Performance review of simple management procedures

Thomas R. Carruthers, Laurence T. Kell, Doug D. S. Butterworth, Mark N. Maunder, Helena F. Geromont, Carl Walters, Murdoch K. McAllister, Richard Hillary, Polina Levontin, Toshihide Kitakado, and Campbell R. Davies

SEDAR49-RD-06

June 2016

ICES Journal of Marine Science Advance Access published November 20, 2015

ICES Journal of Marine Science

ICES International Council for the Exploration of the Sea
CIEM Conseil International pour l'Exploration de la Mer

ICES Journal of Marine Science; doi:10.1093/icesjms/fsv212

Performance review of simple management procedures

Thomas R. Carruthers[1]*, Laurence T. Kell[2], Doug D. S. Butterworth[3], Mark N. Maunder[4], Helena F. Geromont[3], Carl Walters[1], Murdoch K. McAllister[1], Richard Hillary[5], Polina Levontin[6], Toshihide Kitakado[7], and Campbell R. Davies[5]

[1]*Fisheries Centre, AERL, University of British Columbia, 2202 Main Mall, Vancouver, BC, Canada V6T 1Z4*
[2]*International Commission for the Conservation of Atlantic Tunas, Calle Corazón de María, 8, Madrid 28002, Spain*
[3]*Department of Maths and Applied Maths, University of Cape Town, Rondebosch 7701, South Africa*
[4]*Inter-American Tropical Tuna Commission, 8604 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA*
[5]*CSIRO Marine Laboratories, Castray Esplanade, Hobart, TAS 7000, Australia*
[6]*Centre for Environmental Policy, Imperial College London, Silwood Park, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK*
[7]*Tokyo University of Marine Science and Technology, 5-7, Konan 4, Minato-ku, Tokyo 108-8477, Japan*

**Corresponding author: tel: + 1 604 822 6903; fax: + 1 604 822 8934; e-mail: t.carruthers@fisheries.ubc.ca*

Carruthers, T. R., Kell, L. T., Butterworth, D. D. S., Maunder, M. N.Geromont, H. F., Walters, C., McAllister, M. K., Hillary, R. Levontin, P., Kitakado, T., and Davies, C. R. Performance review of simple management procedures. – ICES Journal of Marine Science, doi: 10.1093/icesjms/fsv212.

Received 7 May 2015; revised 19 October 2015; accepted 22 October 2015.

Using a management strategy evaluation approach, we compare a range of new and established management procedures (MPs) for setting catch-limits in fisheries. Performance is evaluated with respect to fish life history type, level of stock depletion, data quality, and autocorrelation in recruitment strength. We quantify the robustness of each MP with respect to the various observation processes. Methods using observations of absolute biomass or stock depletion offer the best overall performance and this is consistent across life history types, data qualities, and stock depletion levels. Simple MPs can outperform conventional data-limited methods and data-rich assessments that use time-series of catch and effort data. MP performance is most sensitive to biases in catch data. Our results indicate that often tuning MPs for specific stocks is important, though this may not be viable in data-poor assessment scenarios because of insufficient data and analysis resources.

Keywords: data-limited, data-poor, fisheries management, harvest strategy, management procedure, management strategy evaluation, simulation, stock assessment.

Introduction

Management advice for the most economically important fish stocks, and increasingly also for bycatch species, is based on stock assessment. Under the stock assessment paradigm, models of stock dynamics are fitted to detailed fishery-dependent and -independent data and then used to assess historical stock status, derive reference points, and, in some cases, predict the likely impact of alternative management options (e.g. total allowable catches and measures to change the relative mortality among age classes). Stock assessments are updated periodically to include new data and to assess whether management recommendations require revision according to changes in estimates of exploitation level, stock status, and productivity. The assumptions of the stock assessment may be updated regularly by scientists, in some cases yearly (Hilborn, 2003).

Management strategy evaluation (MSE, Cochrane *et al.*, 1998; Butterworth and Punt, 1999; Butterworth, 2007) is an alternative fisheries management paradigm, which focuses on the relative performance of alternative management procedures (MPs, also known as, harvest strategies) to meet specified management objectives. It differs from the assessment approach in that detailed fishery data are used to condition multiple operating models (OMs), simulation models that represent alternative plausible hypotheses about fishery and population dynamics, rather than selecting a "best assessment" and conducting sensitivity tests as the basis for management advice.

These simulations are then used to tune, evaluate, and select an MP to be used as the basis for providing management recommendations. Commonly, the data, "assessment model", and decision rules that constitute the MP are much simpler than a conventional stock assessment (Punt and Donovan, 2007). Referred to in this study as simple MPs, these rely only on recent information regarding size composition, trends in abundance, and catch data. Instead of using stock assessment as the primary source of management advice, simple MPs may be used to generate routine management advice while the operating model is updated to accommodate new data at longer time intervals (e.g. CCSBT, 2011). There is increasing evidence that simple MPs can perform as least as well as conventional stock assessments in providing reliable management advice (Geromont and Butterworth, 2015a).

MSE typically adds stability to the management decision process by identifying realistic management objectives through stakeholder participation, followed by a thorough evaluation of trade-offs arising under alternative MPs, where the evaluation encompasses a range of plausible past and future scenarios and sources of uncertainty (e.g. Rockmann *et al*., 2012). MSE can also be used to guide the scientific process by identifying where the reduction of scientific uncertainty is most critical so as to provide the best research, monitoring, and enforcement (Fromentin *et al*., 2014).

In most of the management settings where simple MPs have been developed, the data have also been sufficient to support conventional stock assessments (e.g. Namibian Hake—Butterworth and Geromont, 2001; South African Hake—Rademeyer *et al*., 2008; Southern Bluefin Tuna—CCSBT, 2011; US west coast flatfish and rockfish—Wetzel and Punt, 2011; Australian fisheries—Smith *et al*., 2013). However, there has been an increasing interest in quantitative methods to support management decision-making in data-limited fisheries. [We define 'data-rich' as situations where sufficient data are available to conduct a conventional stock assessment (Punt *et al*., 2011). This includes simple stock assessment methods that typically require >15 years of relative abundance or fishing effort data in addition to catch data. We define all other data situations under the heading 'data-limited', which includes 'data-moderate' and 'data-poor'. Data-moderate situations have some form of current information about stock levels that may be observations of absolute biomass, relative abundance, or stock depletion. Data-poor refers here to situations where only historical catches and some catch composition data (e.g. length data) or life history information are available.] Approaches, such as depletion-corrected average catch (DCAC—MacCall, 2009), depletion-based stock reduction analysis (DB-SRA—Dick and MacCall, 2011), and fishing at a fixed fraction of natural mortality rate (Fratio—Walters and Martell, 2002), are currently used as components of MPs for managing data-limited fisheries and have been subject to simulation testing (Carruthers *et al*., 2014). Many of these are closely related to stock assessments; they are based on comparable biological models and rely on many of the same assumptions (e.g. DB-SRA). However, recent research has sought to develop and test new data-limited MPs that require fewer assumptions about underlying population dynamics and make management recommendations using only recent time-series data such as catches and catch-per-unit-effort data (e.g. Maunder, 2014; Geromont and Butterworth, 2015b).

In this study, we compare the performance of a suite of MPs that have been described in the primary and grey literature. The MPs were chosen to span a range of assumptions and data requirements. We also include new approaches that operate on alternative fishery information. A number of the MPs of this study have been parameterized according to simulations specific to a particular management scenario. We refer to these as "tuned" MPs. Examples include MPs applied in the management of Southern Bluefin Tuna (CCSBT, 2011) and the index slope and target MPs described by Geromont and Butterworth (2015a). We also test "generic" MPs that are intended to operate over a wider range of scenarios by attempting account for broad information about stock life history or sustainable exploitation rate (for example, fishing at a fixed fraction of the natural mortality rate). Many of the generic MPs tested in this study are new approaches that have higher data demands than the tuned MPs and rely on recent observations of absolute biomass, i.e. indices that are treated as measures of absolute biomass that were obtained from a fishery-independent survey (e.g. an acoustic trawl or pelagic egg survey).

We describe a reference set of OMs and identify arbitrary, general performance metrics (i.e. summary statistics). The aims of this study are to reveal the performance trade-offs among MPs, identify the core sensitivities of the MPs to their data inputs and the parameters of the OMs, and identify important interactions between MPs and life history/data quality. We evaluate the performance of the MPs in relation to stock dynamics, in particular longevity, temporal variability in productivity, and stock depletion. We also investigate whether generic MPs could provide comparable or better performance to approaches that are currently applied in both data-rich and data-limited settings.

Methods

Generic MPs

In this study, a number of new MPs are described that aim to use recent observations of absolute biomass B, and total annual catches C, to infer surplus production S, and therefore stock level relative to a productive stock size (i.e. Maunder, 2014 and Figure 1). These MPs are appropriate in data-rich situations where a conventional assessment has been used to provide an estimate of the constant of proportionality, q (that scales a relative abundance index to predicted absolute biomass), or alternatively data-moderate settings where a fishery-independent survey could provide an index of abundance that could be treated as a measure of absolute biomass, though with unknown bias.

Seven generic MPs described in this study (Rcontrol, Rcontrol2, Gcontrol, SPmod, SPslope, Fadapt, and DynF) rely on the same calculation of surplus production S:

$$S_y = B_{y+1} - B_y + C_y. \tag{1}$$

The derivative of surplus production with respect to biomass G (i.e. $\mathrm{d}S/\mathrm{d}B$) may be used to move the stock towards a more productive stock size where $G \approx 0$ (Figure 1). Negative G values imply that stock levels are above the most productive stock size; positive G values imply that the stock is below that size, whereas G values close to zero suggest that the stock is close to that size. This concept does not rely on the assumption of a fixed position of the most productive stock size relative to the unfished level and may be able to adapt to temporal shifts in productivity. The degree to which this is possible will depend heavily on the frequency and duration of productivity shifts and time-lags caused by specific life history characteristics such as age-at-maturity, as well as on the quality and frequency of observations of B.

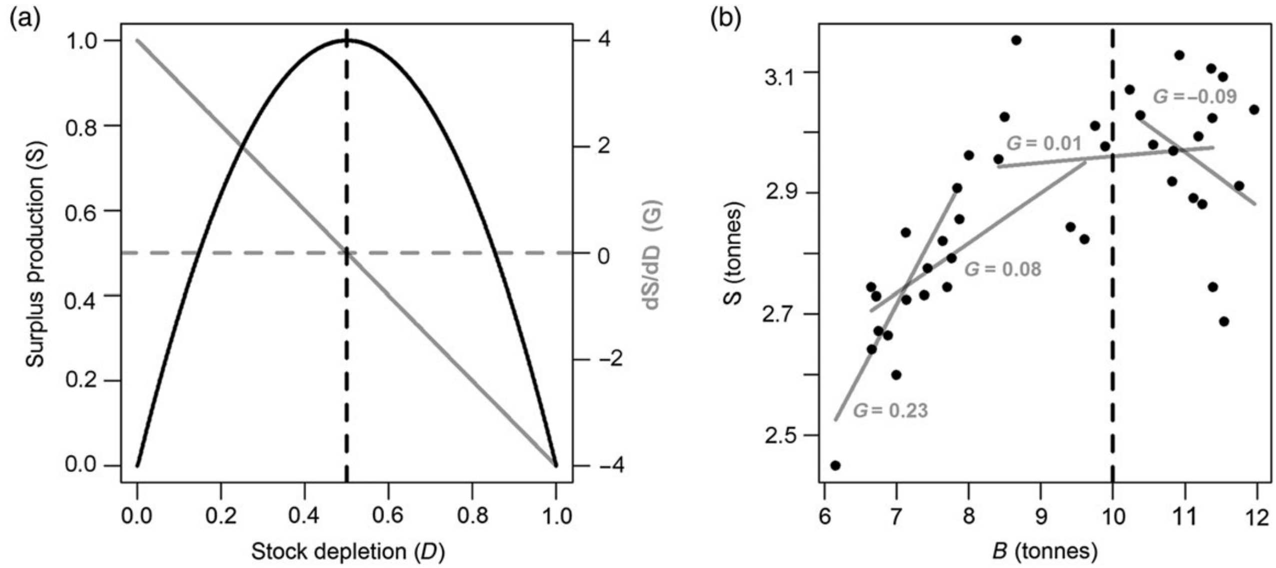

Figure 1. The theoretical model of derivative (G) in surplus production (S) with respect to stock depletion (D, biomass relative to the unfished level) according to the Schaefer production model (a). The vertical dashed line represents the simulated level of biomass at the most productive stock size. Observations of catch and biomass (B) may be used to infer surplus production (S): $S_y = B_y - B_{y-1} + C_{y-1}$ (b). In theory, the derivative of surplus production with respect to biomass can be used to modify management recommendations to move stock levels towards more productive stock sizes where G is close to zero (horizontal dashed line). In (b), estimates of G (grey lines) for four simulated periods are illustrated.

