performance of the delay-difference methods provides food for thought for those analysts seeking to evaluate data-limited methods by comparison with stock assessments. The delay-difference models applied in this analysis assumed perfect knowledge of historical effort, growth, natural mortality rate, and the age that individuals are vulnerable to fishing. Nevertheless, these assessments assume stationary stock dynamics and a linear relationship between historical fishing effort and fishing mortality rate, assumptions that are commonly violated in these simulations. That performance for this method was “mixed” runs contrary to the view of data-rich stock assessments as a “gold standard” against which other approaches may be compared. Our simulation evaluation also confirms that classifying stocks solely according to the amount and types of data available may not be appropriate. A large quantity of data is no guarantee of reliable information on which to base decision making (data-rich stocks are often information poor). The way in which data inform management recommendations relies to a large extent on the validity of the assumptions of the assessment tool. For example, detailed historical data for a short-lived species such as butterfish should not necessarily motivate the use of a conventional data-rich assessment approach that may offer less reliable management advice than a simpler approach using a smaller amount of data that instead, provide information about current stock characteristics.
- Some of the terminology surrounding data-limited methods has the potential to be strongly misleading. One example is the term P^* (probability of overfishing). This simulation study and Punt et al. (2012) found that P^*s of 25% and 50% rarely corresponded to these probabilities of overfishing. Nor did a 25% P^* rule lead to half the probability of overfishing exhibited by a 50% P^* rule. Based on this terminology, decision makers may be led to believe they are choosing a specific outcome and this simulation evaluation reveals that this may not be the case.
- We have evaluated a broad suite of data-limited methods. Certain data-limited methods (e.g., the ‘Robin Hood’ method, the ORCS approach, PSA analysis) have been proposed, but could not be simulation-tested. We recommend that editors of journals who consider publishing new data-poor methods request authors to minimally outline how their method can be tested. Ideally, a reference set of simulation data sets should be made available to allow the results of this paper to be supplemented with those for new data-limited methods.
- Finally, the focus of this paper is on methods that have been identified for use in the management of fish stocks in U.S. waters. However, establishing data-limited methods is particularly relevant to developing countries where there is often less complete reporting of fishery data and fewer resources dedicated to analysis. Moreover, a broader suite of types of assessment methods could be examined for countries which mandate use of control rules, but are less prescriptive regarding the structure of control rules than the U.S. (see, for example, Smith et al., 2009).

Acknowledgements

TRC is grateful for the support of the Gordon and Betty Moore Foundation. AEP acknowledges support from NOAA Grant NA10OAR4320148. This is JISAO contribution 2173. Many thanks to Chris Legault, Jim Berkson and two anonymous reviewers for their helpful 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.2013.12.014.

References

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

Berkson, J., Barbieri, L., Cadrin, S., Cass-Calay, S., Crone, P., Dorn, P., Friess, C., Kobayashi, D., Miller, T.J., Patrick, W.S., Pautzke, S., Ralston, S., Trianni, M., 2011. Calculating acceptable biological catch for stocks that have reliable catch data only. NOAA Technical Memorandum, NMFS-SEFSC-616. http://www.pifsc.noaa.gov/tech/NOAA_Tech_Memo_SEFSC_616.pdf