Ten additional generic rules are included that were proposed by Geromont and Butterworth (2015b; CC1, CC4, LstepCC1, LstepCC4, Ltarget1, Ltarget4, Islope1, Islope4, Itarget1, and Itarget4) for use in data-limited fisheries for which annual catch data are available together with either a relative abundance index (e.g. catch-per-unit-effort data or survey) or catch composition data (e.g. length data). Since their data requirements are modest, these MPs may be particularly appropriate for data-limited settings. We tested slightly modified versions of the Itarget1 and Itarget4 MPs in which the target cpue index varied over time rather being fixed at a historical average cpue.

All the generic MPs have parameters that may be tuned to specific case-studies, such as the sensitivity of management updates to changes in G and the number of years of data used to calculate G. However, in this application, we adopt rules with fixed parameter levels that are intended to operate over a range of population and fishing scenarios. In situations where multiple parameterizations have been proposed, we chose to evaluate two versions that span a range of biological precaution. For example, Geromont and Butterworth (2015b) describe four constant catch MPs: CC1, CC2, CC3, and CC4, which set total annual catches (TACs) according to a declining fraction of average historical catches. In this case, we evaluate the most extreme versions, CC1 and CC4 (100 and 70% of average historical catches, respectively). Table 1 contains a summary of all MPs tested in this study; the equations of the generic MPs are presented in Table 2.

Tuned MPs

In many MSE applications, MPs are tuned through simulation testing to achieve the prespecified management objectives. The tuning is carried out over a range of OMs or a reference OM (CCSBT, 2011). In this simulation evaluation, we chose not to re-tune these MPs to new simulated data. Rather we tested the methods with their published parameter values. The tuned MPs of our analysis come from two sources: the MSE for Southern Bluefin Tuna (CCSBT, 2011) and a recent study by Geromont and Butterworth (2015a) who identified MPs for data-rich fisheries.

In this study, we evaluate a simplified version of the second Southern Bluefin Tuna MP (SBT2) that modifies the TAC to reach a predefined target catch level. The published rule derives target catches from an equation with parameters tuned to the specific SBT simulations, and involves filtering of the two MP input indices via a two-stage relative abundance population model which has both trend and target characteristics (CCSBT, 2011). To make the rule operate in this simulation framework, we assumed that the target catch was *MSY*. To simulate imperfect knowledge in *MSY* as the target catch level, we added bias to the true simulated *MSY* level. In this way, we evaluated a more general version of the SBT2 MP.

Reference methods

To frame the performance of the generic and tuned MPs, we included a series of reference MPs that represent conventional stock assessments or methods currently used in the management of data-limited stocks. To represent a data-rich assessment that uses a time-series of catch and effort data, we included a delay-difference assessment model (Deriso, 1980; Schnute, 1985). More complex stock assessments, such as statistical catch-at-age models, were too computationally intensive to be included in this MSE framework. Additionally, it may be argued that detailed stock assessments involve many subjective decisions regarding data processing and model assumptions that cannot be properly replicated in an automated simulation evaluation. The purpose of including the delay-difference assessment approach was not to mimic an integrated age-structured assessment but rather to evaluate the performance of approaches that rely on a long time-series of relative abundance data that are assumed to represent the exploitation history of the stock. We also test a variant of the delay-difference

Table 1. Overview of the MPs and their data requirements.

Type	Name	Description	Data requirements/inputs	References
Generic MPs	Gcontrol	The derivative of surplus production with respect to biomass is used to update the TAC	Recent catch, recent time-series of stock biomass	
	Rcontrol	Similar to Gcontrol MP that includes an estimate for intrinsic rate of increase to better characterize surplus production	Recent catch, recent time-series of stock biomass, growth model, stock depletion, recruitment compensation, natural mortality rate	
	Rcontrol2	Similar to Rcontrol but assumes a quadratic relationship between surplus production and biomass		
	SPslope	Historical trend in biomass is used to update the TAC recommendation using recent inferred surplus production as a reference point	Recent catch, recent time-series of stock biomass	Maunder (2014)
	SPmod	Similar to SPslope but uses only 1 year of biomass and catch to update the TAC		
	DynF	The TAC is set according to fishing rate which varies according to the derivative of surplus production with respect to biomass	Recent catch, recent absolute biomass, natural mortality rate, ratio of F_{MSY} to natural mortality rate	
	Fadapt	The TAC is set according to the variable rate that is modified according to the derivative of surplus production with respect to biomass		
	CC1	The TAC is recent catch levels	Recent catch	Geromont and Butterworth (2015a)
	CC4	The TAC is 70% of recent catch levels		
	LstepCC1	Incremental changes are made to the TAC in relation to recent changes in mean length in catches	Recent catch, recent catch-at-length observations	
	LstepCC4	As LstepCC1 but starting at a lower initial catch level		
	Ltarget1	Incremental changes are made to the TAC to reach a target mean length in catches	Recent catch, recent catch-at-length observations	
	Ltarget4	As Ltarget1 but reference catch level is lower and target mean length in catches is longer		
	Islope1	Similar to GB_slope: incremental changes in TAC are made to maintain a constant relative abundance index	Recent catch, recent relative abundance index	
	Islope4	As Islope1 but starting catch level is lower and changes to the TAC are made less rapidly in response to trajectory in relative abundance		
	Itarget1	Similar to GB_target: the TAC is set according to a reference catch level that is modified to reach a target catch rate	Recent catch, recent relative abundance index	
	Itarget4	As Itarget1 but reference catch level is lower and target catch rate is higher		
Tuned MPs	SBT1	Simple MP for Southern Bluefin Tuna using a target catch level (simulated *MSY*)	Recent catch, recent relative abundance index	CCSBT (2011)
	SBT2	Adaptive MP for Southern Bluefin Tuna that uses target biomass and catch levels (simulated B_{MSY} and *MSY*, respectively)	Recent catch, recent recruitment strength, *MSY*	
	GB_CC	The TAC is constant catch at *MSY* levels	Recent catch, *MSY*	Geromont and Butterworth (2015b)
	GB_slope	Incremental changes in TAC are made to maintain a constant relative abundance index	Recent catch, recent relative abundance index	
	GB_target	Incremental changes in TAC are made to reach a target catch rate (relative abundance level)	Recent catch, recent catch rate data, *MSY*, target catch rate	
Reference methods	DD	A delay-difference stock assessment	Historical catch, historical relative abundance, growth model, natural mortality rate	Carruthers *et al.* (2014)
	DD_4010	As DD but with a 40-10 harvest control rule superimposed		
	DCAC	Depletion-corrected average catch. The TAC is average catches that are downwards adjusted to account for the "windfall catch" that drove the stock down from unfished levels to biomass at *MSY*	Historical catch, current depletion, B_{MSY} relative to unfished biomass, natural mortality rate, ratio of F_{MSY} to natural mortality rate	MacCall (2009)
	Fratio	The TAC is a fixed ratio of natural mortality rate multiplied by an estimate of current absolute stock biomass	Current biomass, natural mortality rate, ratio of F_{MSY} to natural mortality rate	Walters and Martell (2002)

Table 2. The equations of the generic MPs.

MP name	Parameter values	TAC calculation
Gcontrol	$g^U = 0.5$ $g^L = 2$	$\mathrm{TAC}_y = \begin{cases} \mathrm{TAC}_y^{\mathrm{try}} & g^L < \mathrm{TAC}_y^{\mathrm{try}} < g^U \\ g^L\bar{C}_y & \mathrm{TAC}_y^{\mathrm{try}} < g^L \\ g^U\bar{C}_y & g^U < \mathrm{TAC}_y^{\mathrm{try}} \end{cases}, \quad \mathrm{TAC}_y^{\mathrm{try}} = S_y(1 - 2G_y),$ G_y is the slope in surplus production, S with biomass over the last 10 years $S_y = \bar{B}_y - \bar{B}_{y-1} + \bar{C}_{y-1}$ where $\bar{B}$ and $\bar{C}$ are the biomass and catch predicted by a first-degree Loess smoother with smoothing parameter alpha $= 0.75$, fitted to observations of absolute biomass and catch over the last 10 years
Rcontrol		As Gcontrol except: $G_y = r(1-2D_y)$ where r is the demographically derived prior for the intrinsic rate of increase (McAllister *et al.*, 2001) and D is stock depletion (B_y/B_0)
Rcontrol2		As Rcontrol except: $G_y = \left(\sum_{t=y-9}^{y} S_t/\bar{B}_t - r\right)\left(\sum_{t=y-9}^{y} \bar{B}_t\right)/\left(\sum_{t=y-9}^{y} \bar{B}_t^2\right)$
SPslope		$\mathrm{TAC}_y = \begin{cases} [1 - (\bar{B}_{y-4} - \bar{B}_y)/\bar{B}_{y-4}]C_y^{\mathrm{ave}} & \Delta^B < 9/10 \\ (9/10)S_{y-1} & \Delta^B > 11/10 \\ \mathrm{TAC}_{y-1} & 9/10 < \Delta^B < 11/10 \end{cases}$ $\Delta^B = \bar{B}_y/\bar{B}_{y-4},\ S_y = \bar{B}_y - \bar{B}_{y-1} + \bar{C}_{y-1},\ C_y^{\mathrm{ave}} = 1/4\sum_{t=y-3}^{y} C_t,$ $\mathrm{TAC}_1 = C_1$
SPmod		$\mathrm{TAC}_y = \begin{cases} 4C_y/5 & \Delta^B < 4/5 \\ 6S_{y-1}/5 & \Delta^B > 6/5 \\ \mathrm{TAC}_{y-1} & 4/5 < \Delta^B < 6/5 \end{cases}, \quad \Delta^B = B_y/B_{y-1},\ S_y = B_y - B_{y-1} + C_{y-1}$ $\mathrm{TAC}_1 = C_1$
DynF	$F^L = F_{MSY}/2$ $F^U = 2{\cdot}F_{MSY}$ $F_{MSY} = M(F_{MSY}/M)$	$\mathrm{TAC}_y = \bar{F}_y\bar{B}_y,\ \bar{F}_y = \begin{cases} F_y^{\mathrm{try}} & F^L < F_y^{\mathrm{try}} < F^U \\ F^L & F_y^{\mathrm{try}} < F^L \\ F^U & F^U < F_y^{\mathrm{try}} \end{cases}, \quad F_y^{\mathrm{try}} = F_{y-1}e^{-2G_y}$ As Gcontrol, G_y is the derivative of S with respect to biomass over the last 7 years. $\bar{B}_y$ is calculated as in Gcontrol.
Fadapt		$\mathrm{TAC}_y = \bar{F}_y\bar{B}_y,\ \bar{F}_y = F^L + \mathrm{logit}^{-1}(W_y - G_y)(F^U - F^L)$ $W_y = \begin{cases} \mathrm{logit}(F_y^{\mathrm{ave}} - F^L/F^U - F^L) & F^L < F_y^{\mathrm{ave}} < F^U \\ -2 & F_y^{\mathrm{ave}} < F^L \\ 2 & F^U < F_y^{\mathrm{ave}} \end{cases}$ As Gcontrol, G_y is the derivative of S with respect to biomass over the last 7 years. F_y^{ave} is the average harvest rate (C/B) over the last 7 years.
CC1 CC4	$x = 1$ $x = 0.7$	$\mathrm{TAC}_{y+1} = x\bar{C}_y,\quad \bar{C}_y = 1/5\sum_{t=y*-4}^{y*} C_t$
LstepCC1 LstepCC4	$x = 1$ $x = 0.7$	$\mathrm{TAC}_{y+1} = \begin{cases} \mathrm{TAC}_y + 1/20\bar{C}_{y*} & 1.05 < L_y^{\mathrm{rec}}/L_y^{\mathrm{ave}} \\ \mathrm{TAC}_y & 0.98 < L_y^{\mathrm{rec}}/L_y^{\mathrm{ave}} < 1.05 \\ \mathrm{TAC}_y - 1/20\bar{C}_{y*} & 0.96 < L_y^{\mathrm{rec}}/L_y^{\mathrm{ave}} < 0.98 \\ \mathrm{TAC}_y - 1/10\bar{C}_{y*} & L_y^{\mathrm{rec}}/L_y^{\mathrm{ave}} < 0.96 \end{cases}$ $\mathrm{TAC}_{y*} = x\bar{C}_y\ \bar{C}_y = 1/5\sum_{t=y*-4}^{y*} C_t$
Ltarget1 Ltarget4	$x = 1$ $v = 1.05$ $x = 0.8$ $v = 1.15$	$\mathrm{TAC}_{y+1} = \begin{cases} \mathrm{TAC}_{y*}/2[1 + (L_y^{\mathrm{rec}} - L^0/L^{\mathrm{targ}} - L^0)] & L_y^{\mathrm{rec}} \geq L^0 \\ \mathrm{TAC}_{y*}/2[L_y^{\mathrm{rec}}/L^0]^2 & L_y^{\mathrm{rec}} < L^0 \end{cases}, \quad L^0 = \frac{9}{10}L_y^{\mathrm{ave}}$ $L_y^{\mathrm{rec}} = 1/5\sum_{t=y-4}^{y}\sum_{i=1}^{n_t} L_{t,i}^{\mathrm{obs}}/n_t,\ L_y^{\mathrm{ave}} = 1/10\sum_{t=y-9}^{y}\sum_{i=1}^{n_t} L_{t,i}^{\mathrm{obs}}/n_t, L^{\mathrm{targ}} = vL_y^{\mathrm{ave}}$ n_t is the number of length observations at time t

Continued

Table 2. *Continued*

MP name	Parameter values	TAC calculation
Islope1	$x = 0.8$ $\lambda = 0.4$	$\mathrm{TAC}_{y*} = x\bar{C}_y,\ \bar{C}_y = 1/5\sum_{t=y*-4}^{y*} C_t,\ \mathrm{TAC}_{y+1} = \mathrm{TAC}_y(1+\lambda s_y),$
Islope4	$x = 0.6$ $\lambda = 0.2$	where s is the derivative of log cpue with respect to time over the last 5 years
Itarget1	$x = 1$ $\nu = 1.5$	$\mathrm{TAC}_{y+1} = \begin{cases} \mathrm{TAC}_{y*}/2[1 + (I_y^{\mathrm{rec}} - I^0/I_y^{\mathrm{targ}} - I^0)] & I_y^{\mathrm{rec}} \geq I^0 \\ \mathrm{TAC}_{y*}/2[I_y^{\mathrm{rec}}/I^0]^2 & I_y^{\mathrm{rec}} < I^0 \end{cases},$
Itarget4	$x = 0.7$ $\nu = 2.5$	$\mathrm{TAC}_{y*} = x\bar{C}_y\ ,\ \bar{C}_y = 1/5\frac{1}{5}\sum_{t=y*-4}^{y*} C_t,$ $I_y^{\mathrm{rec}} = 1/5\sum_{t=y-4}^{y} I_t,\ I^{\mathrm{ave}} = 1/10\sum_{t=y-9}^{y} I_t,\ I^0 = 4/5I_y^{\mathrm{ave}},\ I_y^{\mathrm{targ}} = \nu\ I_y^{\mathrm{ave}}$

TAC, a total allowable catch recommendation; *C*, a total annual catch observation; *B*, an observation of absolute biomass; *D*, an estimate of current stock depletion (biomass relative to unfished); B_0, unfished biomass; *I*, an annual relative abundance index or catch rate observation; *M* and F_{MSY}/M are imperfectly known simulated values of natural mortality rate and fishing rate at maximum sustainable yield relative to the natural mortality rate. y^* refers to the first year in which the MP was implemented, n_t is the number of length observations in a particular year *t*, L^{obs} is the observed length of a fish caught and *MSY* is maximum sustainable yield (although this is not known perfectly).

model that is combined with the "40-10" harvest control rule. Under this rule, the stock is not fished when stock size is below 10% unfished biomass and fished at F_{MSY} above 40% of unfished biomass. Between 10 and 40% unfished levels, exploitation rate follows a linear increase from 0 to 100% F_{MSY}.

DCAC is the first of two data-limited methods we included in the performance evaluation (MacCall, 2009). The DCAC provides an estimate of "sustainable catch" based on an estimate of average annual catch and four inputs: depletion (B_{cur}/B_0), the ratio of F_{MSY}/M, M and B_{MSY}/B_0. This method aims to calculate a sustainable catch which takes account of the removal of the "windfall harvest" of less productive biomass that may have occurred as the stock became depleted (the equations are included in Supplementary Appendix A). DCAC is currently used by the Pacific Fishery Management Council to set catch-limits for some data-limited stocks (PFMC, 2010).

The second data-limited method, Fratio, simply aims to fish at a constant exploitation rate that is a fixed fraction of the natural mortality rate (Gulland, 1971; Walters and Martell, 2002). The North Pacific Fishery Management Council uses an Fratio method for managing stock complexes in situations where stock assessments are not available (NPFMC, 2012, 2013). Under the Fratio method, a catch limit is simply the product of the estimate of natural mortality rate M, the ratio of F_{MSY}/M, and a current observation of absolute biomass. Since absolute biomass indices are also required for some of the generic MPs, the simpler Fratio MP provides an useful comparison.

Operating model structure and simulation design

Simulation testing was carried out in the R statistical environment (R Core Team, 2015) using the R-package Data-Limited Methods toolkit (DLMtool v1.35; Carruthers, 2015). The package is freely available and includes all the operating models and MPs evaluated in this study (computer code for reproducing our methods is available online at https://github.com/tcarruth/Carruthers-et-at-2015-MP-MSE).

The operating model of DLMtool (v1.35) is an age-structured, spatial population dynamics model of identical structure to that of Carruthers *et al.* (2014) (a full description of the operating model is given in Supplementary Appendix A). A simple two-area spatial structure is assumed that generates a movement model according to two parameters that are user-defined: the probability of individuals staying in area 1 between model time-steps and the fraction of the population inhabiting area 1 under unfished conditions (i.e. the fraction of habitat in area 1). This structure provides for simple simulation of phenomenon such as spatial fishing restrictions and fishery targeting. However, in this application, values for spatial parameters are consistent with a fully mixed stock and the simulated spatial dynamics of our operating models provided the same results as a single-area operating model.

We constructed operating models using a factorial design encompassing 36 sets of operating model assumptions. The four factors were (i) life history with three levels, (ii) temporal autocorrelation in recruitment with two levels, (iii) starting stock depletion with two levels, and (iv) data quality with three levels (Table 3). For each of the 36 combinations, we carried out 500 simulations for each MP. Each simulation was then projected forward for 40 years adopting the TAC recommendations of each MP. We did not simulate implementation error and assumed that prescribed catches would be taken exactly up to a maximum instantaneous fishing mortality rate of 90%. The MPs were rerun and the TAC updated every 3 years to approximate an assessment cycle.

Three population life history types of varying longevity were simulated based on the outputs of data-rich stock assessments for Pacific herring (DFO, 2012), the eastern stock of Atlantic bluefin tuna (ICCAT, 2012), and Pacific canary rockfish (Wallace and Cope, 2011). The depletion for each stock was not set to the value from the assessment, because the intention was to characterize broad life history types rather than the status of particular stocks.

Previous simulation evaluations have indicated that most of the variability in the performance of MPs occurs in the range of stock depleted below B_{MSY} (Carruthers *et al.*, 2014), which is arguably the most important population level for evaluating performance, at least from a general policy perspective. We simulated two ranges of initial depletion: a rebuilding scenario in which the spawning stock is between 2.5 and 15% of unfished levels (less than $B_{MSY}/2$) and an overexploited scenario in which spawning biomass is between 15 and 35% of unfished levels (between $B_{MSY}/2$ and B_{MSY}).

Autocorrelation in recruitment was simulated to evaluate the performance of the MPs in situations where stock productivity varies over time [an AR1 process, Appendix Equation (B2)]. This may not fully reflect step-changes in recruitment that have been

Table 3. Overview of simulation design.

MSE attribute **Life history**	**Symbol**	**Case 1** **Herring**	**Case 2** **Bluefin tuna**	**Case 3** **Rockfish**
Maximum age	n_a	10	32	64
Natural mortality rate (y^{-1})	M	0.28 – 0.38	0.2 – 0.28	0.04 – 0.08
Recruitment compensation (steepness)	H	0.4 – 0.6	0.6 – 0.9	0.35 – 0.72
Recruitment deviations lognormal SD	σ_R	0.2 – 0.4	0.1 – 0.3	0.2 – 0.5
Initial stock depletion		Rebuilding	Overexploited	
Biomiass relative to unfished	D	2.5 – 15%	15 – 35%	
Non-stationarity in recruitment		Low	High	
Autocorrelation	V	0 – 30%	60 – 90%	
Data quality		Perfect info	Data-rich	Data-poor
Data inputs				
Bias in biomass relative to unfished (lognormal SD)	Υ_D	None	0.2	0.5
Bias in annual catches (lognormal SD)	Υ_C	None	0.2	0.5
Observation error in annual catches (lognormal SD)	σ_C	None	0.2 – 0.4	0.3 – 0.6
Observation error in relative abundance index (lognormal SD)	σ_I	None	0.1 – 0.3	0.2 – 0.6
Observation error in recruitment (lognormal SD)	$\sigma_{R_{obs}}$	None	0.05 – 0.1	0.1 – 0.3
Hyperstablity/hyperdepletion in index	B	None	2/3 – 3/2	1/3 – 3
Bias in absolute biomass	Υ_B	None	1/3 – 3	1/5 – 5
Observation error in absolute biomass (lognormal SD)	σ_B	None	0.1 – 0.5	0.5 – 1
Bias in natural mortality rate (lognormal SD)	Υ_M	None	0.25	0.5
Bias in von Bertalanffy κ parameter (lognormal SD)	Υ_κ	None	0.05	0.1
Bias in B_{MSY} relative to unfished (lognormal SD)	$\Upsilon_{B_{peak}}$	None	0.1	0.2
Bias in *age at maturity* (lognormal SD)	Υ_{Am}	None	0.1	0.2
Number of annual age/length observations	n_{CAA}	10 000 – 15 000	200 – 500	50 – 100
Other control rule inputs				
Bias in ratio of F_{MSY} to M (lognormal SD)	$\Upsilon_{F_{MSY}\_M}$	None	0.25	0.5
Bias in target cpue (B_{MSY}) (lognormal SD)	Υ_{cpue}	None	0.3	1
Bias in target catch (*MSY*) (lognormal SD)	Υ_{MSY}	None	0.2	0.3

Parameter ranges represent the lower and upper bounds of a uniform random variable. SD refers to the standard deviation. For biases, the SD defines a range of potential multiplicative biases that may be sampled among simulations with mean of 1 (no bias on average). For example, given an SD of 20% three sampled biases could be 0.84, 1.05, and 1.34, representing a consistent moderate negative bias for all observations of simulation 1, small positive bias for simulation 2, and a stronger positive bias for simulation 3.

observed in some fishery settings (Vert-pre *et al.*, 2013). However, simulating recruitment autocorrelation does not invalidate the derivation and subsequent use of *MSY* reference points which are basic as inputs to several MPs and also in assessing performance of the MPs.

Bias and imprecision in the knowledge of the simulated system were generated for all the inputs to the MPs (e.g. natural mortality rate and observations of absolute biomass; see Appendix Table A1 that includes a summary of the observation error model). We simulated three data quality levels corresponding to perfect information, data-rich, and data-limited scenarios. The perfect information simulations assume no error in knowledge of inputs to MPs, essentially removing the observation model and revealing performance with respect to operating model parameters, such as age-at-maturity and recruitment compensation, and the specific levels of process error (recruitment variance in this case). Data-rich simulations assume that inputs to MPs are known imperfectly, and may be subject to moderate bias and imprecision. Among simulations, consistent bias in annual catch observations is sampled from a lognormal distribution with mean 1 and a standard deviation, Υ_C, of 20% (Table 3). For example, a value of 0.82 could be drawn for the first simulation and be used as consistent 18% downwards bias for all observed catches in that time-series. In addition to this bias, we superimpose imprecision that is lognormal error in annual catch observations per simulation σ_C, allowing us to separate the effect on performance of both bias and imprecision in inputs. The data-limited simulations included higher levels of bias and imprecision in the inputs to the MPs to simulate the poorer quality of data. Biases were generated from lognormal distributions for all inputs except observations of absolute biomass that were assumed to be log-uniform. This was intended to reflect the greater probability of more extreme biases in these data. Note that biases are also considered for the target cpue and biomass (both intended to reflect *MSY* levels); this is because these quantities are inputs to some of the control rules (Itarget1, Itarget4, and GB_target—see Tables 2 and 4).

Performance criteria

Performance was summarized by five metrics: yield, average annual variability in yield (AAVY), fishing mortality rate, spawning-stock biomass (SSB), and SSB over the long term. For each of these metrics, a reference level was defined in addition to an acceptable rate of obtaining the reference level (Table 5). For example for fishing mortality rate, a statistic *F* was identified that reflects the fraction of simulation years in which fishing mortality rate was less than a reference level of 125% F_{MSY}. An acceptable score for *F* was considered to be 60%: performance was considered acceptable if an MP had greater than a 60% probability of fishing at a rate under 125% F_{MSY}. These performance criteria were established on an *ad hoc* basis and were intended to represent the performance of a hypothetical MSE.

In most fisheries, it is not desirable for catch-limits to fluctuate strongly between years. To address this, we include the performance metric AAVY. AAVY is the mean difference in the yield of adjacent

Table 4. The equations of the tuned MPs.