Beverton, R.J.H., Holt, S.J., 1957. On the dynamics of exploited fish populations. Fishery Investigations Series II, vol. XIX. Ministry of Agriculture, Fisheries and Food, London.
Butterworth, D.S., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56, 985–998.
Butterworth, D.S., Johnston, S.J., Brandão, A., 2010. On pre-testing the likely efficacy of suggested management approaches for data-poor fisheries. Managing Data-poor Fisheries: Case Studies, Models and Solutions 1, 107–130.
Cochrane, K.L., Butterworth, D.S., De Oliveira, J.A.A., Roel, B.A., 1998. Management procedures in a fishery based on highly variable stocks and with conflicting objectives: experiences in the South African pelagic fishery. Rev. Fish. Biol. Fish. 8, 177–214.
Cope, J., Punt, A.E., 2009. Length-based reference points for data-limited situations: applications and restrictions. Mar. Coast. Fish. Dyn. Manage. Ecosyst. Sci. 1, 169–186.
Cope, J., Dick, E.J., MacCall, A., Monk, M., Soper, B., Wetzel, C., 2013. Data-moderate stock assessments for brown, China, copper, sharpchin, stripetail and yellowtail rockfishes and English and rex soles in 2013. http://www.pcouncil.org/wp-content/uploads/.F5a_ATT1_DM_ASSMT_2013_ELECTRIC_JUN2013BB.pdf
Costello, C., Ovando, D., Hilborn, R., Gains, S.D., Deschenes, O., Lester, S.E., 2012. Status and solutions for the world's unassessed fisheries. Science 338, 517–520.
Deriso, R.B., 1980. Harvesting strategies and parameter estimation for an age-structured model. Can. J. Fish. Aquat. Sci. 37, 268–282.
Dichmont, C.M., Deng, R.A., Punt, A.E., Venebles, W., Haddon, M., 2006. Management strategies for short-lived species: the case of Australia's Northern Prawn Fishery 1. Accounting for multiple species, spatial structure and implementation uncertainty when evaluating risk. Fish. Res. 82, 204–220.
Dichmont, C.M., Deng, R.A., Punt, A.E., Venables, W., Hutton, T., 2012. From input to output controls in a short-lived species: the case of Australia's Northern Prawn Fishery. Mar. Freshw. Res. 63, 727–739.
Dick, E.J., MacCall, A.D., 2010. Estimates of sustainable yield for 50 data-poor stocks in the Pacific Coast groundfish fishery management plan. Technical memorandum. Southwest fisheries Science Centre, Santa Cruz, CA. National Marine Fisheries Service, National Oceanic and Atmospheric Administration of the U.S. Department of Commerce, NOAA-TM-NMFS-SWFSC-460.
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.
DiCosimo, J., Methot, R.D., Ormseth, O.A., 2010. Use of annual catch limits to avoid stock depletion in the Bering Sea and Aleutian Islands management area (Northeast Pacific). ICES J. Mar. Sci. 67, 1861–1865.
Doubleday, W.G., Rivard, D., 1981. Bottom trawl surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58, 237.
Gislason, H., Daan, N., Rice, J.C., Pope, J.G., 2010. Size, growth, temperature and the natural mortality rate of fish. Fish Fish. 11 (2), 149–158.
Gulland, J.A., 1971. The Fish Resources of the Ocean. Fishing News Books, West Byfleet, UK.
Gunderson, D.R., 1993. Surveys of Fisheries Resources. Wiley, New York, pp. 248.
Hewitt, D.A., Hoenig, J.M., 2005. Comparison of two approaches for estimating natural mortality rate based on longevity. Fish. Bull. 103, 433–437.
Hoenig, J.M., 1983. Empirical use of longevity data to estimate mortality rates. Fish. Bull. 82, 898–903.
Klaer, N.L., Wayte, S.E., Fay, G., 2012. An evaluation of the performance of a harvest strategy that uses an average-length-based assessment method. Fish. Res. 134–136, 42–51.
MAFMC, 2010. Amendment 11 to the Atlantic mackerel, squid and butterfish fishery management plan. Mid Atlantic Fishery Management Council. US National Marine Fisheries Service http://www.mafmc.org/meeting_materials/2010/February/Tab05_Squid,Mackerel,Butterfish_Amendment_11_Public_Hearing_Document.pdf
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.
MARMAP, 2013. Marine Resources Monitoring, Assessment, & Prediction. South Carolina Department of Natural Resources http://www.dnr.sc.gov/marine/mrri/MARMAP/index.html
Martell, S., Froese, R., 2012. A simple method for estimating MSY from catch and resilience. Fish Fish., http://dx.doi.org/10.1111/j.1467-2979.2012.00485.x.
NMFS, 2011. Fisheries of the Northeastern United States: Atlantic Mackerel, Squid and Butterfish (MSB) Fishery; Revision of 2011 Butterfish Specifications, 76 Fed. Reg. 13887, 13888 (15.03.11).
NMFS, 2013. Assessment methods for data-moderate stocks. Report of the methodology review panel meeting. National Marine Fisheries Service, Alaska Fisheries Science Centre, Seattle, Washington.