MP name	TAC calculation
SBT1	$\mathrm{TAC}_y = \mathrm{TAC}_{y-1} \cdot \begin{cases} 1 + 2X_y/3 & X_y < 0 \\ 1 + 3X_y & X_y \geq 0 \end{cases}$
	where X_y is the derivative of log cpue with respect to time over the last 10 years
SBT2	$\mathrm{TAC}_y = 1/2\mathrm{TAC}_{y-1} + \delta_{MSY}, \delta = \begin{cases} \Delta^{7/4} & \Delta < 1 \\ \Delta^{1/4} & \Delta > 1 \end{cases}, \Delta = R^{\mathrm{ave}}/R^{\mathrm{hist}}$
	$R_y^{\mathrm{ave}} = 1/5 \sum_{t=y-4}^{y} R_{\mathrm{obs}_t},\ R_y^{\mathrm{hist}} = 1/10 \sum_{t=y-9}^{y} R_{\mathrm{obs}_t}$
GB_CC	$\mathrm{TAC}_y = \begin{cases} 4C^{\mathrm{ave}}/5 & MSY < 4C^{\mathrm{ave}}/5 \\ C^{\mathrm{ave}} & 4C^{\mathrm{ave}}/5 < MSY < 6C^{\mathrm{ave}}/5 \\ 6C^{\mathrm{ave}}/5 & 6C^{\mathrm{ave}}/5 < MSY \end{cases}$
	where C^{ave} is the mean historical annual catch
GB_slope	$\mathrm{TAC}_y = \begin{cases} 4C^{\mathrm{ave}}/5 & \mathrm{TAC}_y^{\mathrm{try}} < 4C^{\mathrm{ave}}/5 \\ \mathrm{TAC}_y^{\mathrm{try}} & 4C^{\mathrm{ave}}/5 < \mathrm{TAC}_y^{\mathrm{try}} < 6C^{\mathrm{ave}}/5 \\ 6C^{\mathrm{ave}}/5 & 6C^{\mathrm{ave}}/5 < \mathrm{TAC}_y^{\mathrm{try}} \end{cases},\ \mathrm{TAC}_y^{\mathrm{try}} = C^{\mathrm{ave}}(1 + G_y)$
	where C^{ave} is the mean historical annual catch and G is the derivative of log cpue with respect to time over the last 5 years.
GB_target	$\mathrm{TAC}_y = \begin{cases} 4C^{\mathrm{ave}}/5 & \mathrm{TAC}_y^{\mathrm{try}} < 4C^{\mathrm{ave}}/5 \\ \mathrm{TAC}_y^{\mathrm{try}} & 4C^{\mathrm{ave}}/5 < \mathrm{TAC}_y^{\mathrm{try}} < 6C^{\mathrm{ave}}/5 \\ 6C^{\mathrm{ave}}/5 & 6C^{\mathrm{ave}}/5 < \mathrm{TAC}_y^{\mathrm{try}} \end{cases}$
	$\mathrm{TAC}_y^{\mathrm{try}} = \begin{cases} MSY(1/2 + I_y^{\mathrm{rec}} - I^0/2(I^{\mathrm{targ}} - I^0)) & I^0 \leq I_y^{\mathrm{rec}} \\ MSY/2(I_y^{\mathrm{rec}}/I^0)^2 & I_y^{\mathrm{rec}} < I^0 \end{cases},\ I^0 = 1/25 \sum_{t=y*-4}^{y*} I_t,\ I_y^{\mathrm{rec}} = 1/4 \sum_{t=y-3}^{y} I_t\ I_y^{\mathrm{rec}} = \frac{1}{4} \sum_{t=y-3}^{y} I_t$
	I^{targ} is the abundance index that corresponds to the true simulated B_{MSY}, expressed relative to unfished biomass for which the value is not known perfectly

TAC, a total annual catch recommendation; *C*, a total annual catch observation; *B*, an observation of absolute biomass; *I*, an annual relative abundance index or catch rate (cpue) observation; *R*obs, an estimate of recruitment strength; *MSY*, maximum sustainable yield (though this is not known perfectly).

Table 5. Performance metrics of this simulation evaluation.

Performance metric	Reference level	Target (%)
SSB	SSB is above 50% *MSY* levels evaluated over all projected years	>60
Long-term SSB	SSB is above 50% *MSY* levels evaluated over the last 10 projected years	>60
Fishing mortality rate (*F*)	Exploitation rate is lower than 125% F_{MSY} evaluated over all projected years	>60
Yield	Mean catches are greater than 50% *MSY* evaluated over all projected years	>60
AAVY	AAVY is <30%	>70

Targets are probabilities of achieving a reference level. Note that MP performance is considered acceptable if all targets are met in combination.

projected years (starting from the last historical year) divided by the mean yield over the same period.

$$\mathrm{AAVY} = \frac{(n_p + 1)\sum_{y=n_h}^{n_h+n_p-1} |C_{y+1} - C_y|}{n_p \sum_{y=n_h}^{n_h+n_p} C_y}, \tag{2}$$

where n_p is the number of projected years, n_h is the number of historical years, and C_y is the true simulated catch in year *y*.

Individually, the performance metric targets are not overly stringent. However, they are intended to be interpreted in combination: a suitable MP is one that can satisfy the acceptable rate for all metrics combined (i.e. greater than a 60% chance of obtaining SSB over 50% *MSY* levels and yields greater than 50% *MSY* levels etc., Table 5).

In specific management contexts, a set of performance metrics for an MP might include constraints that are more limiting than those selected for our analysis (Table 5). For instance, acceptability criteria recommended as best practice for Australian Fisheries Management Authority specify a 95% or greater probability of keeping SSB above the limit reference points in simulations over a 20-year period (Sainsbury, 2008).

Evaluating MP robustness and quantifying the value of information

As described above in Section 2.4, 500 simulations were undertaken for each MP in which a range of operating model parameters were sampled and output performance metrics were calculated. In the simulation, each independent variable (e.g. natural mortality rate *M* and observation error in catches σ_C) was sampled from a range of values that were considered to be credible *a priori* (Table 3 and Supplementary Table A1). To evaluate the robustness of the MPs to the quality of their various inputs, the input ranges were divided into a reference region and two contingency regions that represent more extreme values and are not contiguous with the reference region (Table 6). For biases (e.g. bias in estimates of natural mortality rate), the reference region was selected that represents relatively unbiased inputs and contingency regions represented high and low biases in inputs. For parameters controlling imprecision (e.g. the standard deviation of observation error in annual catches), the reference region represents intermediate precision

Table 6. Specification of robustness tests.

Parameter controlling data quality	Symbol	Ranges for observation error level/bias		
		Reference	Low/−	High/+
Bias in biomass relative to unfished (lognormal SD)	D	3/4–4/3	1/3–1/2	2–3
Bias in annual catches (lognormal SD)	Υ_C	3/4–4/3	1/2–2/3	3/2–2
Observation error in annual catches (lognormal SD)	σ_C	0.35–0.45	0.2–0.3	0.5–0.6
Observation error in relative abundance index (lognormal SD)	σ_I	0.35–0.45	0.2–0.3	0.5–0.6
Hyperstability/hyperdepletion in index	B	3/4–4/3	1/2–2/3	3/2–2
Bias in target cpue (B_{MSY}) (lognormal SD)	Υ_{cpue}	3/4–4/3	1/2–2/3	3/2–2
Bias in target catch (MSY) (lognormal SD)	Υ_{MSY}	3/4–4/3	1/2–2/3	3/2–2
Bias in absolute biomass	Υ_B	3/4–4/3	1/4–1/2	2–4
Observation error in absolute biomass (lognormal SD)	σ_B	0.6–0.7	0.5–0.6	0.7–0.9
Bias in ratio of F_{MSY} to M (lognormal SD)	$\Upsilon_{F_{MSY}\_M}$	3/4–4/3	1/2–2/3	3/2–2
Bias in natural mortality rate (lognormal SD)	Υ_M	3/4–4/3	3/4–4/3	3/2–2

Reference regions represent a range of simulations where data were unbiased or of intermediate precision. Two contingency regions were identified that represent simulations of negative (−) and positive (+) bias or low and high precision. The robustness of MPs was evaluated by comparing their performance among these simulated regions.

and contingency regions represent simulations of low and high precision. To evaluate robustness, we compared mean yield and stock depletion at the end of the time-series across the reference and contingency regions. The sensitivity of yield was used as a basis for quantifying the value of less biased and more precise data. (The performance metrics are assessed for combinations of all these bias and precision factors together. Some metrics may be affected in different directions by lower and higher biases, so that composite results may include some cancellation effects. A more detailed analysis of this is, however, beyond the scope of this investigation.)

Results

Rebuilding scenario

For simulations starting below 50% B_{MSY} levels, the probability of recovering to B_{MSY} levels differed widely among the MPs and was lowest over a 40-year projection for the rockfish life history (Figures 2–4). In general, the choice of MP and life history type more strongly determined rebuilding than data quality or autocorrelation in recruitment. Despite living longer than herring (31 years in contrast to 10 years), the bluefin tuna life history type generally recovered faster due to the higher simulated recruitment compensation (steepness values in the range of 0.6–0.9 in contrast to 0.3–0.6 for herring).

The MPs CC1, DCAC, GB_CC, GB_slope, GB_target, Rcontrol2, SBT1, SBT2, SPmod, and SPslope were unlikely to rebuild stocks for the herring, bluefin tuna, and rockfish life histories (Figures 2–4). In contrast, the MPs DD, DD4010, DynF, Islope1, Islope4, Itarget1, Itarget4, LstepCC4, and Ltarget4 were the most likely to lead to stock rebuilding for these life history types. Usually, simulating better quality data quality improved the probability of reaching B_{MSY}. Exceptions include the Fadapt and CC1 rules. For Fadapt, less precise data generated negatively biased estimates of surplus production and therefore reduced the TAC recommendation of this MP and increased the probability of rebuilding. The CC1 rule, which sets the TAC to mean historical catches, generally led to a low probability of rebuilding. Often rebuilding was similar between ‘perfect information’ and ‘data-rich’ data qualities with a much larger difference in rebuilding performance when ‘data-poor’ quality was simulated. For example in the herring and bluefin tuna simulations, the delay-difference MPs, DD and DD4010, were ∼40% less likely to rebuild over 15 years given poor quality data.

Figure 5 illustrates the trade-off between the probability of achieving more than half MSY catches over the last 10 projected years against the probability of rebuilding SSB to above the MSY level. The MPs that achieved the greatest probability of rebuilding did so by underexploitation. At the other extreme, MPs that led to overexploitation led to stock declines and therefore both low probability of rebuilding and low yield over the final 10 projected years (since the simulated stocks had not recovered, e.g. SPmod and SBT2). A handful of MPs offered a balance of trade-offs in which TACs were high enough to allow for both rebuilding and long-term yields (e.g. Fadapt, DynF, DD, DD4010, and Fratio). These results suggest that over the course of a 40-year rebuilding window, only relatively modest yields can be expected: the best-performing MPs could only deliver between 60 and 80% probability of returning yields more than half MSY (Figure 5) and these rebuilt stocks are achieved in <70% of simulations (Figures 2–4).

Overexploited scenario

When starting from a stock depletion of between 50 and 100% B_{MSY}, many MPs satisfied the SSB and fishing mortality rate performance targets but failed to meet the yield target, for example CC4, LstepCC4, Islope4, and Itarget4 (Figures 6–8). The MPs, Islope1, DynF, Fratio, and DCAC on the other hand, satisfied all targets for most of the operating models. Perhaps not surprisingly, the trade-off which was most apparent was between the yield and fishing mortality rate performance metrics, with only a handful of methods able to satisfy both simultaneously (e.g. DD, DCAC, Islope1, and Fratio). Some MPs which performed among the best in the overexploited scenario, such as DCAC, performed among the worst in the rebuilding scenario, which underlines the critical influence of starting depletion on the choice of MP.

For any given MP, the general pattern among performance metrics (e.g. yield vs. SSB) was similar regardless of life history type (Figures 6–8). However, the magnitude in performance differed enough among life history types to affect the choice of MP. For example, the Islope1 MP met the SSB and F performance targets for most operating models for herring and bluefin tuna (Figures 6 and 7, respectively) but not for rockfish (Figure 8). DCAC, on the other hand, performed well for rockfish and bluefin tuna but less well for herring (though note that DCAC was not developed with the intention that it be applied to short-lived species). Similarly, the Fadapt MP failed the F performance criterion

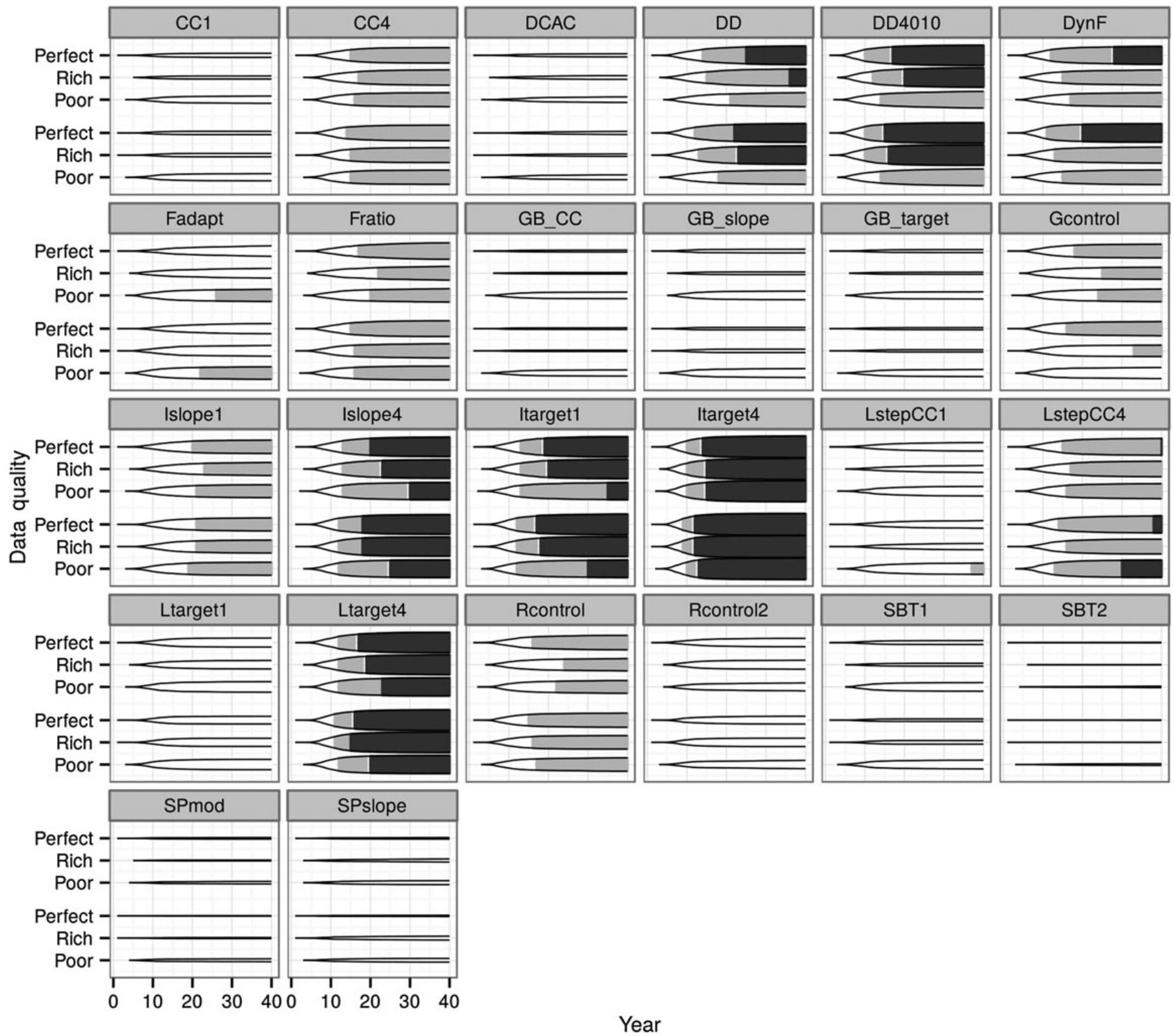

Figure 2. Duration of recovery of spawning biomass to *MSY* levels for the herring life history. The thickness of the horizontal bars represents the cumulative frequency of simulations that have recovered to B_{MSY} levels. The bars are shaded white until 50% of simulations have recovered to B_{MSY} levels, then shaded grey until 75% of simulations have recovered to B_{MSY} levels, after which they are shaded dark grey. The top three bars in each panel represent high autocorrelation in recruitment, and the bottom three bars represent low autocorrelation in recruitment.

by a much larger margin for a greater number of operating models for the rockfish life history. In almost all instances, the simple MPs achieved the AAVY target (greater than a 60% chance of AAVY <30%), making this the least discriminatory performance metric. This result lends support to the findings of Geromont and Butterworth (2015a) that MPs can be better at stabilizing TAC compared with stock assessments (for example, all simple MPs outperformed the DD and DD4010 MPs for the bluefin tuna life history). It is possible, however, that a potential benefit of achieving less variable yields under MPs turns out to be a trade-off of lower yields (or of higher biological risks) compared with management under traditional stock assessment paradigm.

Value of information and sensitivity analysis

For most MPs, long-term yield (average yield over the final 10 years of the projection) was most strongly affected by bias in the observation of annual catches. When observed catches were between half and two-thirds of true values, long-term yields were reduced between 50 and 90% for MPs that require these data (e.g. Gcontrol, Itarget1, Itarget4, Ltarget 1, and Ltarget4). These reductions in yield occurred due to underexploitation. Positive bias in catches of between 50 and 100% led to similar declines in yield for most MPs (between 50 and 95%). However, unlike negative bias in catches, these reductions in yield occurred due to chronic overexploitation and stock declines. When catches were inflated to between 1.5 and 2 times their true value, stock biomass was on average between 50 and 95% lower at the end of the projection for all MPs using these data.

The generic MPs were sensitive to other observation model variables that impact calculations of surplus production. For example, DynF and Fadapt were most sensitive to overestimation of biomass and provided on average 42 and 55% less yield, respectively. Although higher levels of bias were simulated for observations of absolute biomass (Table 6), methods using these data such as DynF and Fadapt provided less pronounced stock declines (additional

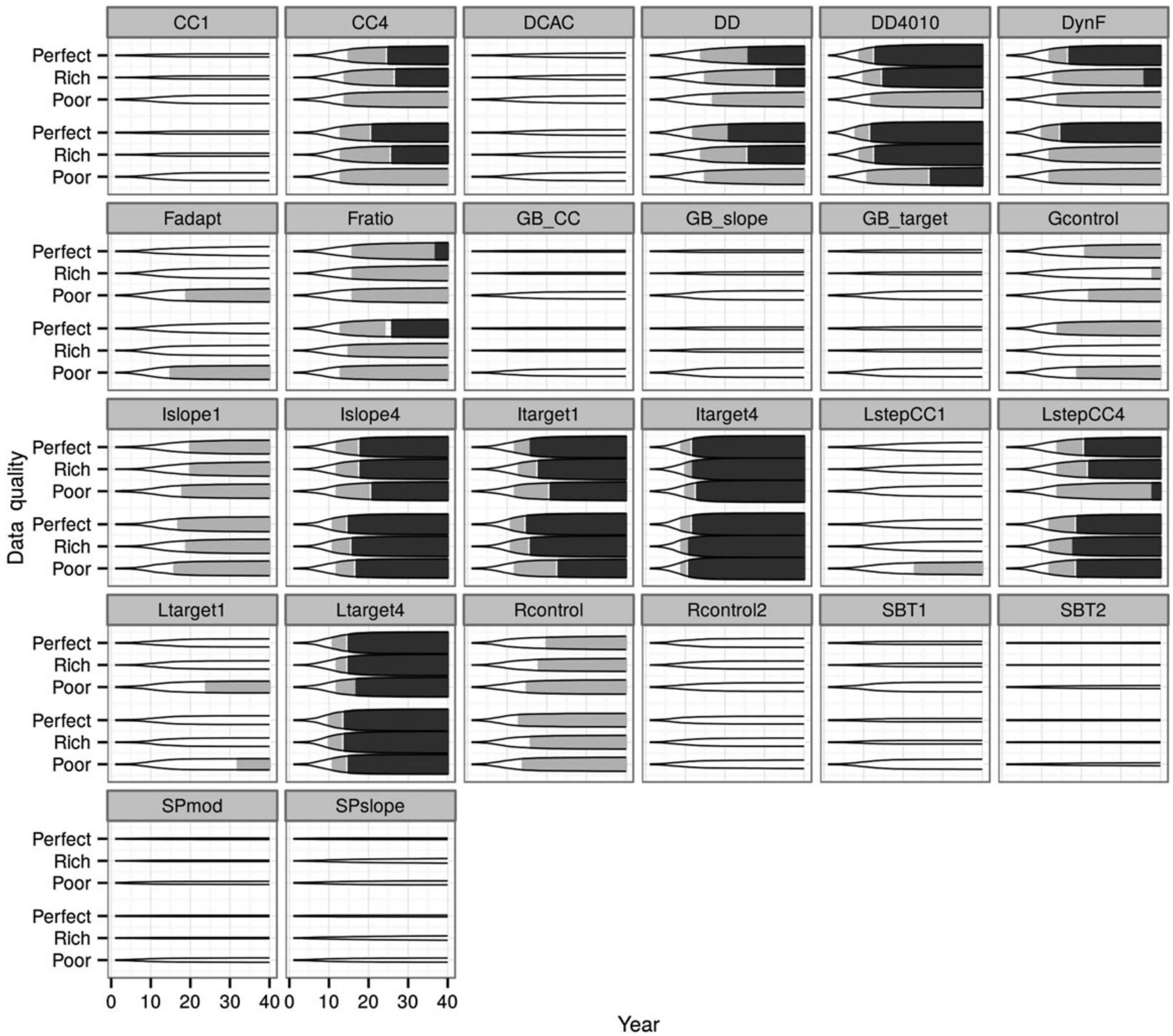

Figure 3. As Figure 2 but for the bluefin tuna life history.

stock declines between 50 and 80%). In previous simulation evaluations, imprecision in catches and abundance indices was indicated to have only a relatively small impact on the performance of data-limited assessment methods (Carruthers *et al.*, 2014). However, the generic MPs were sensitive to high imprecision in catches and the relative abundance index arising from the calculation of surplus production Equation (1). For example, Fadapt and DynF provided 22 and 36% more yield on average given imprecise catch observations. This is because these MPs led to underexploitation in most simulations. Increasing imprecision in catches led to the occasional simulation of strongly positively biased catches which countered the tendency for underexploitation.

Discussion

The simulations indicate that the absolute performance of MPs can vary widely among life history types. Performance rankings, however, were relatively constant and the five best- and worst-performing MPs remained largely the same across life history types. This apparent the lack of interaction between MP and life history type suggests that future analyses could focus on a much smaller subset of MPs. The relative performance of MPs was strongly affected by the initial level of stock depletion. For example, while DCAC performed very well for stocks starting above 50% B_{MSY}, the same MP was unlikely to rebuild stocks starting from below 50% B_{MSY}. This requirement for accurate information regarding depletion is problematic for data-limited scenarios since formally such information is not available by definition.

We found no evidence of a substantial effect of autocorrelation in recruitment on the absolute performance of the MPs. This is perhaps surprising since those MPs rely to varying degrees on assumptions of stationary stock productivity. For example, stationary productivity assumptions are central to reference MPs such as the delay-difference assessments DD and DD_4010. It is likely, however, that the AR1 process used here, even with high autocorrelation, is not appropriate for mimicking regime shifts that are characterized by abrupt then lasting shifts in productivity (e.g. Vert-pre *et al.*, 2013). In future research, this could be addressed by using empirical recruitment data to test MPs.

The best-performing data-limited MPs were generic (e.g. DynF and Fadapt) or reference MPs (e.g. Fratio) that require inputs for

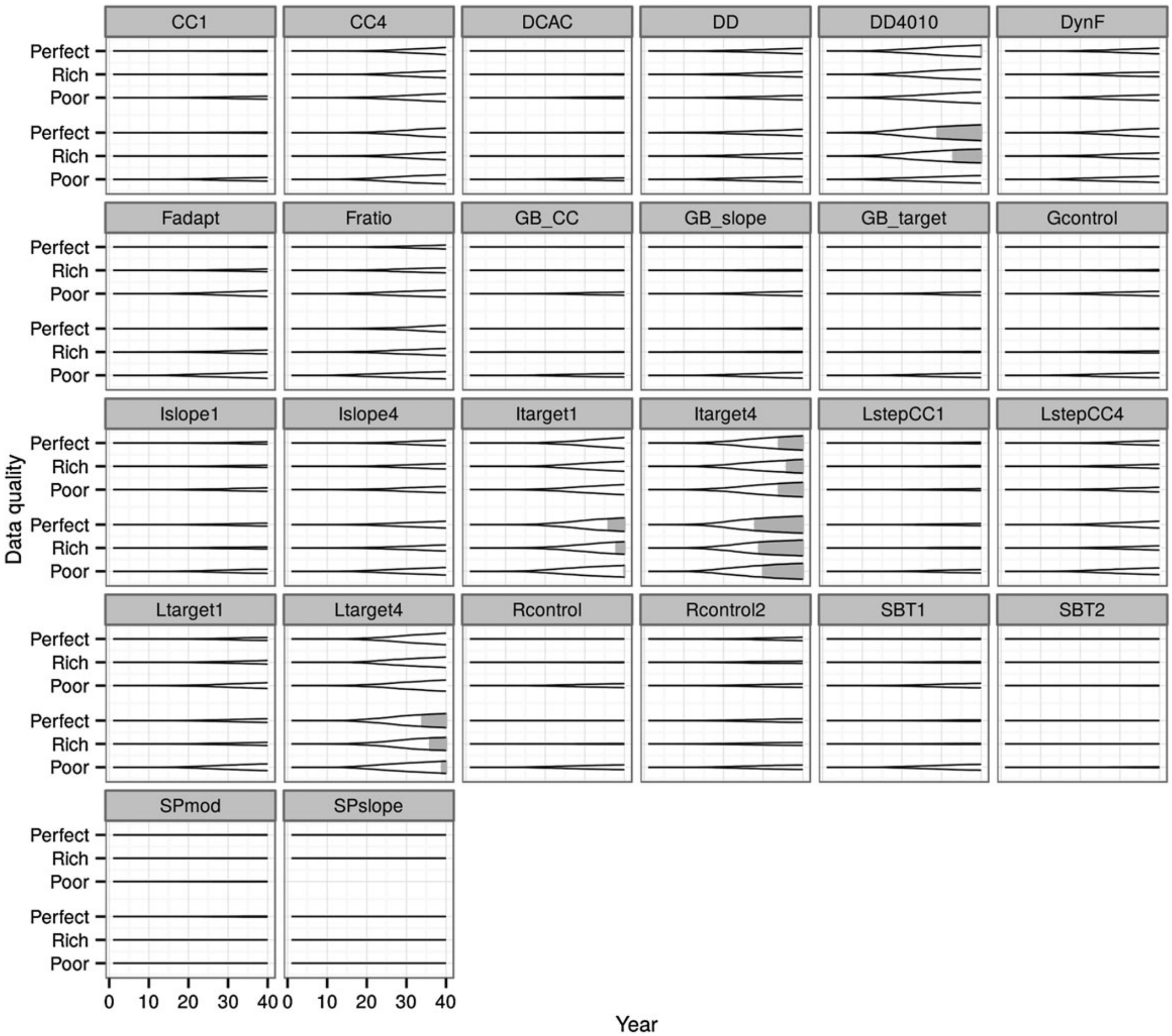

Figure 4. As Figure 2 but for the rockfish life history.