NPFMC, 2012. Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands Management Area. North Pacific Fishery Management Council, Anchorage, Alaska http://alaskafisheries.noaa.gov/npfmc/PDFdocuments/fmp/BSAI/BSAI.pdf
NPFMC, 2013. Stock Assessment and Fishery Evaluation Reports for 2013. North Pacific Fishery Management Council, Anchorage, Alaska http://www.afsc.noaa.gov/refm/stocks/assessments.htm
NSG, 2009. Magnuson–Stevens Act provisions; annual catch limits, National Standard Guidelines, final rule, 50 CFR part 600. Federal Register 74 (11).
Patrick, W.S., Spencer, P., Ormseth, O., Cope, J., Field, J., Kobayashi, D., Gedamke, T., Cortés, E., Bigelow, K., Overholtz, W., Link, J., Lawson, P., 2009. Use of productivity and susceptibility indices to determine stock vulnerability, with example applications to six U.S. fisheries. NOAA Fisheries, NOAA Technical Memorandum NMFS-F/SPO-101.
Pauly, D., 1980. On the interrelationships between natural mortality, growth parameters and mean environmental temperature in 175 fish stocks. J. Cons. Int. Explor. Mer. 39 (2), 175–192.
PFMC, 2010. Pacific Coast Groundfish Fishery Management Plan for the California, Oregon and Washington Groundfish Fishery, as amended through Amendments 23 and 16-5. Pacific Fishery Management Council, Portland, OR.
Punt, A.E., Smith, D.C., Smith, A.D.M., 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES J. Mar. Sci. 68, 972–981.
Punt, A.E., Siddeek, M.S.M., Gdearber-Yots, B., Dalton, M., Rugulo, L., Stram, D., Turnock, B.J., Zheng, J., 2012. Evaluating the impact of buffers to account for scientific uncertainty when setting TACs: application to red king crab in Bristol Bay, Alaska. ICES J. Mar. Sci. 69, 624–634.
R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
RecFIN, 2013. Recreational Fisheries Information Network. Pacific States Marine Fisheries Commission http://www.recfin.org
Restrepo, V.R., Thompson, G.G., Mace, P.M., Gabriel, W.L., Low, L.L., MacCall, A.D., Methot, R.D., Powers, J.E., Taylor, B.L., Wade, P.R., Witzig, J.F., 1998. Technical Guidance on the Use of Precautionary Approaches to Implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act. NOAA Technical Memorandum NMFS-F/SPO-31, 54 pp.
Ricker, W.E., 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board. Can. 191, 382.
SAFMC, 2011. Comprehensive Annual Catch Limit (ACL) Amendment for the South Atlantic region. South Atlantic Fishery Management Council, North Charleston, South Carolina http://sero.nmfs.noaa.gov/sf/pdfs/Comp%20ACL%20Am%20101411%20FINAL.pdf
Schnute, J., 1985. A general theory for analysis or catch and effort data. Can. J. Fish. Aquat. Sci. 42, 414–429.
Schnute, J.T., Haigh, R., 2007. Compositional analysis of catch curve data, with an application to *Sebastes malinger*. ICES J. Mar. Sci. 64, 218–233.
Smith, D., Punt, A., Dowling, N., Smith, A., Tuck, G., Knuckey, I., 2009. Reconciling approaches to the assessment and management of data-poor species and fisheries with Australia's harvest strategy policy. Mar. Coast. Fish. 1, 244–254.
Thompson, G.G., 1993. A proposal for a threshold stock size and maximum fishing mortality rate. In: Smith, S.J., Hunt, J.J., Rivard, D. (Eds.), Risk Evaluation and Biological Reference Points for Fisheries Management. Can. Spec. Publ. Fish. Aquat. Sci. 120. , pp. 303–320.
Thorson, J.T., Branch, T.A., Jensen, O.P., 2012a. Using model-based inference to evaluate global fisheries status from landings, location, and life history data. Can. J. Fish. Aquat. Sci. 69, 645–655.
Thorson, J.T., Cope, Branch, T., Jensen, O., 2012b. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Can. J. Fish. Aquat. Sci. 69, 1556–1568.
Vaughan, D.S., Punt, A.E., Legault, C., Berkson, J., 2012. Peer review of Wilberg et al. "Evaluation of acceptable biological catch (ABC) control rules for mid-Atlantic stocks". Prepared for the Mid-Atlantic Fishery Management Council. http://www.mafmc.org/meeting_materials/SSC/2012-03/REVIEW%20OF%20WILBERG%20ET%20AL%20Final.pdf
Walters, C., Martell, S.J.D., 2002. Stock assessment needs for sustainable fisheries management. Bull. Mar. Sci. 70, 629–638.
Wetzel, C.R., Punt, A.E., 2011. Performance of a fisheries catch-at-age model (stock synthesis) in data-limited situations. Mar. Freshw. Res. 62, 927–936.
Wilberg, M.J., Miller, T.J., Wiedenmann, J., 2011. Evaluation of Acceptable Biological Catch (ABC) control rules for Mid-Atlantic stocks. Report to the Mid-Atlantic Fishery Management Council. UMCES, CBL 11-029. Technical Report Series No. TS-622-11. http://www.mafmc.org/meeting_materials/SSC/2012-03/Final%20MAFMC%20Report%20revised.pdf
Zhou, S., Yin, S., Thorson, J., Smith, T., Fuller, M., 2012. Linking fishing mortality reference points to life history traits: an empirical study. Can. J. Fish. Aquat. Sci. 69, 1292–1301.