absolute biomass or stock depletion. This result confirms the findings of other simulation evaluations (e.g. Carruthers *et al.*, 2014) that these data are particularly valuable. The relative success of MPs using these inputs is perhaps surprising since reasonably large biases were simulated for indices of absolute biomass and stock depletion (in the data-limited biases as extreme as 1/5 or 5 times that of the true values were simulated). This result suggests that simple MPs may offer a relatively inexpensive approach for managing stocks based solely on fishery-independent survey data and historical catches. One example of this is the International Whaling Commission's Revised MP, which is based on historical catches and 5-yearly sighting survey observations of numbers alone (perhaps describable as a data-moderate situation), and is able to achieve reasonable performance only because the range of biases considered plausible for those inputs is not too large (IWC, 1992).

Generic MPs, which use the derivative of surplus production with respect to biomass [*G*, Equation (1)] to update the TAC, generally performed poorly (Gcontrol, Rcontrol, Rcontrol2, SPslope, SPmod, and Fadapt). The simulations reveal that these methods often fail due to bias in observations of current biomass. Since observed biomass is often an order of magnitude larger than the observed catches, even very small positive biases produce *G* values that are overly stable and fail to respond to true simulated biomass levels. Conversely, small negative biases in observed biomass lead to *G* values that are too responsive, leading to MPs that over-correct for changes in true simulated biomass. The DynF MP, which uses *G* to modify an implicit fishing mortality rate constrained within bounds, offered better performance. However, fishing at a fixed ratio of natural mortality rate (Fratio) often delivered higher yields with lower probability of dropping to low stock levels.

The simulations indicate that generic MPs can provide comparable or better performance to (i) data-rich approaches using time-series of catch and effort data and (ii) approaches that are currently used in data-limited assessment settings. In terms of data-rich settings, this conclusion is driven by the relatively good performance of the data-moderate Fratio reference MP that is applied in setting catch-limits for data-limited stock complexes in Alaska (e.g. sculpins; NPFMC, 2012). The Fratio MP often outperformed the delay-difference assessment over a range of data qualities, depletion levels, and life history types. This suggests that an

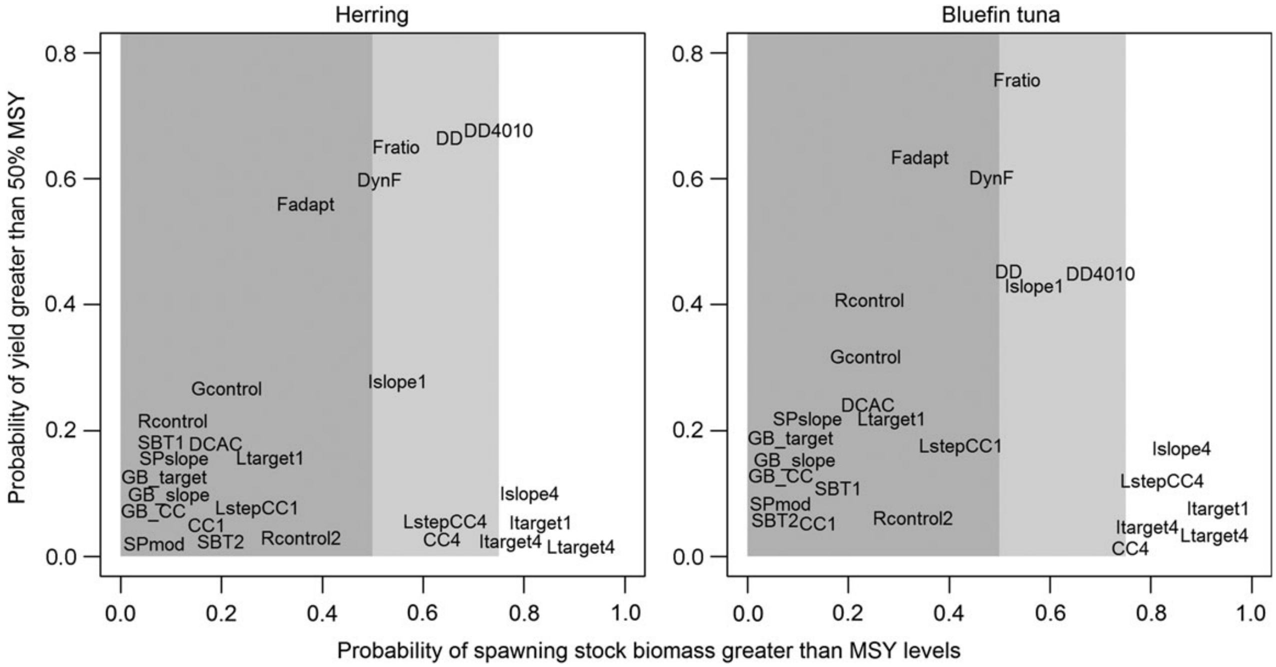

Figure 5. The trade-off between the probability of rebuilding to B_{MSY} levels and the probability of meeting a yield target of 50% *MSY*. Plotted are the results of the data-rich simulations with low autocorrelation in recruitment. Colored shading reflects the probability of rebuilding the stock to above SSB after 40 years. Dark grey regions correspond to less than a 50% probability of rebuilding, and light grey regions less than a 75% probability of rebuilding. Each panel includes a dashed horizontal line at 60% probability of achieving a yield of greater than 50% *MSY*, which corresponds to the target.

index of current absolute biomass is particularly valuable. If such data can be obtained from a fishery-independent survey, or alternatively a relative abundance index can be scaled to absolute biomass by stock assessment, acceptable performance may be obtained from fishing at a fixed fraction of natural mortality rate, which is simple and transparent. However, the simulation tests of Deroba *et al.* (2015) suggest that although stock assessment methods capture relative trends in biomass reasonably well, they provide less reliable estimates of absolute biomass.

In relation to data-limited assessment settings, we compared the MPs with DCAC which is currently used by the Pacific Fishery Management Council to set catch-limits for data-poor stocks (PFMC, 2010). As has been found previously (Carruthers *et al.*, 2014), DCAC can lead to chronic overfishing at the very low stock sizes of the rebuilding scenario (below 15% unfished levels in this analysis). In these circumstances, Islope4 and LstepCC1 MPs often outperformed DCAC by a substantial margin. At more modest levels of stock depletion, MPs such as Islope1 provided comparable performance to DCAC, indicating that simple MPs could be applied more widely (where an index of relative abundance is available), particularly as interim approaches while additional data become available. However, it should be noted that, in some respects, our simulations constitute an unfair evaluation of DCAC which was designed primarily as an interim approach to setting catch-limits for relatively long-lived stocks (natural mortality rates <20%).

It is necessary to underline the importance of operating model and observation model specifications in determining the relative performance of the MPs. The accuracy and bias in simulated data were specified using expert judgement and were consistent with values used for previous analyses (Carruthers *et al.*, 2014). Perhaps, the quality of certain data is overstated, therefore favouring certain classes of MP. This remains a central weakness of simulation evaluations such as ours. It is, however, possible to identify those observation processes that are most important in determining performance for each MP and therefore to highlight those observation processes that are most likely to produce an unfair comparison.

Robustness tests revealed that both positive and negative biases in catch observations can lead to the largest declines in yield among all the observation processes that were simulated. Additionally, virtually all the MPs tested are likely to lead to chronic stock declines if annual catch data are positively biased by >50%. This suggests that when reconstructing of historical time-series (e.g. Zeller *et al.*, 2007, 2015) of catches, caution should be exercised to avoid bias in general and in particular not to overestimate these data.

An ongoing problem in the development, testing, and adoption of MPs is that they are typically established using specific simulations that are often difficult to reproduce. In a new fisheries management setting, it is therefore difficult to evaluate a wide range of MPs comparatively and to select an appropriate MP. In an attempt to address this issue, we identified a reference set of simulations using software and code that are freely available. New MPs may be tested within the same framework, and results could be published that are directly comparable to our results. An additional benefit is that our analysis may be modified to suit particular requirements. For example in a data-limited setting in which new data are to be collected, managers may seek MPs for use over a shorter interim period, and evaluate performance over fewer projected years. Since the simulation data are reproducible exactly, readers could also frame the results using performance metrics appropriate to their particular management framework.

While reference MPs, such as Fratio, appear relatively unaffected by stock status, this cannot be said for tuned MPs that were generally more sensitive to particular depletion levels and life history types. For example, the SBT2 rule that otherwise performed relatively

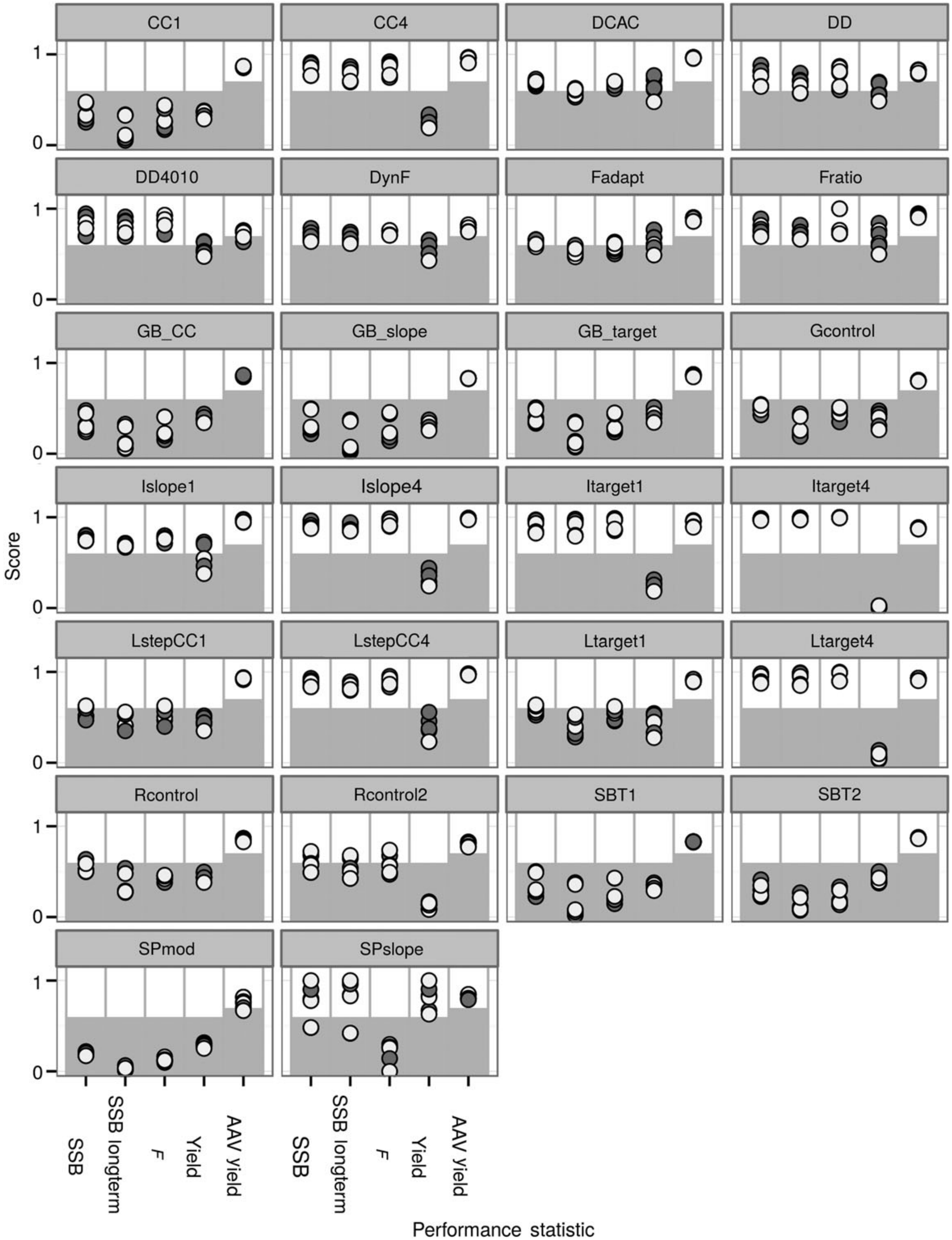

Figure 6. Choice plots (Kell, 2015) summarizing the performance of MPs given the herring life history. Each point represents performance of the MP for a given operating model (data quality and level of recruitment autocorrelation). The score is the frequency of simulations for which a performance goal was achieved. For example an SSB point at 0.91 indicates that 91% of simulations succeeded in keeping SSB above the target level of 50% of *MSY* levels. White areas represent regions that exceed the target performance level. White and grey points represent simulations with low and high recruitment autocorrelation, respectively.

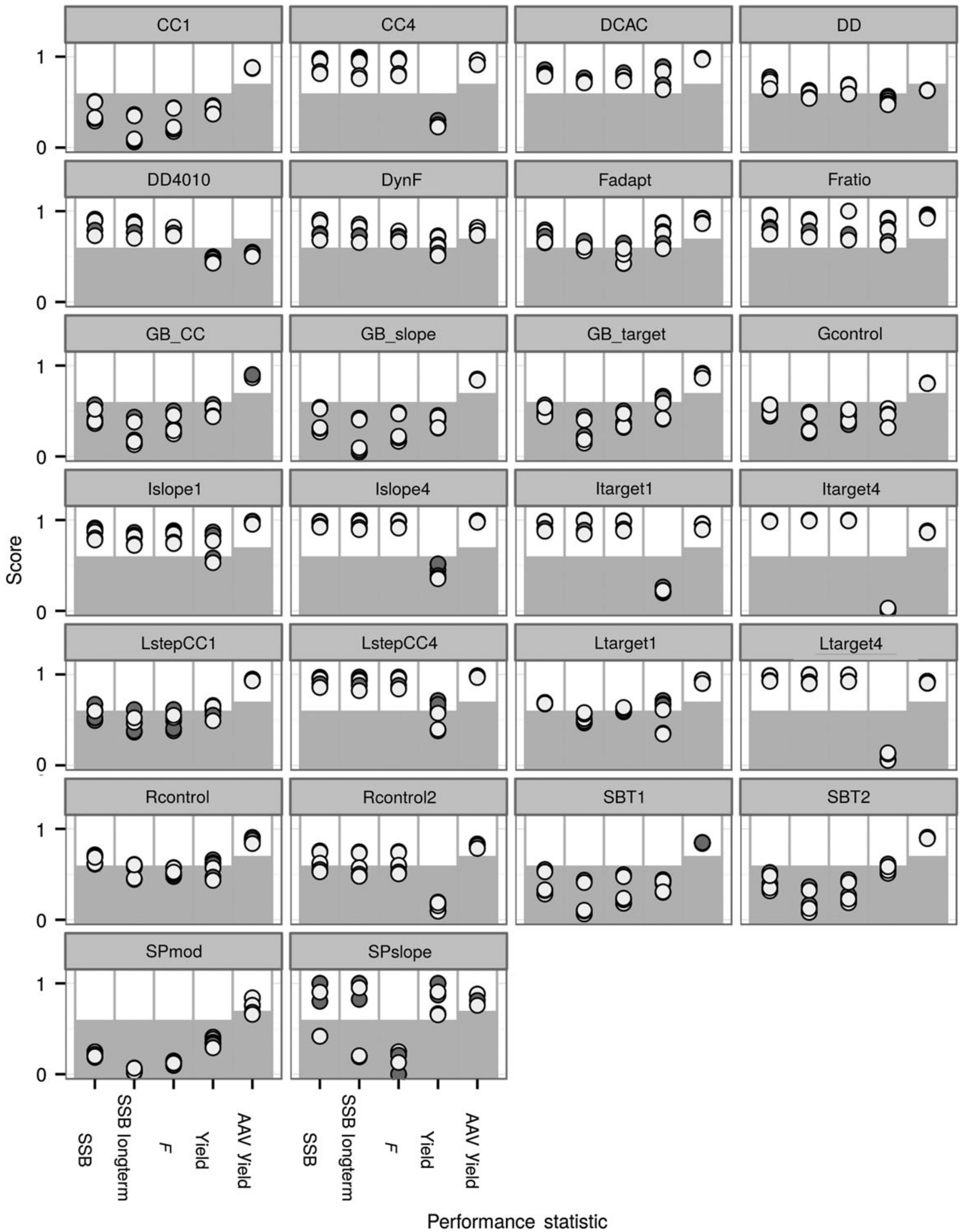

Figure 7. As Figure 6 but for the bluefin tuna life history.

poorly could perform reasonably well for the rockfish life history and the overexploited scenario (between 15 and 35% unfished). The trajectory in fishing effort was an important determinant of the performance of the tuned MPs, further confirming the need to re-tune such MPs to status and exploitation history on a stock-by-stock basis.

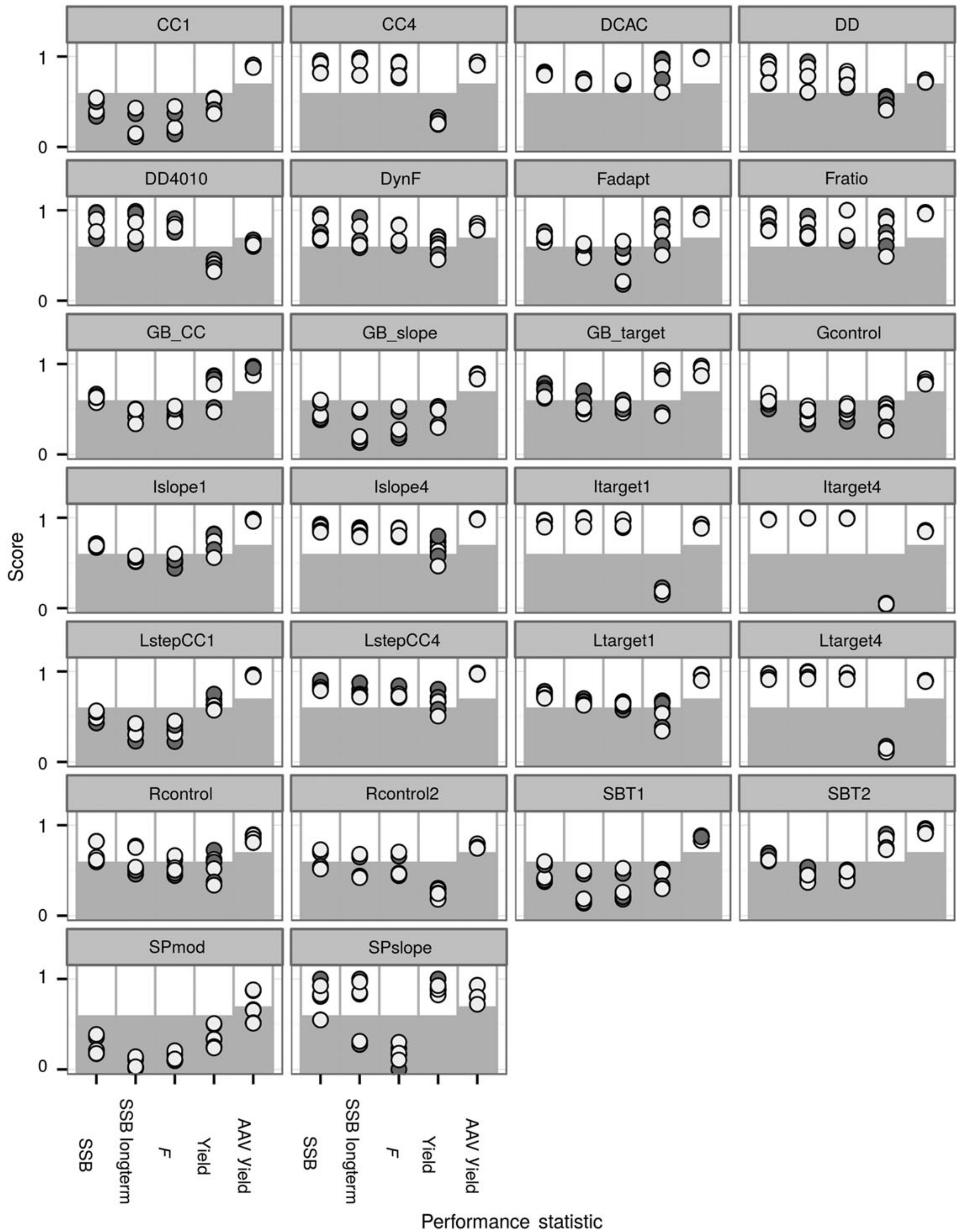

Figure 8. As Figure 6 but for the rockfish life history.

In a data-rich setting, it is clear that tuning an MP to simulations may have large benefits in terms of performance as demonstrated by the variable performance of MPs, such as SBT2, across life history types and depletion levels. Another important advantage of tuning is to standardize, to some degree, the broad management objective-level performance of a suite of candidate MPs. For example, the SBT MP as adopted (CCSBT, 2011; a more complex combination of SBT1 and SBT2 from this study) was one of a suite of candidates—biomass dynamic, empirical, and model-based—and all were tuned to meet the rebuilding objectives for the CCSBT's reference set of OMs. First, this ensures a base-level performance measure: any MP must achieve the objectives for the most probable scenarios. Second, because all the MPs have the same reference performance benchmark, "fair" comparisons for alternative robustness scenarios can be performed.

In data-poor settings, it may not be clear how to specify a suitable operating model since formally, depletion is unknown in these cases. It may be necessary to simulate a wide range of current stock depletion and provide suitable diagnostics of sensitivity in performance. Additionally, it might make more sense in data-poor scenarios to be broader in scope for both the nature of the tuning and the performance criteria. For example in data-poor settings, *MSY* may be hard to quantify although it is a policy objective. Simpler criteria may be more appropriate, such as requiring MPs for (suspected to be depleted) stocks to be able to increase stock biomass (on average) without unduly decreasing catches; or focusing more on *status quo* scenarios where, for example, the simpler MP is a temporary management measure until better data are collected to allow future, more detailed MP or assessment work.

In this simulation evaluation, all the MPs provide management recommendations that are assumed to be implemented perfectly up to a maximum instantaneous fishing mortality rate of 90%. Clearly, scientific TAC recommendations are rarely followed exactly due to a range of management considerations and fishery dynamics. For example, managers may deliberately apply a degree of TAC inertia to prevent sudden declines in catch-limits, TACs may not be fully taken at low stock sizes because it is no longer profitable to fish, or alternatively there may be TAC overages due to insufficient enforcement. Considerations such as these may strongly affect the trade-offs presented in this study, and it must be emphasized that future simulation evaluations attempt to tackle these issues. Some hypotheses can be proposed. Reduction of fishing effort as catch rates decline is likely to reduce the frequency of large reductions in biomass leading to more comparable performance among MPs. Enforcement is likely to vary for alternative management regimes such as catch-limits, size-limits, gear restrictions, effort controls, and spatial closures. Constructing credible enforcement models may be challenging (Coelho *et al.*, 2013), but could strongly alter MP selection if, for example, there was a higher propensity for violation of catch-limits than gear restrictions.

Catch-limits have relatively high information requirements and are therefore most appropriate in data-rich settings. However, many national fishery management organizations are now expected to provide catch-limits for all fisheries in a fishery management plan except some short-lived stocks (e.g. United States, Australia, and New Zealand). For this reason, we have focused in this study on MPs that provide catch-limits. However, input controls, such as gear restrictions and fishing effort, may be expected to provide superior performance to output controls such as catch-limits in a wide range of fishery scenarios (Walters and Martell, 2004) and may be particularly appropriate in data-limited settings. Future simulation testing should be extended to include MPs that generate input recommendations.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Acknowledgements

The development of DLMtool was supported by the Natural Resources Defense Council and the Kingfisher Foundation. Many thanks to Alec MacCall and Clay Porch for their help with the DLMtool initiative. This work was carried out under the ICCAT Atlantic-Wide Research Programme for Bluefin Tuna (GBYP), funded by the European Union, by several ICCAT CPCs, the ICCAT Secretariat, and by other entities (see http://www.iccat.int/GBYP/en/Budget.htm). The content of this study does not necessarily reflect the point of view of ICCAT or of the other funders, who carry no responsibility for it; neither does it anticipate the Commission's future policy in this area.

References

Butterworth, D. S. 2007. Why a management procedure approach? Some positives and negatives. ICES Journal of Marine Science, 64: 613–617.

Butterworth, D. S., and Geromont, H. F. 2001. Evaluation of a class of possible simple interim management procedures for the Namibian Hake fishery. South African Journal of Marine Science, 23: 357–374.

Butterworth, D. S., and Punt, A. E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985–998.

Carruthers, T. R. 2015. DLMtool: Data-Limited Methods Toolkit. v1.35. http://cran.r-project.org/web/packages/DLMtool/index.html (last accessed 5 April 2015).

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

CCSBT. 2011. Report of the Sixteenth Meeting of the Scientific Committee. Commission for the Conservation of Southern Bluefin Tuna, Bali, Indonesia. http://www.ccsbt.org/userfiles/file/docs_english/meetings/meeting_reports/ccsbt_18/report_of_SC16.pdf (last accessed 6 May 2015).

Cochrane, K. L., Butterworth, D. S., De Oliveira, J. A. A., and 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. Reviews in Fish Biology and Fisheries, 8: 177–214.

Coelho, M., Filipe, J. A., and Ferreira, M. A. M. 2013. Modelling enforcement and compliance in fisheries: A survey. International Journal of Latest Trends in Finance and Economic Sciences, 2: 464–469.

Deriso, R. B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Canadian Journal of Fisheries and Aquatic Sciences, 37: 268–282.

Deroba, J. J., Butterworth, D. S., Methot, R. D., Jr, De Oliveira, J., Fernandez, C., Nielsen, A., Cadrin, S. X., *et al.* 2015. Simulation testing the robustness of stock assessment models to error: Some results from the ICES Strategic Initiative on Stock Assessment Methods. ICES Journal of Marine Science, 72: 19–30.

DFO. 2012. Stock Assessment Report on Pacific Herring in British Columbia in 2012. Fisheries and Oceans Canada. Science Advisory Report 2012/062. http://www.dfo-mpo.gc.ca/csas-sccs/Publications/SAR-AS/2012/2012_062-eng.pdf (last accessed 30 September 2014).

Dick, E. J., and MacCall, A. D. 2011. Depletion-based stock reduction analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research, 110: 331–341.

Fromentin, J-M., Bonhommeau, S., Arrizabalaga, H., and Kell, L. T. 2014. The spectre of uncertainty in management of exploited fish stocks: The illustrative case of Atlantic bluefin tuna. Marine Policy, 47: 8–14.

Geromont, H. F., and Butterworth, D. S. 2015a. Complex assessments or simple management procedures for efficient fisheries management: A comparative study. ICES Journal of Marine Science, 72: 262–274.

Geromont, H. F., and Butterworth, D. S. 2015b. Generic management procedures for data-poor fisheries; forecasting with few data. ICES Journal of Marine Science, 72: 251–261.

Gulland, J. A. 1971. The Fish Resources of the Ocean. Fishing News Books, West Byfleet, UK.

Hilborn, R. 2003. The state of the art in stock assessment: Where we are and where we are going. Scientia Marina, 67: 15–20.

ICCAT. 2012. Report of the 2012 Atlantic Bluefin Tuna Stock Assessment Session. International Commission for the Conservation of Atlantic Tunas. SCI-033/2012. http://www.iccat.int/Documents/Meetings/Docs/2012_BFT_ASSESS.pdf.

IWC. 1992. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on Management Procedures. Report of the International Whaling Commission, 42: 87–136.

Kell, L. T. 2015. The quantification and presentation of risk. *In* Fisheries Management Science: An Introduction to Simulation Based Methods. Ed. by C. T. T. Edwards, and D. J. Dankel. Earthscan, Routledge, Oxford, UK.

MacCall, A. D. 2009. Depletion-corrected average catch: A simple formula for estimating sustainable yields in data-poor situations. ICES Journal of Marine Science, 66: 2267–2271.

Maunder, M. N. 2014. Management Strategy Evaluation (MSE) Implementation in Stock Synthesis: Application to Pacific Bluefin Tuna. Fifth Meeting of the Scientific Advisory Committee. Inter-American Tropical Tuna Commission. SAC-05-10b. http://www.iattc.org/Meetings/Meetings2014/MAYSAC/PDFs/SAC-05-10b-Management-Strategy-Evaluation.pdf (last accessed 3 October 2014).

McAllister, M. K., Pikitch, E. K., and Babcock, E. A. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Canadian Journal of Fisheries and Aquatic Sciences, 58: 1871–1890.

NPFMC. 2012. Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands Management Area. North Pacific Fishery Management Council. Anchorage, Alaska. http://www.npfmc.org/wp-content/PDFdocuments/fmp/BSAI/BSAIfmp.pdf (last accessed 6 May 2015).

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.

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., and Donovan, G. P. 2007. Developing management procedures that are robust to uncertainty: Lessons from the International Whaling Commission. ICES Journal of Marine Science, 64: 603–612.

Punt, A. E., Smith, D. C., and Smith, A. D. M. 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: The "Robin Hood" approach. ICES Journal of Marine Science, 68: 972–981.

Rademeyer, R. A., Butterworth, D. S., and Plagányi, É. E. 2008. A history of recent bases for management and the development of a species-combined Operational Management Procedure for the South African hake. African Journal of Marine Science, 30: 291–310.

R Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Rockmann, C., Ulrich, C., Dreyer, M., Bell, E., Borodzicz, E., Haapasaari, P., Hauge, K. H., *et al.* 2012. The added value of participatory modelling in fisheries management - what have we learnt? Marine Policy, 36: 1072–1085.

Sainsbury, K. J. 2008. Best Practice Reference Points for Australian Fisheries. Australian Fisheries Management Authority, Canberra, Australia. 156 pp.

Schnute, J. 1985. A general theory for analysis or catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences, 42: 414–429.

Smith, A. D. M., Smith, D. C., Haddon, M., Knuckey, I. A., Sainsbury, H. J., and Sloan, S. R. 2013. Implementing harvest strategies in Australia: 5 years on. ICES Journal of Marine Science, 71: 195–203.

Vert-pre, K. A., Amoroso, R. O., Jensen, O. P., and Hilborn, R. 2013. Frequency and intensity of productivity regime shifts in marine fish stocks. Proceeding of the National Academy of Sciences, USA, 110: 1779–1784.

Wallace, J. R., and Cope, J. M. 2011. Status Update of the US Canary Rockfish Resource in 2011. Report of the National Marine Fisheries Service. http://www.pcouncil.org/wp-content/uploads/Canary_2011_Assessment_Update.pdf (last accessed 3 October 2014).

Walters, C., and Martell, S. J. D. 2002. Stock assessment needs for sustainable fisheries management. Bulletin of Marine Science, 70: 629–638.

Walters, C., and Martel, S. J. D. 2004. Fisheries Ecology and Management. Princeton University Press, Princeton, NJ.

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

Zeller, D., Booth, S., Davis, G., and Pauly, D. 2007. Re-estimation of small-scale fishery catches for US flag-associated island areas in the western Pacific: the last 50 years. Fishery Bulletin US, 105: 266–277.

Zeller, D., Harper, S., Zylich, K., and Pauly, D. 2015. Synthesis of underreported small-scale fisheries catch in Pacific-island waters. Coral Reefs, 34: 25–39.

Appendix

Table A1. Summary of the bias/error parameters and related distributions that control the accuracy and precision of knowledge of the simulated system that is subsequently used by the data-limited methods and harvest control rules.

Variable	Symbol	Related functions
The standard deviation of the lognormally distributed bias in natural mortality rate M (μ_M varies among simulations)	Υ_M	$M_{\mathrm{obs}} = M \times \mu_M$ $\mu_M \sim$ lognormal($\mu = 1, \Upsilon_M$)
The standard deviation of the lognormally distributed bias in von Bertalanffy growth rate parameter K (μ_K varies among simulations)	Υ_K	$K_{\mathrm{obs}} = K \times \mu_K$ $\mu_K \sim$ lognormal($\mu = 1, \Upsilon_K$)
The standard deviation of the lognormally distributed bias in biomass at maximum sustainble yield B_{MSY} ($\mu_{B_{MSY}}$ varies among simulations)	$\Upsilon_{B_{MSY}}$	$B_{MSY_{\mathrm{obs}}} = B_{MSY} \times \mu_{B_{MSY}}$ $\mu_{B_{MSY}} \sim$ lognormal($\mu = 1, \Upsilon_{B_{MSY}}$)
The standard deviation of the lognormally distributed bias in biomass at maximum sustainable yield relative to unfished B_{peak} (B_{MSY}/B_0, $\mu_{B_{\mathrm{peak}}}$ varies among simulations)	$\Upsilon_{B_{\mathrm{peak}}}$	$B_{\mathrm{peak}_{\mathrm{obs}}} = B_{\mathrm{peak}} \times \mu_{B_{\mathrm{peak}}}$ $\mu_{B_{\mathrm{peak}}} \sim$ lognormal($\mu = 1, \Upsilon_{B_{\mathrm{peak}}}$)
The standard deviation of the lognormally distributed bias in the ratio of maximum sustainable fishing mortality rate to natural mortality rate F_{MSY}_M ($\mu_{F_{MSY}\_M}$ varies among simulations)	$\Upsilon_{F_{MSY}-M}$	$F_{MSY}\_M_{\mathrm{obs}} = F_{MSY}\_M \times \mu_{F_{MSY}\_M}$ $\mu_{F_{MSY}\_M} \sim$ lognormal($\mu = 1, \Upsilon_{F_{MSY}\_M}$)
The standard deviation of the lognormally distributed bias in MSY(μ_{MSY} varies among simulations)	Υ_{MSY}	$MSY_{\mathrm{obs}} = MSY \times \mu_{MSY}$ $\mu_{MSY} \sim$ lognormal($\mu = 1, \Upsilon_{MSY}$)
The standard deviation of the lognormally distributed bias in the age at first maturity Am (μ_{Am} varies among simulations)	Υ_{Am}	$\mathrm{Am}_{\mathrm{obs}} = \mathrm{Am} \times \mu_{\mathrm{Am}}$ $\mu_{\mathrm{Am}} \sim$ lognormal($\mu = 1, \Upsilon_{\mathrm{Am}}$)
Uniformly distributed observation error in recruitment (R_{obs}, varies among years and simulations, $\sigma_{R_{\mathrm{obs}}}$ varies among simulations)	$\sigma_{R_{\mathrm{obs}}}$	R_{obs} = lognormal($\mu = R$, $\sigma_{R_{\mathrm{obs}}}$) $\sigma_{R_{\mathrm{obs}}} \sim U(L_{R_{\mathrm{obs}}}, U_{R_{\mathrm{obs}}})$
The standard deviation of the lognormally distributed bias in the current level of stock depletion D ($B/B0$; D_{obs}, and j_D vary among projected years and simulations; μ_D and σ_D vary among simulations)	Υ_D	$D_{\mathrm{obs}} = D \times j_D$ $j_D \sim$ lognormal(μ_D,σ_D) $\mu_D \sim$ lognormal($\mu = 1, \Upsilon_D$)
Uniformly distributed observation error in current stock depletion μ_D for projected years	σ_D	$\sigma_D \sim U(L_D,U_D)$
The standard deviation of the lognormally distributed bias in catches C (C_{obs} and Υ_C vary among projected years and simulations; μ_C and σ_C vary among simulations)	Υ_C	$C_{\mathrm{obs}} = C \times \Upsilon_C$ $\Upsilon_C \sim$ lognormal(μ_C,σ_C)
Uniformly distributed observation error in catches	σC	$\mu_C \sim$ lognormal($\mu = 1, \Upsilon_C$) $\sigma_C \sim U(L_C,U_C)$
Standard deviation in lognormal error in the relative abundance index for projected years (I and Υ_I vary among years and simulations, σ_I varies among simulations)	σ_I	$I = B^{\beta} \times \Upsilon_I$ $\Upsilon_I \sim$ lognormal(1, σ_I)
The beta parameter controlling hyperstability/hyperdepletion in the abundance index (β varies among simulations)	B	$\sigma_I \sim U(L_I,U_I)$ $LN(\beta) \sim U(LN(\beta_{\min}),LN(\beta_{\max}))$
Log-uniform bias in current biomass (B_{obs} and j_B vary among years and simulations, μ_B and σ_B vary among simulations)	$\min_B$ $\max_B$	$B_{\mathrm{obs}} = B \times j_B$, $j_B \sim$ lognormal(μ_B,σ_B) $\log(\mu_B) \sim U(\min_B,\max_B)$
The maximum standard deviation for lognormal error in current biomass for projected years	σ_B	$\sigma_{B_{\mathrm{cur}}} \sim U(L_B, U_B)$

The lognormal distribution described in the table below where ~lognormal(μ,σ) is the exponent of the normal distribution with mean μ and standard deviation σ, parameters: dnorm($-0.5\log(1+\sigma^2/\mu^2)$, $\sqrt{\log(1+\sigma^2/\mu^2)}$).

Handling editor: Ernesto Jardim